

To what degree are Croatian and Serbian the same language?
Evidence from a Translation Study*

John Frederick Bailyn
Stony Brook University

1. Introduction

The purpose of this article is to argue, on the basis of the comparison of purely grammatical *building-blocks*, exposed through a particular kind of experimental study, that the (standard) languages referred to as Croatian and Serbian are the same language, and should be acknowledged and treated as such.¹ (I assume further, that if this point can be demonstrated for Croatian and Serbian, the argument extends directly to what are now being called Bosnian and Montenegrin.) A central assumption behind this claim, and behind the entire article, is that the question of linguistic identity *can* be discussed within the realm of *structural/formal* linguistics and need not be (only) defined with regard to culture, ethnicity, nationality or politics, despite common sociolinguistic assumptions to the contrary, such as that of Langston & Peti-Stantić (2003), who state that “it is not possible to define what constitutes a language as opposed to a variant or dialect in terms of the inherent features of the language variety itself.” (L&P-S 2003, p. 249). The claim of this article is that it is possible to do so, by comparison of the grammatical building-blocks of the language variants in question, whose similarity (or in this case near-identity) can be exposed by experimental methods.² This results will show that Croatian and Serbian are for all intents and purposes linguistically identical.

This view contradicts much common lore about these two languages. There are two common narratives surrounding the historical and current status of Croatian and Serbian, both of

*This paper is dedicated to my teachers of then Serbo-Croatian, Wayles Browne and Milorad Radovanović, with the explicit caveat that they are in no way responsible for the contents here, except insofar as they both inspired me to study and continue studying BCS and to visit then Yugoslavia in 1989 and current Serbia, Bosnia and Croatia many times since. I am grateful to participants at the Language & Human Rights Conference at Stony Brook University’s Humanities Institute in 2006, as well as to Steven Franks and several anonymous reviewers for invaluable discussion. The article would not have been possible without Danijela Lugarić and her students from the University of Zagreb who sparked the original debate in St. Petersburg in 2006 that inspired the study. Thanks to all the Zagreb students who participated in the translation study. Thanks to Ivana Mitrović and Dijana Jelača for help with the original texts, and to Dijana Jelača for much needed moral support. All mistakes are my own.

¹By ‘standard’ I mean only the contemporary language norms in the two countries as opposed to Croatian and Serbian being defined as “anything spoken by Croats and Serbs respectively”. I do not mean a ‘literary standard’ (see discussion below), nor do I adopt Crystal’s 1997 definition of a standard language, namely, “a particular variety of a language that has prestige within a speech community” (emphasis by an anonymous reveiwer). Because standardization of variants has led to claims of linguistic identity and then to claims of linguistic distinctness, contemporary language norms are as good a place as any to examine the purported differences. If we find linguistic identity in these variants, the point will have been made more generally – namely that cultural/ethnic/national identity is not isomorphic with linguistic uniqueness.

²I will use the term “variants” as a neutral designation for the linguistic systems that are generally referred to as Croatian, Serbian and so on. I will sometimes use the term “Western variant” for what is popularly known as (standard) “Croatian” and “Eastern variant” for is known as (standard) “Serbian”, by way of acknowledging that those distinctions among the variants that exist do not correlate either with national groupings or with political structure. See Alexander 2006 for discussion.

which are based on non-linguistic criteria in defining the relevant languages as distinct. One story is that the formerly unified language Serbo-Croatian disintegrated into “successor languages” in the 1990s, as the Serbo-Croatian speaking areas of Yugoslavia broke apart into the countries of Croatia, Bosnia & Herzegovina, and Serbia & Montenegro, (and now include Montenegro (and Kosovo)). The “successor languages” -- Croatian and Serbian (as well as Bosnian/Bosniak and now Montenegrin)) then acquired their rightful status as independent languages once their nations had gained independent status (Greenberg 2004, Alexander 2006). The other story is that two distinct languages existed before their joint standardization in the 19th and 20th century as Serbo-Croatian. Until the 1990s, they were artificially combined into the language of Serbo-Croatian, and after Yugoslav disintegration were able to regain their former status as independent languages (Langston & Peti-Stantić 2003).

A common view of the recent emergence of successor languages is revealed by a reviewer of this article who states that “what is clear is that as of 1991-2 Serbo-Croatian officially ceased to exist in the Yugoslav successor states. All sides agreed that the unified language was to be jettisoned and probably never again to be resurrected.” However, examination in this article of the *grammatical building blocks* of the variants, exposed by experimental methods, determine that ‘official’ declarations and the desire to ‘jettison’ a unified language do not make its demise a fact, as the same reviewer reveals in the statement that “there are no scholars who claim that Croatian and Serbian are two distinct languages from a formal linguistic point of view.” It is this formal linguistic point of view that is emphasized in the Translation Study I report on in this article.³

It is important to note that while it may be obviously true for some linguists that Croatian and Serbian are the same language, (see, for example Kordić 2005, 2006, 2008), it is also the case that for many linguists, the opposite appears obvious. Thus Langston & Peti-Stantić (2003) state that “on the basis of both the historical development and the current political realities, *there can be no doubt* that they *should* be treated as separate languages.” (L&P-S 2003, p. 249, emphasis mine). This article is intended both to inject “doubt” (back) into the discussion, as well as to question the notion that ‘modality’ (“X *should* be treated as Y”) is appropriate in discussions of the identity of linguistic systems. Naturally, much of this depends on the definitions involved, a point to which I return immediately below.

The issue is very emotional for many people -- it is associated with the complex ethnic, social and political history of the former Yugoslavia and its violent break-up in the 1990s, its previous legacy as a ‘unified’ state, complex language standardization processes (Peti-Stantić 2008), and the cultural and national identities of the successor states, which are naturally intimately intertwined with linguistic identity. I propose stepping away from the emotional and socio-political issues as much as possible and looking at the linguistic systems themselves. The translation study reported here allows us to do exactly that.

³The issue of the *name* of this single language is far from trivial. I will use “BCS” (standing for Bosnian/Croatian/Serbian), simply because other possible English names for the common language (especially “Serbo-Croatian”) have complex historical and political connotations that cloud the debate and raise many non-linguistic issues. As Alexander (2006) discusses, the use of “BCS” is convenient, easy to pronounce, as politically neutral as possible, the 3 letters are arranged in alphabetical order, and it does not force one to take a stand on the complex issue of “Bosnian” vs. “Bosniak” (despite the term BCS not being used (or even known) within the BCS speaking areas, where the names *hrvatski*, *sprski* and *bosanski/bošnjački*, in addition to *naš jezik* (‘our language’) are most common.) See Alexander 2006, chs. 22-26 for important discussion.

Linguists and others speaking out in favor of the identity of the two variants have generally relied on two points of commonality:

1) Two traditional criteria in favor of the claim of a common language:

- a. mutual intelligibility
- b. sharing of a common (literary) standard

Mutual intelligibility, of course, is something that is naturally subjective, and also depends on factors such as exposure to common media sources, political unity or disunity, and so on. That it should not be relied upon as decisive in the case of BCS is true both for those who believe that (standard) Croatian and (standard) Serbian are distinct languages, and for those who believe they are essentially the same. Thus Langston & Peti-Stantić (2003) note that “although mutual intelligibility intuitively would seem to offer an unambiguous criterion for determining whether two speech varieties represent the same or different languages, intelligibility is actually a relative and to some extent subjective feature; studies have shown that speakers’ evaluations of intelligibility may be subject to social and political pressures” (L&P-S 2003, p. 275, fn. 6). And, as Alexander (2006) puts it, “this criterion [mutual intelligibility] is also imprecise, and largely dependent not only on perception but also on emotion (and, to an extent, on the educational level of the speakers)... objective linguists may decide that two speech systems clearly represent the same language; yet if speakers of these two systems are sufficiently convinced by external factors that they will not be able to understand each other, then that will usually turn out to be the case, and the speakers in question will claim that they are speaking different languages.” (Alexander 2006, p. 401)

Mutual intelligibility arguments also fail to be definitive because they are impossible to quantify. Surveys have been done about *impressions* of mutual intelligibility, but it is not possible to quantify the effect in a way that could tease apart intelligibility due to sharing a single linguistic system compared to mutual intelligibility due to years of sharing media and other cultural sources of discourse. Thus the argument of mutual intelligibility is typically convincing only to those who believe independently that there is a single language at issue, but unconvincing to those who feel otherwise. That being said, there is no doubt of the near 100% mutual intelligibility of (standard) Croatian and (standard) Serbian, as is obvious from the ability of all groups to enjoy each others’ films, TV and sports broadcasts, newspapers, rock lyrics etc.⁴

The issue of sharing a common (literary) standard has also been used to motivate positing a single Serbo-Croatian language (see discussion in Greenberg 2004) (as well as to motivate positing distinct languages now). The existence of a common set of literary documents at various historical periods solidified the perceived unity of Serbo-Croatian. The counter-argument here has been that insofar as there ever was a common Serbo-Croatian language, it was *artificially created*, and therefore in some sense never truly existed. This argument could be used to distinguish the BCS case from many language variants that fail the mutual intelligibility test (Arabic variants, North Italian variants, Chinese variants, and so on), but which are nevertheless often regarded as one language because of the generally-accepted existence of a common literary

⁴In this regard, it is often pointed out that Czech and Slovak also show a high rate of mutual intelligibility, but are linguistically distinct languages (as a parallel translation study would quickly reveal). Thus the *fact* of mutual intelligibility is difficult to raise to the level of *linguistic evidence* in favor of linguistic identity, especially because of the confounding factor that if the two were distinct languages, political and cultural reality could easily explain why all speakers of one understand all speakers of the other so easily.

standard. In the case of post-Yugoslav Croatian or Serbian (or Bosnian or Montenegrin), this objection is at times coupled with the claim that the Serbo-Croatian literary standard was based on one of the groups' local variant. Thus the claimed existence of a Serbo-Croatian literary standard as motivation for the claim of a single language now, for example, could be said to prolong the Serbianization of the language(s) under the Novi Sad Agreement of 1954 (and perhaps earlier, depending on the version of the claim), although as Greenberg (2004) shows, this claim distorts the historical situation considerably.⁵

In fact, the initial unification movements of the 19th century under Vuk Karadžić for the Serbs and Ljudevit Gaj for the Croats were attempts to unify dialects that cut across ethnic and religious lines into a common language. Compromises were made as to which dialect to choose as the source of a common standard. The *štokavian ijekavian* dialect took on this role, with the result that most BCS speakers speak a modern version of the *štokavian* dialect. (See Greenberg 2004 and Alexander 2006 for descriptive histories.)⁶ As for the current dialectal situation, “neither the older dialect divisions into Štokavski vs. Čakavski vs. Kajkavski nor the later subdivision into Ekavski vs. Ijekavski vs. Ikavski correspond geographically to the major religious, cultural and political boundaries.” (Browne & Alt 2004, p. 9)

What I want to do in this article, then, is simply to provide *grammatical* evidence that BCS remains a single language, based in part on structural comparison of the internal linguistic levels of Phonology, Morphology and Syntax (or linguistic ‘building-blocks’), but primarily on the results of a Translation Study from “Serbian” into “Croatian”, which casts strong doubt on the idea that Croatian and Serbian are distinct languages. The *lexical* distinctions that do exist come nowhere near to reaching the level of distinguishing the two languages, especially given the essentially identical nature of the entire grammatical system.

The structure of the article is as follows: First, I present the Single Language Hypothesis, to be used in what follows. In Part 2, I describe a pilot Translation Study conducted in 2008. In Part 3, I present results from the Translation Study, supplemented by comparative description, that strongly support the identity claim for BCS. In Part 4, I discuss the important issue of lexical differences, which is often used to support the claim of distinct languages, and show that even in this aspect of language, the Single Language Hypothesis is supported.

⁵The claim of the artificiality of the earlier common standard is a strange argument in the sense that all (literary) standards are to a certain degree artificial or arbitrary – language is in a constant state of flux and change, and dialectal and regional variation is a reality of all language communities. Crucially, there is no biological or natural basis for any literary language, whereas what generative linguists refer to as “I-language” is considered to be part of the natural world, and can therefore be studied using methodology generally accepted in the natural sciences (Chomsky 1981, 1995, 2005). The I-language view would take every speaker to have a distinct internalized linguistic system (an *idiolect*), determined by the complex mapping of linguistic data available during acquisition onto innate structures. “Sharing the same language” reduces to degree of overlap of the resulting idiolects. I do not adopt the extreme generativist position that no two grammars are ever the same, but I do assume an essential distinction between the internalized linguistic ‘knowledge’ of native speakers, a natural phenomenon, and literary standards, which are cultural creations, albeit important ones.

⁶ As Bugarski (2004) points out, the 19th century unification was not without inconsistencies at that time: “Although in essence a single system structurally, it still displayed non-negligible differences in script and orthography, in points of pronunciation and grammar, and especially in lexicon.” The latter points are well-known; the former point, of it already being “a single system structurally”, is the historical basis for the *natural* unity of the current BCS system that this article exposes through a Translation study.

(2) provides the definition of linguistic identity I will work with:

2) **The Single Language Hypothesis: (SLH)**

The degree to which 2 language varieties can be considered the same linguistic system correlates with the degree to which their building-blocks are the same, that is, the degree to which their internal linguistic systems employ the same grammatical components.

Grammatical components relevant to the SLH in the case of BCS are given in (3):

3) Grammatical Components relevant to the SLH:

- a. Phonology: the same phonemic inventory
- b. Morphology
 - i. Derivational Morphology: identical derivational morphological devices for the same kinds of derivations
 - ii. Inflectional Morphology:
 - identical distinctly represented morphological categories (case, number, gender for nominals; person, number and tense for verbal categories, etc.)
 - identical form of the actual bits of inflectional morphological for the same inflectional categories
- c. Syntax: identical settings of major syntactic parameters
- d. Lexicon -- degree of identity in:
 - i. lexical categories (N, V, Adj and their combinatory requirements – case, selection, etc)
 - ii. functional/grammatical categories (P, C, D, Neg, T, adverb,...)
 - iii. functional/grammatical combinations (PP temporal and special modifiers; verbal government, selection, etc.)

In what follows I will show that the Single Language Hypothesis is strongly upheld for BCS.

To test the Single Language Hypothesis for BCS, a Translation Study was undertaken in the Summer of 2008 (see below). Results of the Translation Study corroborate the descriptive fact that there are practically no significant differences between Croatian and Serbian with regard to any aspect of (3)a (phonology), (3)b (morphology), and (3)c (syntax). As for (3)d (the lexicon), we will see that although Croatian and Serbian are well-known to differ to a certain degree with regard to the first part of (3)di, that is, there are certainly quite a few instances of distinct lexical terms for identical concepts or notions, the % of such variation is relatively small, as we will see from the Translation Study. More importantly, we will also see that the second part of (3)di, namely what requirements particular lexical items make on the elements they combine with, are practically identical. Finally, *and crucially for the maintenance of the Single Language Hypothesis*, we will see that the other major component of the lexicon, namely the functional/grammatical markers of the language (3)dii, are for all intents and purposes identical, as are their required combinatorics (3)diii. Given the near-identical nature of the entire internal linguistic system, it could be argued that the relevance of the lexical items that do differ between Croatian and Serbian for mutual understanding is reduced considerably, since the grammatical frames in which the differing lexical items find themselves more than compensate for the lexical distinctions themselves, in ways that are absent when we are in fact dealing with distinct linguistic systems, even those as similar as Czech and Slovak.

2. The Translation Study

In order to test the extent to which Croatian and Serbian have distinct linguistic systems, a pilot Translation Study was conducted in 2008 with 16 adult native speakers of the Croatian variant of BCS. The speakers were all from the Croatian capital Zagreb, and 15 of them ranged in age from 20-38, (with the 16th being a 55-year-old).⁷ All speakers except one were born and educated entirely in Croatia; all completed high school in Croatia and attended or currently attend the University of Zagreb. The group consisted of 11 women and 5 men. The subjects were asked to “translate” 9 short texts from Serbian into Croatian. The texts themselves were taken from various registers and sources, in the Eastern *Ekavian* variant spoken in Serbia. The texts were checked with adult Serbian speakers from Novi Sad to verify their authenticity.⁸

4) List of texts used:

<u>number</u>	<u>theme</u>	<u>source</u>	<u>length</u>
1	local educational news item	internet	129 words
2	personal monologue	blog	176 words
3	literary text	Milovan Glisic, <i>Redak Zver</i>	166 words
4	baggage instructions	airline website	106 words
5	technology instructions	e-mail server	87 words
6	political news (Zimbabwe)	internet	90 words
7	recipe	cooking website	95 words
8	a story (priča)	unknown	147 words
9	sports news	internet	68 words
		total:	1,064 words

The Croatian subjects were given the texts (in Latin alphabet) and asked to translate them into Croatian as closely as possible, without avoiding any opportunity to show a distinction between the original and the translation. The distinctions were emphasized in the directions to the subjects deliberately, so that those results showing identical or near-identical forms could be reliably considered indicators of true identity. Results will be presented for each area of language as we move forward through the various linguistic levels.⁹

⁷ A reviewer asks about the relatively small size of the group. Naturally, more data, and a bi-directional study of translations would reveal more. However, this study still involved over 17,000 words (across all participants) across a range of registers.

⁸ A reviewer observes that the texts do not cover scientific or government/political styles, which (s)he claims would show more differences. However, a survey of the website of the government of Bosnia & Herzegovina (<http://www.fbihvlada.gov.ba>), which is presented in 3 varieties, shows a remarkable convergence in all areas of language, including lexicon. Clearly a larger study, covering a wider range of texts, would be of interest in determining which styles of the different varieties show more or less lexical distinctions, especially if it included translations in both directions.

⁹ A systematic survey of attitudes toward the issue of Croatian and Serbian language was not conducted with these speakers. However the younger speakers (9 of the 16 are in their 20s) all came into the study determined that it would support their strong conviction of Croatian and Serbian as distinct languages.

3. One Linguistic System

An examination of all the major ‘levels’ of language show that BCS is clearly a single language with a single grammatical system. In this section I briefly review the relevant levels and discuss any apparent distinctions between the purported different languages, referring to the Translation Study where relevant to provide empirical support.

3.1. Phonology

The Eastern and Western variants of BCS have the identical set of phonemic *distinctions*. The vowel system consists of the following phonemes:

5) BCS Vowel Phonemes (all variants): /a/, /e/, /i/, /o/, /u/

The consonant system consists of the following 25 phonemes:

6) BCS Consonant Phonemes (all variants): /r/, /v/, /j/, /l/, /lj/, /m/, /n/, /nj/, /f/, /s/, /z/, /š/, /ž/, /h/, /dj/, /dž/ /c/, /č/, /ć/, /b/, /p/, /d/, /t/, /g/, /k/

This is not to say that the variants do not show some systematic distinctions. As is well-known, the three major *štokavian* sub-dialects – *Ekavian*, *Ikavian* and *Ijekavian*, are so-named for how the older Slavic vowel *jat* /ě/ developed in the modern version of the variants. The Croatian subjects in the Translation Study all replaced Serbian /e/ stemming from historical /ě/ with the appropriate combination, either /-ije/ or /je/. This is shown in (7):¹⁰

7) Translation Study: cognate words in Text 1, showing Serbian /e/ → Croatian /ije/ or /je/

<u>Text</u>	<u>Original</u>	<u>Translation</u>	<u># of occurrences</u>	<u>Comments</u>
1	primenjenih	primijenjenih	64 of 64	4 instances in text 1
1	umetnosti / umetnika	umjetnosti umjetnika	48 of 48	3 instances in text 1
1	obeležavanja	obilježavanja	32 of 32	two changes here
1	najuspešnijim	najuspješnijim	16 of 16	
1	odseka	odsjeka	16 of 16	

totals: 176 of 176

Crucially, these systematic distinctions do not introduce any phonemes that are not present in the other variants. The only relevant phonemes in these variant forms are /e/, /i/ and /j/, all of which exist independently in all variants. /ě/ itself has been lost in all modern BCS variants, replaced

¹⁰(7) only shows the results for Text 1. I do not provide the statistics for the same phenomenon from the other texts. It is clear from the absolute 100% rendition of these instances of Serbian /e/ into Croatian /je/ or /ije/, that the correspondence is systematic. In what follows, I will also use samples of the 9 texts to make the relevant points rather than providing the data from all of them on every point of comparison.

either by /e/, /i/, or a combination of /j/ and the two remaining vowels.¹¹ Furthermore, the modern variation described above does not correlate with “Serbian” and “Croatian” in a direct way. Serbian spoken in Serbia itself is mostly, but not entirely *Ekavian*.¹²

It is true, however, that there are a few other systematic phonological distinctions between Western and Eastern variants with regard to other phonological issues beyond the modern-day representation of the older *jat* phoneme. There is partly systematic variation between western /h/ and Eastern /v/ in cases such as the following:

8) Eastern /v/ vs. Western /h/ in some lexical items:

English	Eastern	Western
tobacco	duvan	duhan
to cook	kuvati	kuhati
dry	suvo	suho
deaf	gluvo	gluho

Thus in Text 2 and Text 7, the masculine adjective *suvi* ‘dry’, appear, in the context *suvi informacija* (‘dry pieces of information’) in Text 2 and *suvi vrat* (‘dry pork’) in Text 7. All 16 Croatian speakers translated this as *suhih and suhi*, just as above. However, Text 2 also contains 24 other instances of the phoneme /v/, and Text 7 contains 15 other instances of /v/, all of which are rendered as /v/ by all of the Croatian subjects. The same goes for /h/, which is found 13 times in the Text 2 and twice in Text 7 and remains as /h/ in the translations in all cases where the same lexical item is used, except for 1 – the lone instance being the word for ‘chemistry’ -- *hemija* in the Serbian original, translated as *kemija* by all 16 Croatian speakers. Therefore we must conclude that cases such as *suvi~suhi* (and *hemija~kemija*) are lexically specified, and do not represent any *phonemic* distinction with regard either to Eastern /v/ or Western /h/.

One more note about phonology. Browne & Alt 2004 contains 90 pages of linguistic description of the various levels of BCS, covering a wide range of complex phonological processes. In none of these cases, (with the minor exception of the ability of Western vs. Eastern speakers to discern tonal distinctions that are no longer a distinctive part of the phonological system), are any dialectal issues even mentioned (because there are none of relevance). A short example of the complexity and uniformity of one such situation should suffice:

The BCS "old jotation" resulting from the Proto-Slavic jotation is: (1) labials add lj, thus p-plj, b-blj, m-mlj, v-vlj ; the newer sound f also becomes flj. (2) s, z alternate with š, ž. (3) t, d alternate with ć, đ. (4) k, g, h alternate with č, ž, š ; as in the first palatalization's reflex, c has also come to alternate with č. (5) l, n alternate with lj, nj. (6) r and other consonants (palatals of various sorts, also the group št) are unaffected. (Browne & Alt 2004, p. 18)

¹¹Greenberg (2004) and Browne & Alt (2004) identify one dialect in the West of Serbia that appears to have a vowel between /i/ and /e/ as a development of former /ě/, in addition to /i/ and /e/. However this potentially real phonemic distinction does not constitute a difference between standard Croatian and standard Serbian. The 2 variants have identical phonemic systems.

¹²Bosnian and Croatian Serbs primarily speak *Ijekavian*, and attempts by the nationalist Bosnian Serb leadership during the 1990s to impose *Ekavian* on the local Serbs did not succeed, and was abandoned as an official project in 1998 (Bugarski 2004, Greenberg 2004).

It is well-known that Slavic “jotation” systems are extremely complex. The paragraph quoted here succinctly summarizes what BCS has inherited from Proto-Slavic in one such instance. (No other Slavic language has anything like the same set of realizations, due to differences in the phonological histories). Half a dozen developments of this sort are described in detail in Browne & Alt 2004. In any instances where there are dialectal distinctions, they are described as well. However no dialectal variations are mentioned here, *because there are none*. Clearly, if Croatian and Serbian had distinct developments in areas of the phonology such as these, we might be able to support the claim that the phonological systems were distinct to some degree. However, this is not the case – the phonological systems are essentially identical in this and all other similar regards. This strongly supports the Single Language Hypothesis.

3.2. Morphology

3.2.1. Derivational Morphology

Derivational morphology is practically identical in the Eastern and Western variants of BCS. A good example of this identity is found in the derivation of perfective/imperfective verb pairs. As in other Slavic languages, there exists a complex set of possible morphological relations between perfective and imperfective verb pairs. In some cases, it is a matter of simple prefixation, where the perfective form surfaces with a prefix that is absent in the imperfective form (*pisati~napisati* ‘to write’). Another device is the use of suppletive stems in imperfective and perfective pairs such as *nalaziti~naći* (“to find”). Finally, there are derivational morphological devices such as infixation/suffixation, as in *kupovati~kupiti* (“to buy”) and vowel alternations, such as *zatvarati~zatvoriti* (“to close”). (See Browne & Alt 2004 for exact descriptions.)

If the Western and Eastern variants had truly diverged to the point of being distinct linguistic systems, one might expect each to have developed distinct ways of deriving (at least some of their) imperfective/perfective pairs. Such a distinct development would indeed cause a significant ripple effect across the derivational morphology of the language, including affecting nominalizations and other areas of vocabulary expansion, and in its interaction with the lexicon, it is not difficult to imagine this leading rather quickly to the advent of truly distinct verbal systems, and eventually even to a lack of full mutual intelligibility. After all, although the various Slavic languages share tendencies in their system of imperfective derivation, they are all quite different systems, and each requires a unique grammatical description. One such description suffices for all of the BCS variants exactly *because they follow identical rules of derivational morphology*. If the behavior of imperfective derivation were in any way exceptional in this regard we could perhaps maintain a case for distinct linguistic systems. But in area after area of the derivational morphology, we find the same level of identity and no need to even discuss regional variation, for the most part.

Nominalization is another area of derivational morphology where there is no indication of any difference at all in how the process works among any of the BCS variants – all nominalize verbs productively (far more productively than Russian does, for example), using the suffix *-nje* (*pisanje* ‘writing’ from *pisati* ‘to write’). Croatian and Serbian are identical in this regard. No changes in this regard are found in the translations.

As with the phonological systems, there are some minor areas of semi-productive distinctions in the derivational morphology systems. For example, in verbs based on foreign borrowings, Western variants tends to use the formative suffix */-irati/*, whereas Eastern variants

use /-ovati/. At the same time, the final form below shows that even this pattern is not productive. This is shown in (9):

9) Semi-productive distinctions in formatives with borrowed roots:

English	Bosnian	Croatian	Serbian
to organize	organizirati organizovati	organizirati	organizovati
to construct	konstruisati ?konstruirati	konstruirati	konstruisati
But:			
to analyze	analizirati	analizirati	analizirati

Some of these distinctions come up in the Translation Study, Thus in Text 2, 15 of the 16 Croatian speakers render the original form *reprodukujemo* ('we reproduce') from *reprodukovati* as *reproducirati* ('to reproduce').¹³ (The 16th speaker provided a distinct lexical item – *ponoviti* 'to repeat'). This essentially conforms to what is shown in (9) above, with the minor caveat that the foreign root ending /k/ is transformed by the Croatian speakers into /c/ before the /ir/ suffix, in keeping with standard velar palatalization rules. However, it should be noted that this distinction relates only to borrowed words that surface as verbs, and does not interfere with other derivational processes (all such verbs can nominalize as discussed above), or with verbal conjugation, or other inflectional processes. Thus we can conclude that the system of derivational morphology supports the Single Language Hypothesis.

3.2.2. Inflectional Morphology

In what follows, I will discuss the practically identical nature of every instance of inflectional morphology across the BCS variants. In discussing inflectional morphology, it is useful to first categorize the inflectional *distinctions* represented by the language (a potential source of variation), and then the actual morphophonemic manifestations of the various inflectional morphemes. As we will see, there is almost no variation here of any kind, and this is central to the Single Language Hypothesis.¹⁴

3.2.2.1. Morphological Categories

Browne & Alt (2004) present the overall BCS system as follows:

¹³Note that the conjugated vs. infinitival forms here are the result of the verb being in a subordinate clause, which typically show conjugated form vs. infinitival variation. See section on syntax, below.

¹⁴A reviewer expresses surprise that this discussion need to be included in the article at all, because "the dialectal basis [of the two variants] is the same." However, this same reviewer maintains the position that the two variants correspond to distinct *languages*. The paradoxical view emerges from that review that the dialect is the same but the languages differ. Clearly, terminology is at issue here. This reviewer is in agreement with the basic claims of this article about linguistic 'building blocks', while at the same time relying on a purely *sociological* definition of language, in the absence of a purely linguistic definition. We therefore appear only to differ on the definition of 'language' for which I rely on the Single Language Hypothesis as opposed to extra-linguistic criteria. Under other definitions, of course, other conclusions are possible.

BCS distinguishes masculine, neuter and feminine genders in singular and plural; the 2/3/4 form opposes masculine-neuter to feminine.... There are three main sets of case-and-number endings or declension types. One has -o, -e or zero in the nominative singular and -a in the genitive singular. It includes most masculine and all neuter nouns. A second has nominative singular - a, genitive -e. It contains most feminine nouns and small classes of masculines. The third type ends in zero in nominative singular, -i in genitive. It includes all feminines apart from a-stems. (p. 28)

Needless to say, this statement holds in its entirety for both Croatian and Serbian. A full list of distinctions in inflectional morphology made (in all BCS variants) is given in (10):

10) Categories expressed by BCS inflection:

a. Nominal:

- i. Case: Nom, Gen, Dat, Acc, Instr, Loc, Voc
- ii. Number: Sg, Paucal, Pl
- iii. Gender: M, F, N
- iv. Class

b. Adjectival

- i. Attributive: Concord with modified N for Nominal categories listed in (10)ai-iii
- ii. Short vs. Long form morphology in Masc. Sg. Nom A's to express definiteness
- iii. Predicate adjectives: Number (Sg. Paucal, Pl); Gender (M,F,N)

c. Verbal:

- i. Person: 1st, 2nd, 3rd
- ii. Number: Sg, (Paucal), Pl
- iii. Tense: Pres, Past, Future (compound form), Aorist, (Imperfect)
- iv. Aspect: Pf, Impf
- v. Other: Imperative, participial, subjunctive (morphologically identical with past tense)

Practically none of this inventory differs among variants.¹⁵ The Translation Study shows remarkable uniformity of morphological form between the two variants. Case endings are the same, predicate agreement is the same, concord is the same, verbal endings are the same, and so on. In verbal categories, there were sporadic instances of changed prefixes, *naliti* ('to pour') changing to *doliti* ('to pour') for 2 of the 16 Croatian speakers in both instances in Text 7, and for another 2 speakers in one instance but not the other. The vast majority of verbal prefixes remained unchanged.¹⁶

¹⁵ It is true, as Greenberg (2004) points out, that there are some southern Serbian dialects that are losing case distinctions, presumably under the influence of Macedonian and Bulgarian, and as such represent transitional variants to those distinct languages. Of course, this occurs naturally in many contact areas, and does not provide any support to the notion that the Eastern and Western variants themselves contain any significant difference on this score.

¹⁶One minor point of variation concerns the 3rd sg copular form *jeljest*. As Browne & Alt (2004) point out: "the 3rd person singular is *jest* in the Croatian standard, *jeste* in Serbian, both in Bosnian, but all standards use the expression *to jest* 'that is, i.e.' In asking a question with *li*, the 3rd person singular is *Je li.*" (Browne & Alt 2004, p. 49). However, it should be noted that all dialects use *jeste* in short answers and emphatic contexts, and all use *je* in reduced and unstressed contexts.

The only instances of a minor change in morphological sub-category of nominals between the Serbian originals and Croatian translations involved nouns that show a declension class switch, which in most cases also involved a change in gender. (Note that all the feminine nouns shown here are declension Class I, and all the masculines are Class II, except *vezista* in the original text which is a Class I masculine, a small but well-attested category.)

11) All instances of declension class and/or gender switch (5 words total (over all 9 texts) out of 269 nouns, 4,304 tokens):

<u>Text #</u>	<u>English</u>	<u>Original</u>	<u>Translation</u>	<u># of tokens</u>	<u>rate of change</u>
1	criterion	<i>kriterijum</i> (m)	<i>kriterija</i> (f)	16 of 16	100%
1	asst professor	<i>docent</i> (m)	<i>docentica</i> (f)	2 of 16	12.5%
3	half	<i>po</i> (m)	<i>pola</i> (f)	30 of 32	93.75%
5	luggage	<i>prtljag</i> (m)	<i>prtljaga</i> (f)	80 of 80	100%
9	signalman	<i>vezista</i> (Class I)	<i>vezist</i> (Class II)	8 of 16	50%
				total: (these 5 items)	136 of 160 85%
				total: (all items)	136 of 4,304 3.2%

Thus we see systematic gender variation in only 3 nouns out of 269 and partial variation in 2 others. The vast majority remain identical in the translations.

3.2.2.2. Inflectional Morphophonemics

It is clearly the case that Croatian and Serbian share the same set of morphological category *distinctions*. But are there any significant divergences in what *forms* are used to represent the different categories, that is, is there evidence that, for example, the Genitive case forms for Croatian speakers are different from those for Serbian speakers? Here, we turn again to the Translation Study. I base the results presented here about inflectional morphology on the translations of Text #3.

12) Inflectional morphemes found in Text #3:

	<u>types</u>	<u>occurrences</u> (in text 3)	<u>tokens</u> (in translations)	<u>differences</u> (# tokens)	<u>convergence</u> <u>rate</u> (%)
Nominal endings	27	52	832	6 (1 type)	99.28%
Verbal endings	12	31	497	0	100%
Total	39	83	1,329	6 (1 type)	99.55%

The rate of convergence is over 99%. In fact, there is only one instance in Text 3 of any change in the output of inflectional morphology: There are two instances of the masculine Accusative animate adjective form *-og* in the original text. This occurs after the preposition *za* ('for') which takes Accusative. The original of Text 3 has the phrases *za jednog* ('for one [Englishman]') and *za drugog* ('for another [Englishman]'). 2 of the 16 translations change *-og* to *-oga* in the first instance and 4 of them change it to *-oga* in the second instance. Thus the total number of changes from *-og* to *-oga* is 6 out of 32. However, it is not clear at all that this represents even a distinction between variants so much as a register distinction. The two forms are considered acceptable in all variants to such a degree that standard textbooks list the endings as *-og(a)*. The

same holds for other adjectival endings that are polysyllabic, such as dative *-om(e)* and plural *-im(a)*. No other instances of morpho-phonemic change are found in any other nominal forms, and no adjectival or verbal forms show any variant in inflectional morphology of any kind. The SLH is strongly supported in this most crucial of building-blocks, the productive inflectional morphology.

The significance of this overwhelming identity in inflectional morphophonemics cannot be overemphasized. Croatian and Serbian (and Bosnian and Montenegrin) share 99% of morphological categories expressed overtly and 99% of the specific morphophonemic forms. Under conditions of 99%+ identity of morphological devices and forms, a much higher level of lexical variation is possible than otherwise. The grammatical frame and context in which lexical items find themselves is identical. Culturally-carried meanings of words may vary by national culture, but the internalized computational system that allow us to put them together into larger meanings is the same. Lexical distinctions of up to 25 or 30% of the entire non-grammatical lexicon might be expected to be easily tolerated in this situation and the Single Language Hypothesis would still lead us to take seriously the idea that the variants in question are a single language. In the BCS case things are much clearer than that, for as we will see, the lexical distinctions make up at most 7% of the open-class lexicon and under 1% of the functional/grammatical (closed-class) lexicon.

3.3. Syntax

With regard to the major parameters by which languages differ from each other syntactically (Chomsky 1981), all variants of BCS share identical settings in practically all instances. The one well-known syntactic difference between Eastern and Western variants (*da* + present tense verbs (Eastern) vs. Infinitival constructions (Western)) will be discussed below after an overview of essential parameter settings. In this section, I discuss some aspects of syntax that are not reflected in the texts simply to exemplify the unified system of BCS syntax.

3.3.1. Parameter settings

In the following chart, I show the BCS parameter setting for 7 major aspects of syntax (for discussion of these and many other syntactic issues, see Rudin 1988, Franks 1995, Bošković 2001, 2002, Browne & Alt 2004, Alexander 2006, Bailyn 2007 and references therein). In none of these areas, do we find any difference between the Western and Eastern variants.

13) Some major parameter settings for BCS:

<u>Parameter</u>	<u>Croatian</u>	<u>Serbian</u>	<u>variation?</u>
Multiple WH-fronting	+	+	none
Superiority in multiple WH	-	-	none
Pro-drop (dropping of pronoun subjects)	+	+	none
<i>svoj</i> required for 3 rd person subjects	+	+	none
Subject condition on anaphor binding	+	+	none
Availability of long-distance antecedent	+	+	none
Clitics in second position	+	+	none

In fact, there are only two areas of syntax where any variation was found in the Translation Study – one is well-known -- *da* clauses vs. infinitives, and the other less known – the choice of pronominal vs. reflexive possessive adjectives.¹⁷ Let us begin with the latter:

3.3.2 Binding differences

The translations show various instances of a second person possessive pronoun *tvoj/tvoja/tvoji...* or *vaš/vaša/vaši* ('your') being replaced in the translation by the reflexive possessive pronoun *svoj/svoja/svoji* ('self's'). Examples are given in (14):

14) 2nd person possessive pronouns vs. reflexive possessives:

a. Text 2 (original):

Jedanput kada to shvatite na dobrom ste putu da ostvarite vaše snove.
Once when that catch on good aux-2pl path that realize **your** dreams
“Once you understand that, you are on the way to realize your dreams.”

b. translations: ...da ostvarite **vaše** snove (non-reflexive: 9 of 16)
that realize **your** dreams

... da ostvarite **svoje** snove (reflexive: 7 of 16)
that realize **self's** dreams

15) 2nd person possessive pronouns vs reflexive possessives:

a. Text 5 (original):

možete ostati bez **vaših** privatnih poruka
can-2pl remain without **your** private messages
“you could end up without your private letters”

b. translations: ... bez **vaših** privatnih poruka (non-reflexive: 8 of 15)
without **your** private messages

... bez **svojih** privatnih poruka (reflexive: 7 of 15)
without **self's** private messages

16) 1st person (plural) possessive pronoun vs reflexive possessive:

a. Text 2 (original):

...da nas osposobe da realizujemo **naše** snove
that us help that realize **our** dreams
“to help us realize our dreams.”

¹⁷Alexander (2006) gives examples the placement of second position clitics varying between Croatian and Serbian speakers. Such variation was not found in the Translation Study, but this issue requires further experimental studies specifically designed with constructions of this kind in mind.

b. translations:	...da	realizujemo/ ostvarimo/	naše	snove (non-reflexive: 9 of 13) ¹⁸
	that	realize	our	dreams
	... da	realizujemo/-iramo ostvarimo/	svoje	snove (reflexive: 4 of 16)
	that	realize	self's	dreams

Thus it seems that there is a slight preference for the reflexive possessive pronoun to replace the personal reflexive pronoun in about half of the possible cases. However, it remains unclear to what degree this is a matter of *construction choice* (see discussion below on lexical variation) or a true syntactic difference between the Eastern and Western variants. If it can be shown that there is real complementary distribution here, that is that the Eastern (original) variants disallow the use of *svoj* in the constructions showing *vaš* and that the Western (translated) versions *only* allow *svoj*, then we might be dealing with a true syntactic difference. In 3 texts, the Eastern variant original shows 5 instances of the reflexive pronoun *svoj* – twice in Text 1, twice in Text 4 and once in Text 8 – and these instances of *svoj* are translated as *svoj* in 100% of cases. This result cuts both ways. That is, on the one hand it clearly shows that the Eastern variant uses the reflexive possessive regularly, and that it would be too strong to say it is an element only used or favored by the Western variant. On the other hand, the fact that there are 0% of changes in the direction of reflexive → non-reflexive, whereas we have a fairly high percentage of changes from non-reflexive → reflexive may or may not represent some sort of syntactic difference between the two variants. As a reviewer points out, a follow-up study with translations from Croatian into Serbian might illuminate this issue.

3.3.3 Infinitives and *da*-clauses

The best-known syntactic distinction between Eastern and Western variants of BCS concerns the tendency in the East to prefer tensed complement clauses with *da* and in the West to prefer infinitives, whenever possible. Typical examples are given in (17):

17) Variation in “control clauses” (from Text 2)

a. <u>Eastern</u>		b. <u>Western</u>	(15 of 16)
zašto... treba da znam		zašto... trebam znati	
why need-imps that know-1sg		why need-1sg know- infin	
“why...I need to know”		“why... I need to know”	
c. <u>Eastern</u>		d. <u>Western</u>	(5 of 16)
preferiraju da nas zatrpa ju...		preferiraju nas zatrpa ti	
prefer that us overwhelm- 3pl		prefer-3pl us overwhelm- Infin	
“[they prefer] that they overwhelm us”		“[they prefer] to overwhelm us”	

¹⁸3 of the 16 translations employed non-verbal constructions in this second clause that exclude the possibility of reflexive because of the lack of a subject antecedent. I therefore did not count these cases as instance of non-reflexive, since no reflexive option is grammatically available here.

Interestingly, only 5 of the 16 speakers transformed the clause in (17)c into an infinitival clause of the kind shown in (17)d, despite the complete availability of the infinitival alternate. This indicates that we are most likely dealing with a matter of construction *choice* rather than a real parametric syntactic distinction. Serbian speakers are perfectly comfortable with both variants, and it appears from the high numbers of Croatian translations that maintained the *da*+ tensed verb clauses in the translations, that their grammars also allow both constructions. That the Western variant does not *disallow* the *da* + tensed V clauses, despite the common lore to the opposite, is seen in cases where for independent reasons the infinitival clause is unavailable.

18) Cases where infinitival replacement is not possible: (from Text 2)

a. Original (Eastern)

da nas osposobe da realizujemo naše snove
that us make possible that realize-1pl our dreams
“that make it possible for us to realize our dreams”

b. Translation (Western)

da nas osposobe da realiziramo naše snove
that us make possible that realize-1pl our dreams
“that make it possible for us to realize our dreams”
(14 out of 16 instances x 2 *da* clauses)

(There is a standard *-ovati ~ -irati* distinction in the verb *to realize*, see above.) Both *da* + conjugated verb clauses are maintained in the translation, because there is no viable (verbal) alternative. In an effort to avoid the *da* clause, some Croatian speakers paraphrased the entire construction with a nominalized verb in a prepositional phrase for the second *da* clause. Most made no changes, however (14 of 16), because of the lack of a syntactic infinitival option.

I therefore maintain that the best-known instance of an apparent syntactic distinction between variants is nothing more than case of preference for one available construction over the other. Indeed, Miskeljin (2005) states that for every Serbian *da*-clause, an infinitival variant exists that is preferred in Western variants, though both are grammatical in both. In fact, however, there is one instance of infinitival clauses that are notoriously absent in all variants of BCS, despite their full availability in other Slavic languages. These involve object control structures, such as English (19) (subject control cases are given for comparison)

19) Subject control in English:

- a. Mary wants to go.
- b. Jovan is trying to sleep.

20) Object control in English:

- a. The generals ordered Jovan to go
- b. The generals convinced Mary to stop

21) Subject control in BCS: (Eastern variant: *da* + conjugated verb)

- a. Marija hoće da **ide**
Marija wants-3sg that go-3sg
“Marija wants to go”

- b. Jovan pokušava da **spava**
Jovan tries-3sg that sleep-3sg
“She wants to go”

22) Subject control in BCS: (Western variant: Infinitive)

- a. (ona) hoće **ići**
wants-3sg to go-**Infin**
“She wants to go”

- b. Jovan pokušava **spavati**
Jovan tries-3sg sleep-**Infin**
“She wants to go”=

23) Object control in BCS: (*da* + conjugated verb only: all variants!)

- a. Naredili su mu da ide
ordered aux him that go-
“They ordered him to go.”

- b. *Naredili su mu **ići** (infinitive construction ungrammatical: all variants!)
ordered aux him go-**Infin**
“They ordered him to go.”

These examples show that there are syntactic constraints on the Western infinitival constructions that rule out the possibility of most infinitival object control sentences in BCS in any variant, despite their availability in many related languages.¹⁹ An explanation of this gap in the BCS paradigm is beyond the scope of this article. But there is strong agreement among all BCS speakers that (23)b is ungrammatical (including those that will always choose an infinitival construction over a *da* + verb construction in every instances when it is possible.) This fact show several important things about this purported syntactic distinction – first, the infinitival variants are more restricted than in most other languages with infinitives. Second, Serbian and Croatian speakers share the same judgments about this unusual restriction, a similarity far more intriguing and potentially important than the differences in construction choice that speakers make. Third, the difference between the two variants in this area cannot be reduced to a parameter, whereas the similarity (when its underlying nature is discovered) presumably can.

¹⁹ Browne & Alt (2004, p. 75) note that “in Croatian two verbs permit an infinitive to refer to their object: 'teach' and 'help'” and give the following examples:

- i) Učio sam ga plivati.
'I taught him to swim.'
- ii) Pomogli smo mu graditi kuću.
'We helped him to build a house.'

These are the exceptions that prove the rule. As B&A note, what remains “good in all standards” is:

- iii) Pomogli smo mu da gradi kuću.
'We helped him that he build a house.'

4. Lexical Differences and Similarities

In this section, I discuss the issue of the lexicon with regard to the often heard claim that this is the area where the two variants are the most distinct, and that this qualifies them for separate language status. The Translation Study results show that the percentage of such lexical variation is relatively small, even in the open-class categories. This is especially true in the closed lexical classes, or the so-called “functional categories”, where we will see that the two variants are *identical*. It is only in the open classes that we find variation, and given the absolute identity of the rest of the entire grammatical systems, these differences, many of which may simple be a matter of preference, are also insignificant. It should be noted that a far wider set of lexical distinctions can be commonly found across dialects considered the same language.

4.1 Functional (closed-class) Categories

To call the two variants ‘similar’ is to completely understate the situation with regard to the functional categories of the language. Consider the statistics from the Translation Study. The 9 original translation texts contain the following breakdown of lexical and functional elements:

24) Parts of speech in the 9 translation texts:

<u>category</u>	<u>class type</u>	<u>total # of instances</u>	<u># of tokens translated</u>	<u># of instances changed</u>	<u># of tokens changed</u>
noun	open	269	4,304	24	250
main verb	open	144	2,304	9	125
adjective	open	99	1,584	4	25
adverb	semi-open	68	1,088	0	0
preposition	closed	87	1,392	0	0
determiner	closed	62	992	1	11
auxiliary	closed	61	976	0	0
conjunction	closed	54	864	0	0
complementizer	closed	39	624	0	0
WH-phrases	closed	29	464	0	0
particle	closed	23	368	0	0
negation	closed	13	208	0	0
clitics (pron)	closed	12	192	0	0
pronouns (full)	closed	11	176	0	0
	<u>totals:</u>	971	15,536	38 (3.91%)	411(2.65%)

The open-class categories (noun, main verb and adjective) are discussed below. Of the other categories, including adverbs, the only functional category that shows any semi-systematic variation is the feminine 3rd person possessive pronoun *njen-* which is translated by the Croatian speakers in 11 of 32 instances as *njezin-*.²⁰ This represents a mere 1.61% of all determiners and

²⁰ There are scattered individual cases of changes that appear to be neither systematic, nor exclusive for the speakers at hand. Thus although the vast majority of instances (13 of 16) involving Eastern *posle* (‘after’) are rendered as *poslije*, with the standard Ijekavian change, 3 speakers translate it as *nakon* (‘after’). Similarly, in Text 8, the phrase *po ispadanju* (‘after relegation’), and *po* is also translated as

it occurs in less than half of the tokens where *njen-* appears in the original (11 of 32). Overall, then, determiners are changed 1.11% of the time. No other grammatical or functional elements show any alternations at all between the originals and the translations. That represents a total of 11 individual changes out of the total number of functional/grammatical category tokens of 7,648 (0.21%). Or, to put it the other way, functional and grammatical lexical items in the original Serbian texts remain *identical* after translation into Croatian in 99.79% of cases. Clearly, the Single Language Hypothesis is strongly supported.

4.2. Lexical Categories

The results of the Translation Study in the realm of lexical distinctions in the open class lexical categories (noun, verb and adjective) are quite important in evaluating the one possible area in which the Single Language Hypothesis is not overwhelmingly obvious. Here is the distribution of open class lexical differences brought out in the Translation Study:

25) Open class categories in the 9 translation texts:

<u>category</u>	<u>total # of instances</u>	<u>total # tokens</u>	<u># of instances changed</u>	<u>% of instances changed</u>	<u># of tokens changed</u>	<u># of tokens changed</u>
noun	269	4,304	24	8.92%	250	5.81%
main verb	144	2,304	9	6.25%	125	5.43%
adjective	99	1,584	4	4.04%	25	1.58%

Nouns show the highest percentage of lexical difference between the Serbian originals and the Croatian translation – 8.92%. This means that 8.92% of all 269 noun occurrences in the 9 texts were in some translation or other translated with an entirely lexically unrelated word. However, the fact that only 5.81% of all tokens of those nouns (across the 16 translations) were changed shows that even of the 24 nouns that showed a lexical difference, not all were systematically changed in the Croatian translations, though some were. In fact, of the 24 nouns, 9 verbs and 4 adjectives that showed non-cognate lexical distinctions, only 8 nouns, 5 verbs and 2 adjectives were treated systematically as different by a high percentage of the 16 translations (here I have included all of those for which 10 or more of the 16 translations agreed on the lexical distinction). The other lexical items that showed distinctions showed them in fewer than 10 of the translations.

It is also notable that of the 990 words that appear in the 9 texts (which constitute 15,840 word tokens, across the 16 translations) there were only two instances in which the translator said that s/he did not know a word in the original texts. That is a knowledge rate of the original

nakon by 5 of 16 speakers, though the other 11 maintain *po*. However, this is not included in the cases above because it is sporadic, and also because the form *nakon* is in standard usage in Eastern dialects as well. Naturally, we would need further research to determine whether instances of *nakon* in original Eastern texts would remain as *nakon*, showing a tendency toward replacing *posle* with *nakon*, or whether those forms might also be translated as *poslije*, which would support then idea that some speakers make conscious (or not so conscious) changes whenever two interchangeable forms are available (see Alexander 2006, chs. 22 and 26 of discussion of such tendencies). There is no evidence in the case of *posle~nakon* that the issue is even one of dialectal variation. Further research on specific tendencies with these kinds of alternations is required.

words of 15,838 out of 15,840 tokens, or 99.9%. Still, passive recognition could be possible of words from neighboring language with significant language contact. But the extremely low rate of overall lexical change, and the absolute lack of such distinctions in the closed-class vocabulary is strong evidence that the Single Language Hypothesis applies clearly in this case, more clearly than in cases of variants of a common written language whose spoken version have indeed diverged to be on the edge between language and dialect. The BCS lexicon contains practically no linguistic evidence of there being a case to be made for distinct language status.

Furthermore, as discussed in Section 3, given the inflectional nature of the language and the near 100% identity in both derivational and inflectional morphology, the minor lexical variation that exists can almost always be determined by context – modifiers are the same, idioms mostly the same, syntactic and morphological frames in which open-class words find themselves will be almost always identical. Under these circumstances, determining the meaning of an unfamiliar word is not unlike what we experience in childhood as we learn words in our native language at an amazing rate – we determine its possible range of meanings from its grammatical (and real world) context and within 2 or 3 occurrences of the word, we are able to situate it in our conceptual world. With the absolute identity of the closed-class vocabulary and inflectional morphology, even a far higher rate of lexical distinctions would not have any significant effect on separating the variants into 2 separate languages.

5. Conclusion

Since Chomsky 1957, it has become more and more clear that what we know when we know a language is a set of building-blocks and rules of combination, some universal and some language-specific, which allow us to create an infinite variety of novel sentences and utterances. The infinite possibilities our finite system gives us is one of the great sources of creativity in the human species. This view sees language as part of the natural world. Internalized language systems are the natural and unique result of the salient features of the speech to which each generation of children is exposed mapped onto the linguistic component of the mind. Internalized linguistic systems change naturally over time, and dialects evolve into distinct languages regularly, but slowly. The idea that the variants of a language could dissolve into distinct languages in a 10-year period is excluded by everything linguists have discovered about language change since the 19th century and before. Distinct literary standards are a different story, and here readers are referred to Peti-Stantić (2008) and other specialists on processes of language standardization, literary norms and so on.

On appropriate definitions of language, of course, one could (and many people have) come to the opposite conclusion about Croatian and Serbian – they could be (and have been) defined simply as the language of the Croatian people and the language of the Serbian people, respectively. The Croatian and Serbian nations and people have distinct cultural identities, so it is natural to associate each with a distinct language. The tendency to define nationhood and cultural identity through language goes back at least as far as Dante (Fishman 1999, Joseph 2006). The view that distinct peoples in distinct countries speak distinct languages, (a correlation that is often absent in the world -- many countries have multiple local languages and some languages span many countries and peoples), is often supported by reference to regional distinctions among these “languages” in pronunciation, grammar, and especially in vocabulary (lexicon). We have seen that such linguistic distinctions are essentially absent in the case of Croatian and Serbian. Thus a crucial consequence of the findings of the study reported here is a

necessary disavowal of any *exclusive* connection between cultural and linguistic identity. Of course there is a connection, as the very fabric of cultural and personal creativity is dressed in language. And distinct ethnic, cultural, national, religious groups regularly create aspects of their identity using language in unique ways that distinguish them. But that does not make it impossible that the language they use might not be used by other people as well. Since Chomsky 1957, 1965 it has been generally accepted that what we *do* with our language is not the same as its structure. Thus, nothing in this article concerning the near identity of the two linguistic systems under discussion endangers the notion of distinct cultural and ethnic identification through one's language (used for literature, law, education, government and so on), nor threatens the cultural uniqueness of any of the peoples involved. Rather, the issue at hand has been a narrow linguistic one – how similar are the linguistic systems, and is there any linguistic justification for defining them as unique. We have seen that there is not.

Of course, the Western and Eastern variants of BCS might well still develop into different languages. Usually, such divergence happens when rule-governed sound changes lead to paradigmatic shifts in the morphological system, which might in turn affect the typology of the language (e.g. V2 changing to SVO). English diverged from other Germanic languages when sound change led to the loss of significant inflectional morphology, which in turn put pressure on the case system, and led to the fixing of SVO word order and the disappearance of V2. What is instructive about such examples is that the change begins with changes to some aspect of the internal system's building blocks (in the English case, in the inflectional morphology, lost through regular sound change), and spreads to other levels of the system. In the BCS case, our translations study has shown that the internal grammatical system is where the Single Language Hypothesis is supported most strongly.

jbailyn@notes.cc.sunysb.edu

References

- Alexander, Ronelle (2006) *Bosnian, Croatian, Serbian: A grammar with sociolinguistic commentary*. Madison, WI: University of Wisconsin Press.
- Bailyn, John F. (2007) "A derivational account of microvariation in Slavic binding". Richard Compton, Magdalena Golezinska and Ulyana Savchenko, eds. *Annual Workshop on Formal Approaches to Slavic Linguistics: The Toronto Meeting*. Ann Arbor: Michigan Slavic Productions, 25-41.
- Benson, Morton (1971 and later editions) *Serbo-Croatian-English Dictionary*. Philadelphia: University of Pennsylvania Press.
- Bošković, Željko (2001) *On the nature of the Syntax-Phonology interface: Cliticization and related phenomena*. Oxford: Elsevier Science (North-Holland Linguistics Series: Linguistic Variations).
- Bošković, Željko (2002) "On Multiple WH-fronting". *Linguistic Inquiry* 33(3): 351-383.
- Browne, Wayles (1993/2002) "Serbo-Croat". Bernard Comrie and Greville Corbett, eds. *The Slavonic languages*. London: Routledge, 306-387.
- Browne, Wayles and Theresa Alt (2004) *A handbook of Bosnian, Serbian, and Croatian* SEELRC <http://seelrc.org:8080/grammar/mainframe.jsp?nLanguageID=1>

- Bugarški, Ranko (2004) "Language and boundaries in the Yugoslav context". Birgitta Busch and Helen Kelly-Holmes, eds. *Language, discourse, and borders in the Yugoslav successor states*. Clavedon, England: Multimedia Matters Press.
- Bugarški, Ranko & Celia Hawkesworth (eds) (2004) *Language in the Former Yugoslav Lands* Slavica Publishers.
- Busch, Birgitta and Helen Kelly-Holmes (eds) (2004) *Language, discourse, and borders in the Yugoslav successor states*. Clavedon, England: Multimedia Matters Press.
- Cardinaletti, A. & L. Repetti (2008) "The Phonology and Syntax of Preverbal and Postverbal Subject Clitics in Northern Italian Dialects" *Linguistic Inquiry* 39(4), 523-563.
- Chomsky, Noam (1957) *Syntactic Structures*. Berlin: Mouton de Gruyter.
- Chomsky, Noam (1965) *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Chomsky, Noam (1981) *Lectures on Government and Binding*. New York: Foris.
- Chomsky, Noam (1995) *The Minimalist program*. Cambridge, MA: The MIT Press.
- Chomsky, Noam (2005) Three Factors in language design. *Linguistic* 36(1), 1-22.
- Crystal, David (1997) *A dictionary of linguistics and phonetics*. 4th edition. Cambridge, MA: Blackwell.
- Fishman, Joseph (1999) (ed) *Handbook of Language and Ethnic Identity*. Oxford: Oxford University Press.
- Franks, Steven (1995) *Parameters of Slavic Morphosyntax*. Oxford: Oxford University Press.
- Greenberg, Robert D. (1999) "In the aftermath of Yugoslavia's collapse: The politics of language death and language birth". *International Politics* 36(2):141-158.
- Greenberg, Robert D. (2000) "Language planning in the Federal Republic of Yugoslavia: The crisis over the future of Serbian". *Slavic Review* 59(3): 625-640.
- Greenberg, Robert D. (2004) *Language and identity in the Balkans: Serbo-Croatian and its disintegration*. Oxford: Oxford University Press.
- Joseph, John E. (2006) *Language and Politics*. Edinburgh: Edinburgh University Press.
- Kačić, Miro (1997), *Croatian and Serbian: Delusions and Distortions*. Zagreb: Novi Most.
- Klajn Ivan (2001) "Neologisms in present-day Serbian". *International journal of the sociology of language* 151:89-110.
- KordiĆ, Snežana (2005) "I dalje jedan jezik" *Sarajevske sveske* 10.
- KordiĆ, Snežana (2006) "Jezik i identitet" (review of B. Busch & H Kelly-Holmes *Language, Discourse, and Borders in the Yugoslav Successor States*) *Knjevena Republika* 4.
- KordiĆ, Snežana (2008) "Jezični Nacionalizam" (review of Stjepko Težak's *Hrvatski nas (ne)podobni*), *Slovo* 4(17): 186-205.
- Langston, Keith and Anita Peti-Stantić (2003) "Attitudes towards linguistic purism in Croatia: Evaluating efforts at language reform". In Mirjana N. Dedaić and Daniel N. Nelson, eds. *At war with words*. Berlin: Mouton de Gruyter, 247-282.

- Lehiste, Ilse and Pavle Ivić (1986) *Word and Sentence Prosody in Serbocroatian*. Cambridge, MA: MIT Press.
- Miškeljin, Ivana (2005) "PRO and Control in Serbian" MA Thesis, University of Novi Sad.
- Peti-Stantić, Anita (2002) "Funcionalno preslojavanje suvremenoga standarda". Riječki Filološki Dani. Rijeka: Filozofski Fakultet, 365-373.
- Peti-Stantić, Anita (2008) *Jezik naš i/li njihov*. Zagreb: Srednja Europa.
- Radovanović, Milorad (1992) "Standard Serbo-Croatian and the theory of language planning". Ranko Bugarski and Celia Hawkesworth, eds. *Language Planning in Yugoslavia*. Columbus, OH: Slavica Publishers, 93-100.
- Radovanović, Milorad (1993). "Yugoslavia: language situation." R. E. Asher and J. M. Y. Simpson, eds. *The Encyclopedia of Language and Linguistics*. Oxford: Pergamon, 5077-5078
- Radovanović, Milorad (ed.) (1996). *Srpski jezik [na kraju veka]*. Opole: Uniwersytet Opolski Instytut Filologii Polskiej (Belgrade: Institut za srpski jezik SANU Službeni glasnik).
- Radovanović, Milorad (2000). "From Serbo-Croatian to Serbian." *Multilingua* 19.
- Radovanović, Milorad and , Randall A. Major, eds. (2001) "Serbian Sociolinguistics" *International Journal of the Sociology of Language*, 151 (1).
- Rudin, Catherine (1988) "On multiple questions and multiple WH Fronting" *Natural Language and Linguistic Theory* 6(4): 445-501.
- Skiljan, Dubravko (2004) "From Serbo-Croatian to Croatian: Croatian linguistic identity" Ranko Bugarski and Celia Hawkesworth, eds. *Language Planning in Yugoslavia*. Columbus, OH: Slavica Publishers, 67-85.

Appendix

- Original texts used:

Text 1: (News Item: Novi likovni pogledi i tonovi Autor: M. V. | 15.07.2008 - 06:00

Izložba radova studenata završne godine Fakulteta primenjenih umetnosti „Diploma 2008“ biće otvorena večeras (19) u Muzeju primenjene umetnosti, a trajaće do 15. avgusta.

U godini obeležavanja 60 godina postojanja FPU, 94 mlada umetnika predstaviće se svojim najuspešnijim radovima, a učestvuju studenti završne godine svih deset specijalizovanih odseka i osam ateljea Fakulteta - Zidno slikarstvo, Primenjeno vajarstvo, Unutrašnja arhitektura, Primenjena grafika, Scenografija, Kostim, Tekstil, Keramika, Industrijski dizajn, Konzervacija i restauracija. U katalogu izložbe dr Ivana Kuzmanović-Novović, docent FPU, ističe da izlazak na javnu scenu diplomcima daje i novu ulogu, a to je podizanje nivoa estetskih kriterijuma u društvu. „U poplavi neukusa i kiča, zadatak ovih mladih i obrazovanih umetnika je da, ispoljavajući svoj talenat, oblikuju savremenu likovnu i modnu scenu i postanu kreatori ukusa, pri tome poštujući osnovna načela umetnosti.

Text 2: (From a blog)

Ne sećam se da mi je bilo ko tokom mog školovanja objasnio zašto ja u stvari treba da znam sve te stvari iz matematike, hemije, biologije, fizike, istorije, geografije i ostalih predmeta koji su tu jer je tamo neki pisac školskih kurikuluma tako zamislio. Tek kasnije sam shvatila da je sve to potrebno da bih shvatila kako svet funkcioniše. Jedanput kada to shvatite na dobrom ste putu da ostvarite vaše snove. Izgleda da niko nije objasnio našim edukatorima da je njihova uloga da nas osposobe da realizujemo naše snove. Oni radije preferiraju da nas zatrpaju masom suvih informacija bez ikakvog objašnjenja za njihovu primenu i onda da nam daju loše ocene ako ne uspemo da na nekom testu reprodukujemo iste a bez ikakvog dokaza da smo mi zaista i razumeli o čemu se tu zaista radi. Mnogi od njih sprovode vrhunsku ličnu, etičku i profesionalnu degradaciju prodajući ispite i diplome. Pri tome i kupci istih takođe vrše ličnu degradaciju plaćanjem, u stvari, bezvrednog papira koji svedoči o nečemu što oni nisu. I tako se krug zatvara.

Text 3: (Literature) Milovan Glišić, *Redak Zver*:

Englezi su nekako nastrani ljudi. Mnoge priče pričaju se o njima. Tako vele da je nekakav Englez po veka preležao na svom krevetu i pljuckao u tavan, vežbajući se da pogodi u jedno mesto. Za drugog kažu da je voleo jesti bifteke na lađi, pa se čitavo po godine vozio preko Temze tamo i amo, samo da se najede bifteka. Neki, opet, Englez video je u obližnjoj kući s prozora kako se jedan čovek hoće da ubije: nategne pištolj sebi u prsi, pa vrati, ne sme da skreše. Englez otrči u kuću onom čoveku i upita ga: "Šta ćeš to?" - "Hoću", veli, "da se ubijem, pa ne mogu!" - "Daj ovamo taj pištolj! Ovako se ubija!" rekne Englez i skreše pištolj sebi u prsi. Za jednog, opet, pričaju kako je nekoliko godina putovao železnicom, samo da bi mu se dogodila kakva nesreća, ili da prsne kazan ili da se sudare vozovi. I tako dalje i tako dalje. Ele, otprilike tako su nastrani ljudi ti Englezi.

Text 4: (Airport Instructions)

Procedura predaje prtljaga (korisni saveti)

Tokom boravka na aerodromu:

* Svoj prtljag, zaključan, obeležen nalepnicom ili priveskom sa imenom i adresom, nemojte ostavljati bez nadzora do predaje na šalteru za registraciju putnika.

* Savetujemo Vam da ručni prtljag koji unosite sa sobom u putničku kabinu bude pod vašim stalnim nadzorom.

* Savetujemo Vam da ne preuzimate i ne predajete tuđ prtljag kao svoj, zbog stvari koje bi Vam mogle ugroziti bezbednost ili Vas izložiti zakonskoj odgovornosti.

Avio kompanije zasebno propisuju određenu veličinu i težinu prtljaga koji se unosi u kabinu. Prilikom kupovine avio-karte obavezno se detaljnije raspitajte o dozvoljenoj težini i veličini kabinskog prtljaga.

Text 5: (Technology Instructions) (Tehnologija)

Zbog problema sa mejl serverima može da se desi da mejlovi koje šalje forumski softver ne stignu na odredište. Ako arhivirate privatne poruke slanjem na vašu email adresu i tom prilikom uključite opciju za njihovo brisanje, možete ostati bez vaših privatnih poruka.

Zbog toga vam predlažemo da arhiviranje vršite na sledeći način:

1. Prvo arhivirajte jednu poruku bez brisanja za probu.
2. Proverite vaš mailbox, ako je poruka stigla znači da je sve u redu. Ako ne, nemojte arhivirati poruke.

Trudićemo se da rešimo ovaj problem u najkraćem mogućem roku.

Text 6: (Politics) Politika

U Zimbabveu su održani predsednički i parlamentarni izbori, na kojima su birači odlučivali da li će aktuelni predsednik Robert Mugabe ponovo biti izabran za šefa države.

Glavni kandidat na izborima za predsednika Zimbabvea je šef države Mugabe, koji je nedavno napunio 84 godine i 28 godina se nalazi na čelu zemlje, a kome bi, ukoliko bude izabran, to bio šesti mandat. Izborna komisija je juče odbacila optužbe opozicije da je pripremljena velika izborna krađa i da je u biračkim spiskovima upisan veliki broj nepostojećih glasača, koji treba da obezbede pobedu ZANU-PF.

Text 7: (Recipe) Recept

Na vrelom maslinovom ulju prodinstaš malo crnog luka i ubaciš krupno sečene (ili lomljene) pečurke, propržiš ih 3-4 minuta uz stalno mešanje i naliješ litar vode. Krckaš to na tihoj vatri dok 3/4 vode ne ispari. Onda ubaciš mleveno svinjsko meso i naliješ još pola litra vode. Opet krckaš sve dok 3/4 vode ne ispari. Kad voda ispari, ubaciš sitno seckanu slaninu, sitno seckan suvi vrat, pospeš začinicima (đumbir, majčina dušica, peršun) i mešaš dok sva voda ne ispari... Onda naliješ pola litre kečapa, sipaš origano i na jakoj vatri mešaš dok kečap ne proključa...

Text 8: (a story) Priča

Naime, Vera se upoznala, ko zna kako, sa jednim mladim poljskim lekarom, dr Stanislavom Pujdakovskim. Poznanstvo se pretvorilo u obostranu ljubav, koja je bila krunisana brakom. Ali, zamalo!

Iako nije poznato zašto, pretpostavlja se da su uzrok bili ratovi, Balkanski i Prvi svetski. Dr Pujdakovski je žurno napustio Novi Sad i pohrlio u otadžbinu. Vera je ostala, uz majku, braću i sestru u Novom Sadu. Razlog nije bila nesloga, već siromaštvo i nezbrinutost porodice koju je Vera nesebično pomagala i pre i dugo godina posle svog bračnog brodoloma. Mladi poljski lekar se više nikad nije javio. Vera više nikada nije ispevala novu pesmu. Jedino je njena pesma "Pre rastanka" (danas "Jesenje lišće") obnovljena melodijom, prijateljski poručivala nekom u daljini da je život borba, da je retka sreća i da u njemu treba istrajati. Bio je to on u njenim mislima, mladi lekar, o kome nije nikada više ništa saznala...

Text 9: (Sportske vesti)

Brazilski fudbalski internacionalac Matuzalem nastaviće karijeru u rimskom Laciju u kojem će sledeću sezonu provesti kao pozajmljen igrač.

Ovaj 28-godišnji ofanzivni vezista se tako vratio u Italiju gde je od 1999. do 2004. godine nastupao za Napoli, Parmu, Pjaćencu i Brešu. Potom je sa dosta uspeha igrao za Šahtjor, da bi prošlu sezonu nezapaženo proveo u Saragosi koja je po ispadanju iz Primere odlučila da ga ustupi Laciju.