Indo-European '*ego*',
Slavic *ja* = Runic *ek*, and Celtic Ø

ERIC P. HAMP

Professor Emeritus, Department of Linguistics, 1010 East 59th Street, Chicago, IL 60637, linguistics@uchicago.edu

Razprava podaja novo rekonstrukcijo razvoja osebnega zaimka za prvo osebo ednine v indoevropskih jezikih in ugotavlja naslednja območja inovacij: (1) anatolijsko < *VK*; (2) jugovzhodno indoevropsko (indoiransko, armensko) *eg'-H-ém*; (3) grško, latinsko, venetsko *eg'-(ó)H*; (4) severno indoevropsko (albansko, baltsko, slovansko, germansko, traško, toharsko) *eg'*.

The paper gives a new account of the development of the first person singular pronoun in Indo-European languages, finding innovating areals (1) Anatolian *VK*; (2) South-East Indo-European (Indo-Iranian, Armenian) *eg'-H-ém*; (3) Greek, Latin, Venetic *eg'-(ó)H*; (4) North I-E (Albanian, Baltic, Slavic, Germanic, Thracian, Tocharian) *eg'*.

Ključne besede: etimologija, osebni zaimki, indoevropski jeziki, balto-slovanski prajezik, keltski jeziki

Key words: etymology, personal pronouns, Indo-European, Proto-Balto-Slavic, Celtic

The twin Slavic forms *ja* and *jazъ* remain a riddle; see for example Sławski 1956: 477*sq.* In fact, one may fairly say that the matter is little advanced from the position represented by Meillet (1924, 394*sq.*): "*ce jaz s'est en général réduit à ja.* Cette forme est énigmatique… La longue initiale, étant propre au slave, ne se laisse expliquer que par des hypothèses incertaines." The numerous attempts at explanation ring all imaginable changes in detail but scarcely vary in principle; for a rich summary with references see Трубачев, ed., 1974: 100–3. So far as I can see, there is one approach — the correct approach — that has not yet been seriously considered.[1]

---

[1] The discussion expands on the previous versions of the paper which were published in *Indogermanische Forschungen* (1976) and *International Journal of Slavic Linguistics*

As with all other problems of reconstruction we must pay close attention to the question of dialectology among descendant branches. The North European dialects of I-E showed the simplest attested preform for this pronoun, *eg'*. The originally endingless shape of this form in Germanic is guaranteed for us by the Runic *ek*, which is found on the horn of the Gallehus beside the accusative neuter *horna* < *k'r̥nom* (cf. Ind. *śr̥nga-* 'cornu'). Likewise Baltic furnishes us a clearly endingless form in OPruss. *as ~ es, OLith. eš < ež*. Hittite *ú-uk* probably represents a revocalization (on the second person) of the same preform; see now my article in Bombi 2006 (*Festschrift Gusmani*), and remarks below. On the basis of the dialect geography of I-E we must therefore expect a (pre)-Slavic *ez*, to match its fellow North European dialects. The regular open syllable rule of Slavic would then lead us to expect a form *(j)e*. It seems clear that *ja* must somehow be the descendant of this inherited I-E form, and not an (illicit) apocopation of *jaz*. It remains only to explain the deviant quality of the vocalism.

Diverging from my earlier abbreviated account of this pronoun (*IF* 81, 1976, 39–40), it is now no longer necessary to depart from Occam's razor and look for an external source for the apparently long *ě* vocalism, in view of W. Winter's formulation in J. Fisiak ed., *Recent Developments in Historical Phonology*, Mouton 1978, 431–46, esp. 433. We see below also that the Proto-Slavic *ě* would have fitted well phonetically with the interfering early Iranian vocalism. The assumption of *ě* does not however preclude the possibility of a coexisting *e*, since the Baltic evidence points indubitably to a short reflex. We therefore assume that at an early date *eg'-* must developed a sandhi devoiced alternative before voiceless obstruents, thus yielding a later short *(j)e* not lengthened by following voicing.

We may pause here to note that on similar grounds a solution to the puzzling Albanian pronoun is also found; for I am convinced that Albanian is also to be classed as a North European I-E dialect (see, e.g., *IJSLP* 16, 1973, 1–6; since that time I have argued this relation on various grounds (e.g., in *Hylli i Dritës* in press 2008) which I consider logically stronger than and indeed prior in correctness to such claims made earlier by Jokl, which I have discussed in my 1966 article, Position of Albanian, 118*sq* (Hamp 1966).[2] The standard and literary *Shqip* form *unë* is to be segmented *u-në* (cf. dialectal *ti-në* 'thou' from my field notes), and the archaic simplex is preserved in Geg of Buzuku's language (1555) and the *Tosk Arbëresh* and *Arvanítika* enclaves as *u*; see my remarks *Studies presented to Whatmough* (1957; see now *Studime krahasuese për shqipen* 2007) 78*sq*. Now it is possible to derive this vocalism *u* in pretonic

---

and Poetics* (1983). An earlier version of the following appeared as *IJSLP* 27, 1983 '*JA* = Runic *EK*'- *Addenda* not entered in proof are here inserted in this revised version, which takes account of matters I have learned over the past two decades. The first version of this article was written in 1975; *cf.* a brief start *IF* 81, 1976, 39ff.

[2] Trans. now in Albanian in E. P. H. *Studime krahasuese për shqipen*, përgatiti dhe redaktoi Rexhep Ismajli, Prishtinë 2007, #4, 115–145 = *Studime* 13, 2006 (Prishtinë), 45–72. Dr. Rrahman Paçarizi.

or (later) closed syllables (*e.g.*, in sandhi) from an earlier diphthong **uo*, which is the normal Albanian product of **ō*. This *o~na~ue* in turn may come from *ē*. Such an Albanian form could be derived only from the endingless **eg'*, since otherwise the obstruent would leave a reflex in the attested language. We need then only assume that in pre-Albanian **eg'* was lengthened to **ēg'*. As I show *in extenso* elsewhere, this pronoun furnishes one of the pieces of evidence that Albanian and Illyrian shared with Balto-Slavic in the lengthening of syllabics before original mediæ, which Winter has demonstrated.

By contrast with the North European I-E etymon, Greek, Latin and Venetic show **eg'-óH*, perhaps **eg'- óH_o* (with H_3). I have argued (*JAOS* 90, 1970, 230) that Armenian *es* reflects **eg'H*, and that such a form is required to explain the Armenian *s* (and not *c* or *j*). We see then that the South European or Mediterranean dialects are characterized by a laryngeal suffix. I have further suggested (*Papers from the 7th Regional Meeting of the Chicago Linguistic Society*, 1971: 442) that the Tocharian A feminine *ñuk* and the Tocharian B *ñäś* reflect **m-yek-(-w-)*. But this has also been contested; it may well need revision (see D. Q. Adams, *Tocharian Historical Phonology and Morphology*, 1955, 153*sq*). Perhaps **m-yek-we* (f. **-wā*) would account for the data. If that is so, it is possible that **yek* reflects a North European **eg'* which agrees with my present view (2009) of Tocharian.

Venetic *ego* is important for classifying Venetic among I-E dialects. It should be noted explicitly that in this respect Venetic goes with Latin (and Greek), and not with the North of Europe. Regarding **mego*, which valuably confirms *ego*, and the parallel that has been noted with Germanic *ik/mik*, Sommer's theory of an analogy on the nominative seems to me by far the best, regardless of whether the forms are genetically one or, as I think, parallel. The Hittite situation stands quite apart, but tends to illustrate how a like result can be produced independently, yet independently in the same paradigm and thus related by complex context.

It is time now to turn to *jaz*; surely we must accept the arguments which have been adduced claiming the initial yod as the older form and the OCS and Bulg. *azъ* as a regional development with loss of yod. We must note immediately that for such an I-E form it is totally unmotivated, particularly morphologically, to seek a preform in **ō*; this means that the natural avenue to a solution bearing yod will be via the vowel **[ě]*. We also see clearly now that the doubly suffixed form **eg'-H-Vm*, perhaps **eg'-H_o-ém*, is characteristically part of a distinct Indo-Iranian dialect feature: Skt. *ahám*, Avestan *azəm*, OPers. *adam*; cf. Skt. *tvám*, *vám* (du.), *yuvam* (du.), *vayám*, *yūyám*, Av. *tvəm*, *vā* (du.), *vaēm*, *yūžəm*, OPers. *tuvam*, *vayam* – almost all attested with the same ancient enclitic added. It seems reasonable to equate this -*ám* with Latin emphatic *em*, and ultimately with Welsh [*ef*], which I have discussed elsewhere. It will be seen that I differ here with Wackernagel (*Altindische Grammatik* III, 2 453−67), who regards the first singular as inherited from I-E in this complex form and the other pronouns as formed upon that analogy; I see them all as similarly formed innovations, with enclitic added. Burrow appears to follow Wackernagel. On the

basis of the above reasoning I therefore suggest that *jaz* cannot reasonably be inherited by Slavic from I-E as a unit. Yet we find **jazъ* attested in Slovene in *jàz* (*to sem jàz*), but *ja*, *jä* in Resia (Friuli) and *ja* dialectally in Eastern Styria and Bela krajina (Pleteršnik 1894) and perhaps in Bulg. *знам я* 'of course I know'. These marginal Bartoli distributions point to a very early diffusional adoption of * *jazъ*.

Ever since Rozwadowski, and increasingly with the recent work of my lamented late friend Gołąb (*For Wiktor Weintraub*, 1975: 151–9), it has been clear that important and intimate relations between the Slavs and Iranian speakers must be assumed to have held over a considerable time period.

We may speculate how the Iranian borrowing may have been natural in Slavic. We must further recall that in these westerly dialects of Iranian the descendant of Iranian short *a* was markedly fronted; *cf.* Ossetic *æz*, Kurd *ä*. Moreover, since the old long and short vowels had diverged in quality (height etc.), thereby destroying the original Iranian correlations of quantity, it was possible for new contextual phonetic quantities to arise. Hence we may suppose than an *æ* in open syllables sounded somewhat long to a Slavic ear. A bilingual Slav could therefore have easily matched [æzəm] with his native **ĕz*, attributing to the former a derivational morphology comprising a suffix and perhaps a functional relation of emphasis or topicalization or the like.

Let us now summarize the reconstructions and groupings arrived at. It is not important whether or not this pronoun actually had an initial laryngeal, which could have been only the first laryngeal, or *H_1.

Figure 1.

Areal	Branches		Genetic
	Anatolian: Hitt. *ú-uk*		*VK*
South (-East) I-E	Indic: Skt. *ahám*		
	Iranian: Av. *azəm*		*eg'-H-ém*
	OP *adam*		*-óH*
	Armen. *es, *eg'H*		
	Greek ἐγώ *egoH*		
	Latin *egō*		*eg'-(ó)H*
	Falisc. *eko*	*egō*	
	Venet. *ego*		
	Oscan *íív?*		
North Europ. I-E	Alban. *u < *ē*	*ēg'*	
	Slav. *ja < *ě*		*ēg'*
	Balt. *eš, *ež*		~*eg'* *eg'*
	Gmc. ek		*eg*
	?Toch. *yek?		
	Thracian		
	"Cimmerian"		?
North Europ. I-E	Prehellenic		
	Messapic →?→ Alban.		
	Phrygian →?→ Italic		

We see immediately that the *VK* and *eg'* branches represent what we may call "conservative *I-E*" for this feature. If Celtic goes with North European, as I think it does, and especially if it is not Indo-Iranian or Pontic or Italic (which are clearly all true), and this is "residual I-E", then we would expect Celtic to show reflexes of unsuffixed *eg'* > "centum" *eg*.

What we actually find can be illustrated by Old Irish. I cite Thurneysen's inventory for OIrish (and also for other forms and certain comments) from the informative and thoughtful article of Diarmuid Ó Sé (1996); the tabulation of Thurneysen (Ó Sé 1996: 20) is, showing basic and surviving contrastive ("emphatic") forms:

Figure 2.[3]

	Singular	Plural
1.	*mé/me-sse*	*sní/si-nní*
2.	*tú/tu-ssu*	*sí/si-b(-si)*
3. m.	*é/é-som*	*é/é-ssom*
f.	*sí/si-s(s)i*	
n.	*ed*	

Ó Sé remarks on the ambiguities and asymmetries in this system (20*sq.*) and then goes on to show how new balances were instituted over time and in the dialects. We may summarize the main lines of formal development (26 and 30):

Figure 3.

	Scotto-Manx	parent Mod. Ir.	Early Mod. Ir.~Erris, Mayo
sg. 1.	*mé (mí?)/mi-se*	*mé/mi-se*	mé
2.	*t(h)ú/t(h)u-sa*	*t(h)ú/t(h)u-sa*	*tú/thú*
3. m.	*é/é-san*	*sé, é/sei-sean, ei-sean*	*sé, é*
f.	*í/i-se*	*sí, í/si-se, se*	*sí, í*
pl. 1.	*sinn/sinn-e*	*sinn/sinn-e*	*muid~sinn*
2.	*sibh/sibh-se*	*sibh/sibh-se*	*sibh*
3.	*íad/íad-san*	*síad, íad/síad-san, íad-san*	*siad, iad*

Ó Sé brings out the fact that in an Old Irish sentence (apart from proclitic possessive, enclitic redundant NP and VP echo-forms, and suffixed prepositional objects) personal marking is accomplished by:
 1. suffixed inflexion of the verb for the subject;[4]

---

[3] Hyphens are mostly my own.

[4] The deviant marking on *olse, olsi* (and *olseat*) 'inquit' results because *ol* was a deictic, and not a verb, in origin. Its personal marker was therefore a clitic, no doubt a predicate phrase (with the relative copula *as* < *est-to*) at bottom. Note that Welsh *heb(yr)* was on its way to the same end from the opposite (verbal Latin *in-quit*) direction.

2. clitic insertion on the verb for the pronominal object;
3. independent personal pronouns as predicates to the copula.
   The copula itself was proclitic.

On this basis we may say that the I-E origins of these forms of marking were:
1. proclitic possessives < *I-E* clitic genitives;
2. suffixed prepositional objects < *I-E* enclitic personal and deictic datives and accusatives, and frozen or fossilized locatives, locativals, and instrumentals; note also that *I-E* independent personal pronouns in the dative had been replaced by prepositional or focused-particle (directional or relational) constructions;
3. verbal suffixed inflexion < *I-E* suffixed personal inflexion and enclitic deixis;
4. clitic pronominal objects < *I-E* clitic pronouns in Wackernagel's position;[5]
5. independent predicate pronouns < I-E topicalized (accented) pronouns.

We see then that when we investigate the fate of *eg in Celtic we must inspect the forms and behaviour of the last-named class of marking.

Generally, the forms we find in this class:

2. *tú* < I-E *tu
m. *é* < *ei
f. *sí* < **siH$_a$* ← **seH$_a$* = *sā* (see my note in current *Studia Celtica*)
3 pl. *é* < *ei-es

These forms are derived from old nominatives, which is appropriate for predicates in a copula sentence. The neuter *ed* < **i-d(-a ...)* as a nominative-accusative is equally appropriate. The 1 pl. and 2 pl., *sní* and *sí* < *swí* (= Welsh *chwi*), can also be plausibly accounted as nominatives within our explication of their divergence from cognates in other *I-E* branches.

The remarkable gap, alone among well attested branches of I-E, is *eg. Why should this be, and why do we instead find *mé* < *me (or **?me*) accusative (and perhaps in Middle Irish *mí* < **?mi* dative?)?

We must recall the much noted and discussed change that took place in Insular Celtic sentence structure, whereby verb-first order (VSO) took over. This had two great consequences relevant to our problem:

(1) *I-E* *Cop(ula) + Pers(onal predicate) could survive;
(2) *I-E* *# Pers$_x$ > ICeltic *# Cop + Pers$_x$, or was transformed

The resulting structures were: (nom. = nominative; obl = other case)

---

[5] In accord with this distribution, I would note my claim that in Wackernagel's position in compound verbs we find the nominative only in the form of the anaphoric *-*s*, which in Insular Celtic has the effect of blocking lenition. I announced this finding in my Thurneysen lecture in Bonn 1989.

(1) *Cop + Pers$_{nom}$

(2) *Cop + Pers$_{nom, obl}$

(2) would result especially if REL(ative marker) did not show case inflexion, as it sometimes did not. Thus the case form of Pers became nom (as we have already seen for most instances and for good reason), or nom-acc. (*ed*), or potentially obl = *obliquus*. We assume that the last was mostly forced out, as being in conflict with inherited Cop(ula) syntax.

Now consider the phonological fate of **eg*. By final obstruent loss (**tod* neut. sg. nom-acc. > *tó* 'yes') **eg* > **e* (an similarly for British Celtic). Then by lengthening in monosyllable *Auslaut*, **e* > **é*. Meantime **ei* > ICeltic **ē* in 3 m. sg. **ei* and 3 pl. **ei(-es)* > **ei-^{h}*.

The ambiguity of 3 sg. m. and pl. and 1 sg. was too great. The otherwise occurring **obl* (acc.) *mé*, and perhaps old dat. *mí*, was preferred, and replaced **é*.

Thus the shift to VSD structure in Insular Celtic must have had a powerful effect in disfavouring the descendant of **eg'*. In turn, the phonological aspect of our argument makes it strongly likely that Celtic once possessed the North European unsuffixed centum form **eg*.

REFERENCES

Douglas Q. ADAMS, 1988: *Tocharian Historical Phonology and Morphology* (=American Oriental Society, vol. 71). New Haven: American Oriental Society.

Raffaella BOMBI (ed.), 2006: *Studi linguistici in onore di Roberto Gusmani*. Alessandria: Edizioni dell'Orso.

Victor ERLICH (ed.), 1975: *For Wiktor Weintraub: Essays in Polish Literature, Language, and History Presented on the Occasion of His 65th Birthday*. Berlin: Mouton de Gruyter.

Eric P. HAMP, 1966: The Position of Albanian. *Ancient Indo-European Dialects*. Eds. H. Birnbaum and J. Puhvel. Berkeley & Los Angeles: University of California Press.

– –, 1976: Western Indo-European Notes. *IF*, 36–40.

– –, 1983: *Ja* = Runic *ek*. *IJSLP*, 11–13.

– –, 2007: *Studime krahasuese për shqipen*. Rexhep Ismajli, ed. (= *Botime të veçanta* LXXXIII, *Seksioni i gjuhësisë dhe i letërsisë*, Libri 37). Prishtinë: Akademia e Shkencave dhe e Arteve të Kosovës.

Indogermanische Forschungen (= *IF*).

International Journal of Slavic Linguistics and Poetics (= *IJSLP*).

Journal of the American Oriental Society (= *JAOS*).

Antoine MEILLET, 1924: *Le slave commun*. Paris: Champion.

Diarmuid Ó SÉ, 1996: Personal Pronouns in Gaelic Dialects. *Éigse* XXIX, 19–50.

Papers from the 7th Regional Meeting of the Chicago Linguistic Society. Chicago: University of Chicago, 1971.

Maks PLETERŠNIK, 1894/95: *Slovensko-nemški slovar.* Ljubljana: Knezoškofijstvo.

Ernst PULGRAM, 1957: *Studies Presented to Joshua Whatmough on his Sixtieth Birthday.* The Hague: Mouton.

Franciszek SŁAWSKI, 1956: *Słownik etymologiczny języka polskiego.* Cracow: Towarzystwa Miłośników Języka Polskiego.

Олег Н. ТРУБАЧЕВ (ed.), 1974: *Этимологический словарь славянских языков.* Moscow: Nauka.

Jacob WACKERNAGEL, 1896–1957: *Altindische Grammatik*, Bände I–III. Göttingen.

Werner WINTER, 1978: The distribution of short and long vowels in stems of the type Lith. *ésti : vèsti : mèsti* and OCS *jasti : vesti : mesti* in Baltic and Slavic languages. *Recent Developments in Historical Phonology.* J. Fisiak, ed. The Hague, New York: Mouton. 431–44.

PRAINDOEVROPSKO *'EGO'*, SLOVANSKO *JA* = RUNSKO *EK* IN KELTSKO *Ø*

Razprava, ki razširja in dopolnjuje prejšnji različici članka, objavljeni v revijah *Indogermanische Forschungen* (1976) in *International Journal of Slavic Linguistics and Poetics* (1983), prinaša nova spoznanja na področju indoevropskega jezikoslovja; avtor podaja novo rekonstrukcijo razvoja osebnega zaimka za prvo osebo ednine v indoevropskih jezikih, pri tem pa ugotavlja štiri območja inovacij: (1) anatolijsko < **VK*; (2) jugovzhodno indoevropsko (indoiransko, armensko) **eg'-H-ém*; (3) grško, latinsko, venetsko **eg'-(ó)H*; (4) severno indoevropsko (albansko, baltsko, slovansko, germansko, traško, toharsko) **eg'*.