

# On the Need for History in Doing Balkan Linguistics

**Brian D. Joseph**

The Ohio State University  
[joseph.1@osu.edu]

**ABSTRACT:** The field of Balkan Linguistics is concerned with the patterns of structural and lexical convergence that can be found to hold in and among various languages of Southeastern Europe. It is in some sense quintessentially an historically oriented discipline, since virtually all of the convergence to be noted represents a divergence in each language away from an earlier quite different state. At the same time, though, many investigators into matters of the Balkan ‘Sprachbund’ convergence seem to take a purely synchronic typological approach, measuring the languages against one another with regard to various existing structural features they may show. It is argued here that while such a line of inquiry into the Balkan languages is useful, it cannot be done without some sense of history accompanying it. Several examples are discussed here in which a failure to take into account the historical background of a given convergent feature leads to quite erroneous results. In the end, history is vindicated, and is shown to play a key role in understanding Balkan convergences, in all their dimensions.

## I. Time and Speakers in Balkan Linguistics

1. Various languages in the Balkans constitute what has been called a **Sprachbund** (“linguistic area”) sharing a number of structural features, e.g.:

- how they form their future tense
- how they construct subordinate clauses

that plausibly have been attributed to centuries of sustained, intense, and intimate contact among speakers of these different languages.

2. My goals:

- explore the tension between synchrony and diachrony in studying the Balkan Sprachbund
- argue that history is an indispensable part of our bag of tricks as Balkan scholars.

3. Synchrony versus diachrony

- a. What it is (à la Saussure 1915): **synchrony** involves looking at how the contemporary situation is currently arrayed and structured, and **diachrony** involves looking at and describing how the contemporaneous situation got to be the way it is.
- b. What we can do with it: deduce/infer past states of affairs from present states of affairs; make reasonable guesses (hypotheses/inferences) about history based on the present
- c. Why we do that: why not? But also, curiosity – we have a desire to know, and so we try to know.

4. More reasons to draw (historical) inferences from synchrony:

- answers to questions about the history of an event reveal how participants internalize and act in the matter at hand, how they transform the situation as it “plays itself out” through them.

- especially if we realize that a “synchronic state” is really just a way-station to another synchronic state, which itself is a way-station to another synchronic state, and so on, so that diachrony becomes the progression through successive synchronic states.
- speakers are involved in synchronic states, so change in language and the transformation of synchronic states into something different do not happen without speaker involvement in some way, and seeing how synchronic states are altered tells us something about the speakers themselves and how they relate to their language.

5. On speaker involvement in language change (cf. Janda & Joseph 2003):

- it mirrors the comment during US Civil War (involving Union forces of the North versus Confederate forces of the South) by Confederate general George Pickett’s answer to inquiries about what went wrong in his ill-fated “Pickett’s Charge” at the battle of Gettysburg in 1863: “I think the Union Army had something to do with it”.
- similarly, regarding language change “I think that speakers have something to do with it”.
- note also James Milroy’s remark responding to Roger Lass’s notion of “exaptation” whereby languages make us of leftover otherwise useless pieces of older systems, what Lass referred to as “bricolage”; Milroy asked (1999: 188) “If there is bricolage, who is the bricoleur? Does the language do the bricolage independently of those who use it? If so, how?”.

6. A concrete case from variation in contemporary American English

- two forms for the past tense of *ask*: [æst] and [æskt] differing as to informal, casual-speech style [æst], without a *k*, and formal, more careful-speech style [æskt], with a *k*.
- generally, an informal speech form is felt to be derived from the careful form (cf. serious reductions like [ajmɔ̃nəlɪvnaʊ] for ‘I’m **going to** leave now’)
- the reductive change mapping from [æskt] to [æst] is one that on independent grounds is expected in a fast or casual (informal) style speech form (note similar cases e.g. [tEksts] (texts) surfacing as [tEkss]).
- thus quite plausibly [æst] derives from [æskt] synchronically
- and, from that we can infer that diachronically, [æst] derives from an earlier [æskt] and that the reductive process remains in effect today as part of the repertoire of style-shifting rules that speakers internalize.

7. (6e) is somewhat right, but not wholly so, and the difference between what “true” history tells us and what our reconstructed historical guess-work tells us is very instructive:

- no “straight line” connection between an earlier form [æskt] and the modern variant [æst]
- rather, the historical record (from *Oxford English Dictionary*) shows that [æst] has been a part of English for many centuries: (note highlighted part especially):

1 ásci-an, ácsi-, áhsi-, áxi-, áhxi-, áhxi-, áhxi-, áxsi-an, - an, -ǣan, æcsian; 2-3 axien, accien, 3 æxi, axi, 4 acsi, acsy, oxi, oxy, oxsi, oksī, 3-5 axen, (5 axse, exe,) 4-6 axe, ax, (6-dial. ax). Also <sup>f</sup>. 2 esci-, eski-en, 3 easki, (*Orm.*) asskenn, 3-5 ask-en, 3-7 aske, (5 haske, ascke, axke,) 4- ask. Also <sup>v</sup>. 3-4 **esch(e, esse, 3-5 asch(e, 5 ashe, 5-6 asshe, (north. asse, pa. tense ast).** [Common Teut.: OE. *áscian* was cogn. w. OFris. *âskia*, OS. *êscôn*, *êscan*, OHG. *eiscôn*, MHG. *eischen*, Ger. *heischen*, OTeut.

\**aiskōjan*: cf. Skr. *ish* to seek, *ichchh*<sup>ī</sup> wish. The original long *á* gave regularly the ME. (Kentish) *xi*; but elsewhere was shortened before the two consonants, giving ME. *a*, and, in some dialects, *e*. The result of these vowel changes, and of the OE. metathesis *asc-*, *acs-*, was that ME. had the types  $\bar{c}x$ , *ax*, *ex*, *ask*, *esk*, *ash*, *esh*, *ass*, *ess*. The true representative of the orig. *áscian* was the s.w. and w.midl. *ash*, *esh*, also written *esse* (cf. *æsce* [ASH](#), *wæsc(e)an* [WASH](#)), now quite lost. *Acsian*, *axian*, survived in *ax*, down to nearly 1600 the regular literary form, and still used everywhere in midl. and south. dialects, though supplanted in standard English by *ask*, originally the northern form. Already in 15th c. the latter was reduced dialectally to *asse*, pa. tense *ast*, still current dialectally

c. thus, modern [æst] is not a newly created informal/fast-speech form but rather either:

- *askedlast* variation has existed for a long time (so that in a sense, [æst] is not a recent entry into the English repertoire but represents an old informal/fast-speech form), or
- based on an inherited fast-speech form of past tense [æst] paired with a present tense [æsk], a fully “regular” past tense form [æskt] could easily, **at any time**, be re-created, analogically levelling out the difference between [æsk-] and [æs-], and thus giving the appearance of a pairing of careful speech [æskt] with casual, fast-speech [æst] at any synchronic stage.

8. A similar example: strong plural pronoun *them* in modern English and its relation to its corresponding weak (perhaps more informal style) variant 'em:

- a. a synchronic connection between the two forms is plausible, deriving the weak form as a reduction from the strong form (like masculine singular 'im versus its corresponding strong form him)
- b. an historical connection can be inferred from (a), with 'em deriving from them
- c. but, as I say elsewhere (Joseph 2004):

“It turns out, though, that such a connection and the accompanying inference of change are counterfactual: 'em is the inherited form (cf. Old English oblique *him*) while *them* shows effects of Norse contact (borrowed, replacing OE *hie*). Thus the “reduction” process, while plausible synchronically for relating *them*/'em, has no historical basis, and indeed the borrowing account allows one to avoid the embarrassing stipulation of getting rid via “reduction” of a relatively robust consonant like the initial *th-* (whereas reducing *him* to 'em involves only the loss of a relatively weakly articulated and acoustically depleted *h*).

9. Still, no harm no foul: what have we risked in (7) and (8)? Having some knowledge of the history simply corrects the record; we can make false inferences without it and draw false conclusions about directions of change, about the age of particular forms, and so on. But those are relatively benign errors; nothing crucial hinges on them other than getting the facts right.

## II. The Role of Typology in Balkan Linguistics

**10.** Synchronic structural facts yield a “Balkan type” of language, based on common features, (“Balkanisms”), covering phonology (a-b), morphology (c-d), syntax (e-g), and lexicon (h) (drawing on Miklosich 1861):

- a) the “prominence” of schwa, including reduction of unstressed a to schwa
- b) nasal consonants as syllable onset before other consonants
- c) a future tense formed with ‘want’ as an auxiliary
- d) no formal distinction between a genitive and a dative case
- e) lack of infinitive, with subordination instead involving a finite verb plus a conjunction
- f) the postposing of the definite article
- g) weak object pronouns co-occurring with (“doubling”) fuller objects of a verb
- h) formation of teens with ‘ten’ and the digits with an intervening preposition.

**11.** Diachronic vs. synchronic formulations:

- (d) could be stated as “there has been a merger of genitive and dative”
- (e) could be stated as “an earlier infinitive has been lost”
- Zuzana Topolinska (2004): the inventory of Balkanisms can be reformulated “starting from the positive, constructive, functional point of view”, i.e. on purely synchronic typological grounds.

**12.** Scorecard approach to “Balkanicity” (van der Auwera 2002):

Table 2. Balkanisms in the Balkan languages

Features	C.b.	C.e.	A.b.	B.b.	B.d.	A.a.	B.a.	C.d.	C.a.	B.e.	Tot
<i>Languages</i>											
Bulgarian	1	1	1	1	1	1	1	1	1	1	10
Tosk Albanian	1	1	1	1	1	1	1	1	0.5	1	9.5
Macedonian	1	1	1	1	1	0.5	1	1	1	1	9.5
Romanian	1	1	1	1	0.5	1	1	1	0.5	1	9
Geg Albanian	1	1	0	0	1	1	1	1	0.5	1	7.5
Greek	1	1	1	1	0	0	1	0	1	0	6
Serbian/Croatian	0	0.5	1	1	0.5	0.5	0	1	0.5	0.5	5.5
Romani	1	0.5	1	1	0.5	0.5	0	0	1	0	5.5
Turkish	1	0	0	0	1	1	0	0	0	0	3
Total	8	7	7	7	6.5	6.5	6	6	6	5.5	

**13.** Consequences of such typologizing for the Balkans: inferences and claims about contact among speakers that led to the convergences (understanding real history via evidence that is purely linguistic in nature), e.g. about:

- a. the nature of the mechanisms involved in language contact
- b. possible outcomes of contact (e.g. whether phonological rules per se can be borrowed, whether phonological segments individually can be borrowed, whether affixes can be borrowed, etc.)

c. psychological/cognitive implications, e.g. as to how deeply embedded in the grammar certain features are.

**14.** How to characterize a language if we did not know its history (Joseph 1999, 2006):

- a. re Greek: typical to say “no distinction between dative and genitive cases” (as if based on knowledge that an earlier stage of Greek had these cases) versus (more neutrally) marking of indirect objects with the same morphology as is used in marking possession
- b. re Algonquian languages (of Canada and US): no one says “no distinction between accusative and dative” (direct objects and indirect objects are encoded in the same way on the Algonquian verb), as if based on knowledge that case marking is irrelevant for Algonquian. (NB: case marking per se is irrelevant for Bulgarian and Macedonian yet they are brought in under the rubric of languages with no distinction between genitive and dative)

**15.** Typologizing and inferring in an historical vacuum leads one astray from the key point of *Balkan linguistics*, as opposed merely to *the linguistics of the Balkans* (Joseph 1999a); cf.:

- a. Schaller's 1975 distinction: "language of the Balkans" (a purely geographic designation) and "Balkan language" (a designation for those languages of the Balkans that participate in the Balkan Sprachbund and show parallels due to language contact).
- b. extended usage: "comparative syntax of the Balkan languages" (= examining the syntax of individual languages of the Balkans in comparison with other languages of the Balkans and elsewhere) vs. doing "comparative Balkan syntax" (examining the syntax of Balkan languages, keeping the Sprachbund in mind)
- c. More generally, "the linguistics of the Balkans" (= an interest in the languages as languages) vs. "Balkan linguistics" (= a focus on the Sprachbund phenomenon, on contact-induced features)

### III. Some Case Studies regarding Consequences of Ahistorical Typologizing

**16.** Fleeting vowel(s) in Turkish, in Balkan Slavic, in Albanian, and (maybe) in Greek:

- in Turkish 'city' is *şehir* when the *-r* closes a syllable, e.g. in the nominative singular *şehir* or the nominative plural *şehir-ler*, but *şehr* when the *r* is a syllable onset, as in the dative singular *şehr-e* 'to the city' or the definite accusative *şehr-i* 'the city'
- in Bulgarian, as a representative of Balkan Slavic, oversimplifying somewhat, the stem for 'silver' when the *-r* closes a syllable is *srebăr* (cf. *srebăr-nik* 'silver coin'), but *srebr-* when the *-r* is an onset (cf. *srebr-o*)
- so also in Albanian nouns like *emër* 'a name' (NOM) / *emr-i* 'the name' (NOM), where the *ë* (again oversimplifying a bit as there are similar words where *ë* does not disappear, e.g. *emëroj* 'I name')
- in Northern Greek similar alternations, e.g. *máθma* 'lesson.NOM' / *maθímatos* 'of-a-lesson.GEN'
- possible inference: historical (contact-based) relation among these phenomena, and possibility then that phonological rules can be borrowed

## 17. History of (ɪ):

- for Slavic, the fleeting vowel phenomenon results from several sound changes, most importantly the loss of the Proto-Slavic "jers" (ultra-short high front and back vowels) in so-called "weak" positions as opposed to their preservation (ultimately giving /e/ and /ə/, respectively) in "strong" positions, and other related syllable-based adjustments (e.g. insertion of /j/ "to break up stem-final consonant clusters terminating in liquids or nasals" (Scatton 1993: 193). Since similar alternations are found throughout Slavic more generally, the impetus behind the developments leading to the Balkan Slavic alternations is to be located in Proto-Slavic.
- for Turkish, words with the "fleeting vowel" are all borrowings from Arabic with word-final clusters that were otherwise not found in Turkish that were nativized in the borrowing process by the insertion of an epenthetic vowel.
- for Albanian, the alternation is simply the result of a quite natural change (loss of an unstressed vowel word-medially) that demands no special (i.e. contact-related) explanation
- for northern Greek, if it is considered here at all, the conditions are stress-related, not syllable-based
- the result: the fleeting vowel "parallel" is just a mirage – history clarifies this

## 18. Stressed schwa – can segments be borrowed?

- loan words in Balkans show the schwa (e.g. Albanian *këndoj* 'sing' from Latin *cantō*)
- also native developments, i.e. regular sound changes yield a schwa (e.g. in Tosk Albanian, from earlier nasal vowels, cf. *është* 'is' vs. Geg *është*).
- importantly, in some of the languages, there are several different sources for the schwa:
  - northern dialects of Macedonian (e.g. Tetovo, Kumanovo, Kriva Palanka) have schwa from Proto-Slavic jers, east central dialects (e.g. Tikveš-Mariovo) have schwa from vocalic *l*, many peripheral dialects (e.g. Ohrid-Prespa) have schwa from the Proto-Slavic back nasalized vowel, etc.
  - Romanian mid central vowel *ă* derives when stressed from Latin *o* in some contexts, as in *contra* 'against' > *cătră* or *foras* 'outside' > *fără*, and the high central vowel *î*, when stressed, derives from (stressed) *a* before a nasal, as in *cîmp* 'field' from Latin *campus*, and from *i* after *r*, as in *rîpă* 'cliff' from Latin *ripa*.
- thus, one would have to assume that words with schwa were borrowed from one language (that had schwa) into another that did not have schwa and that speakers of the recipient language adopted such forms without altering the schwa to a native sound. Such nonadaptation of loan words does occur, but it is perhaps less usual than their being adapted to the phonology of the borrowing speakers.
- moreover, one would then have to assume as well that schwa made its way into native vocabulary in the borrowing language. Again, that is possible and can happen but is again perhaps less usual.
- even more, though, the schwa in native vocabulary results from several earlier sounds in some of the languages in ways that look like regular sound changes, in that the changes affect a number of lexical items without exceptions.
- thus under a contact-induced account of the emergence of stressed schwa in the Balkan languages, the mere occurrence of a few stressed schwas in loan words has to be assumed to

be sufficient to lead different sounds in each of the recipient languages in the direction of schwa in such a way as to appear to be a regular, exceptionless sound change.

- further, if interference or substratum effects are claimed as the cause and not diffusion from lexical borrowings, how to explain that speakers of a language with schwa carried over their schwa-producing habits into their pronunciation of the words of another language, including in stressed positions where vowel reduction in the direction of a schwa is less expected.
- finally, even if all this were plausible, it is hard to link the developments in one language, say Albanian, with those in another, say Bulgarian, where different sets of second-language speakers would be involved.
- thus despite the insistence over the years that the overall schwa developments could well be a significant shared feature in the Balkans, the history of the individual languages makes it not at all clear that it has anything to do with language contact, other than occurring in some loanwords.

### 19. A grammatical example: *m*-negators in the Balkans

a. Greek, Albanian, and Romani (the language of the Rom, the “Gypsies”) all show negative forms (apparently words) that start with *m*-:

- Greek: *mi*
- Albanian: *mos*
- Romani: *ma*

b. their function: more than negation, but within the sphere of negation, they mark negative commands (prohibitions) and other sorts of modal negation (as opposed to negating simple statements of fact), thus a striking convergence at first glance

c. Adding Turkish to the mix: suffixal negator *-me-/ma-*, e.g. in negative commands:

- *gitme!* ‘Don’t go!’ (cf. *git-ti-m* ‘I went’)
- *alma!* ‘Don’t take!’ (cf. *al-di-m* ‘I took’)

but also with a broader function, negating just about any verb, e.g. *al-ma-di-m* ‘I didn’t take’ or *git-me-di-m* ‘I didn’t go’)

d. Since the usual direction for the flow of elements involving Turkish in the Balkans is from Turkish into the other languages, this would appear to allow these inferences:

- Greek, Albanian, and Romani got at least the nucleus of their *m*-negators from Turkish
- affixes can be borrowed
- grammatical affixes (such as negation) can be borrowed
- such forms can be liberated into full words.

### 20. Clarification from history: (19) is just wild nonsensical speculation:

- *mi* is documented as part of Greek since ancient times (Ancient Greek *mé*.)
- the *m*-negator in these three Indo-European languages in the Balkans has related forms in other IE languages outside of the Balkans (including Tocharian, in the far eastern realms of the IE world), so it is part of the inherited material from Proto-Indo-European into Greek, Albanian, and Romani respectively.

- within Turkic, the *m*-negator is found elsewhere (e.g. in Uzbek), so it is an independent Turkic element that has nothing to do with the other Balkan *m*-negators
- thus, the Balkan *m*-negator convergence is just an accidental convergence not one born of language contact within the Balkans.

**21.** A syntactic example: the postposed definite article

a. Modern Greek is generally excluded from consideration regarding this Balkan feature, since the usual position for its definite article is prenominal and first in the noun phrase, as in:

- ο *kalos anthropos* ‘the good man’

b. but, a few constructions in Greek appear to have postpositive definite articles:

- noun phrases with *olos* ‘all’, e.g. *olos ο kosmos* ‘all the world’
- noun phrases with the demonstrative, e.g. *ekinos ο anthropos* “that the man” (= ‘that man’)
- the doubled article construction with adjectives, e.g. *ο anthropos ο kalos* “the man the good” (= ‘the good man’)

c. So, are these facts to be included in the Balkan mix, and does Greek at least somewhat show this feature, in the Balkan way (with consequent (further) inferences about the possibility of borrowing syntax in language contact)?

**22.** Once again, history trumps typology, and provides an answer here:

a. whatever the history of the postposed article in Albanian, Romanian, and Balkan Slavic, these facts from Greek can’t be relevant for the simple reason that these constructions have largely been a part of Greek since approximately the 5<sup>th</sup> century BC (e.g. *ekeinos ho anthro:pos* ‘that man’ (literally “that the man”), *holon to derma* ‘all the skin, the skin as a whole’), i.e. from long before there was a Balkan Romance or a Balkan Slavic, and probably well before the article developed in Albanian.

b. nor can Greek be responsible for the article facts in the other languages, since it is just in these very limited contexts that Greek shows anything like a postposed definite article: in all other cases, i.e. in all the contexts where Albanian, Balkan Romance, and Balkan Slavic show a postposed article, Greek does not.

c. thus, history tells us here that the patterns in (21b) are not new to Greek and could not have been the basis for what is found in the other languages; there is thus no relationship between the Greek facts and the facts from other Balkan languages.

**23.** Conclusion: history gives us a more nuanced view of things than typology could ever offer.

**IV. Some deep historical closing thoughts on longevity**

**24.** Proto-Indo-European \*Hoyu with reflexes (involving slightly different root shapes but all within the bounds of what PIE allowed) all across the Indo-European family, in a large and wide-ranging set of words;



a. derivatives that refer to life and longevity:

Sanskrit *āyu* ‘life’ (\*Hoyu-)

Latin *aevum* ‘age, eternity’, *ae-t-as* ‘age’, *ae-t-ernus* ‘eternal’ (\*Heyw-o-(t)-)

Old English *ā* ‘ever’ (\*Heyw-)

b. derivatives involving the opposite end of the spectrum, referring to youth, e.g.:

English *youth*, *young* (\*Hyu-)

Latin *iuvenis* ‘young’ (\*Hyu-)

c. (a) and (b) are connected via a meaning ‘vital force’ or ‘youthful vigor’ (so *American Heritage Dictionary*, and elsewhere, e.g. Grassmann 1873 on *āyu* and most recently Southern 2006): the more vitality and youth you have, the longer your life ought to be.

**25.** One reflex in the Balkans of particular interest here: Greek *hygieia* ‘health’, a relative of *hygie:s* ‘healthy’

a. from \*Hyu-g<sup>w</sup>iy-es-, with \*Hyu- ‘long-life, life force’ element, with \*g<sup>w</sup>(e)i- (another root for ‘life’, as in Greek *bios*, English *quick*, etc.; cf. Weiss 1994), thus etymologically ‘having long life’

b. Ancient Greek *hygieia* gives Modern Greek *geia*, pronounced [ja] and ubiquitous in the greeting *ja su* for ‘hello’, but literally ‘health to-you’

c. thus Modern Greek *ja* is connected to the ‘long life’ nexus.

**26.** Another related Balkan word: Albanian *jetë* ‘life’,

a. also involved in a greeting, and also connected to the ‘long life’ semantic nexus.

b. a borrowing from Latin *aetas*, though with some reanalysis on the Albanian side according to Hamp 1968, and thus it derives ultimately from a derivative of the ‘long life’ root

c. *jetë* reflects (and renews) the ancient “long life” semantics in the phrase *përjetë* ‘forever’ (literally “for life”) and the derivative *përjetësi* ‘eternity’ (though admittedly some “naturalness” of semantic extension might be involved, cf. German *fur’s Leben* with similar composition and similar semantics).

d. Albanian *jetë* figures directly in a usage that pertains to ‘longevity of life’: the common greeting *tungjatjeta* ‘hello’: literally “*të u ngjattë jeta*” (= ‘May be-lengthened life-the’, with modal marker *të*, past nonactive marker *u* (from old reflexive), 3SG of optative of causative/factitive verb *ngjat-* ‘lengthen’ derived from *gjatë* ‘long, broad’, and definite form of *jetë* ‘life’))

e. thus *tungjatjeta* shows the longevity trope of \*Heyw- overtly, by juxtaposing ‘long’ and ‘life’.

**27.** There are thus ancient Indo-European semantic echoes, probably millennia-old, in the greetings Modern Greek *geia* and Albanian *tungjatjeta*

**28.** But wait, there’s more!

a. Greek *geia* is also used as a drinking toast either by itself or in the more formally crafted *s tin ija mas* ‘to our health’ (literally: “to the health our”).

- b. and, Eric Hamp (p.c.) has reported that Albanian *tungjatjeta* is used dialectally as a toast in drinking (he encountered it in the north of Albanian-speaking territory many years ago).
- c. Thus functionally these two greetings have exactly the same range, suggesting a persistent Balkan interest in ‘long life’ (admittedly a likely universal, though not always expressed this way – cf. Vedic *śatam śarādas* with a different trope, ‘hundred autumns’).

29. Moreover, it could well be – as suggested by Mark Southern (p.c. 6/2/05) – based on the very interesting analysis (Southern 2006) of PIE \*Hyew-o- ‘grain’ and \*Hoyw- ‘life force’ by which ‘grain’ is a derivative of ‘life force’ (grain being the “staff of life”), that Vedic Sanskrit exclamatory *yos* ‘health’ belongs here as well.

30. Sanskrit *yos* occurs primarily in the asyndetic expression *śam yos* ‘Glück (und) Heil’ (‘fortune (and) welfare’) in the Rig Veda, the oldest attested Sanskrit

- 20 occurrences, always with *śam*, 17 in this fixed form, and never inflected
- as neuter forms (*śam* originally meant something like ‘effort, power’, perhaps), the pieces *śam* and *yos* occur in this form as a direct object of a verb (their typical use), generally with ‘give’ or ‘bestow’, and the contexts in which they occur appear to be fairly solemn.
- still, in a few instances, there is no verb that obviously governs the phrase, and while it could be that a verb is to be understood, perhaps the phrase instead has an exclamatory value, as in

RIG VEDA 7.35.1

śam na indrāgnī bhavatām |  
fortune to-us Indra-and-Agni may-they-be

avobhiḥ na indrāvaruṇā rātahavyā |  
with-refreshments for-us Indra-and-Varuna the-recipients-of-the-oblation (may they be)

śam indrāsomā suvitāya ŚAM YOḤ  
fortune Indra-and-Soma for-welfare (may...be) ŚAM YOḤ(!)

śam na indrāpūṣaṇā vājasātau  
fortune to-us Indra-and-Pūṣan (may...be) in-gaining-of-booty

- without wanting to over-interpret here, an exclamation exhorting or invoking ‘long life’ is indeed what toasts are about.
- so, following a suggestion made by Southern (p.c. 6/2/05), an exclamatory use for *śam yos* would allow for the inference that an invocation of ‘long life’ in exclamatory contexts (such as a toast) in the Balkans could be a remarkably old bit of history preserved to this very day; it would itself (somewhat iconically) be an instance of remarkable longevity!

31. Conclusion: we do need history in doing Balkan linguistics, if for no other reason than to know the truth about our toasts (cf. *in vino veritas*, though meant in a different way); as we raise a glass to celebrate and to recognize the Balkans and their importance to all of us here, let us all say either **YA** or **TUNGJATJETA**, as the case may be.