

Echoes of Indo-European cultural semantics in Albanian

Brian D. Joseph
The Ohio State University

Abstract

Hopper (1991: 28) offers a key element in the study of diachronic semantics, namely what he refers to as the principle of “persistence”. The original meaning of a morpheme (or word) is said to “persist” after it enters a different domain. While Hopper largely discusses shifts into the grammatical domain, I demonstrate here that it applies equally well to the lexical sphere. This principle means that there will be what can be called as “semantic echoes” in various forms that one encounters in a language, that is to say, there will be traces in later usage of the earlier semantics of a form. I discuss three case-studies of such persistence in Albanian, whereby details of usage for various present-day Albanian words and phrases reveal aspects of what can plausibly be reconstructed as Proto-Indo-European cultural semantics, echoed in continuations of the language more than six millennia later.

Keywords: semantic change; diachronic semantics; principle of persistence; language ideology; evidentiality; tongue; Albanian; Greek; Hittite; Latin; Sanskrit; Indo-European

1. Introduction

A key finding in diachronic semantics is what Hopper (1991: 28) refers to as the principle of “persistence”, namely that original semantics of a morpheme (or word) “persist” after it enters the grammatical domain. This principle means that there will be what can be referred to as “semantic echoes” in various forms that one encounters in a language, that is to say, there will be traces in later usage of the earlier semantics of a form. These earlier semantics can be either directly attested or inferred and thus reconstructed for a previous stage of the language.

An example of this persistence can be seen in the Albanian non-active¹ marker *-h-*, as in *la-h-em* ‘I wash myself / I am washed’ (cf. active *la* ‘I wash (something)’). This *-h-* derives from Proto-Indo-European (PIE) **-ské/o-*, a verbal suffix which ends up having an inchoative meaning in Latin (e.g. *rubē-sc-ō* ‘I grow red’), an iterative sense in Hittite (e.g. *walh-iški-zzi* ‘beats repeatedly’), but also a simple present, as (ultimately) in Sanskrit (e.g. present stem *gaccha-* ‘go’ to the root *gam-*, equatable with Greek βάσκε-, thus both from PIE **g^wm-ske-*). All of these uses are presential in nature, so that even though it is hard to reconcile the differences in meaning in these various forms in the daughter languages to come up with a single meaning for the proto-language suffix, it is

¹ This verbal category is also referred to in the literature as “mediopassive” or “middle”, but given that the range of semantics for this category is rather broad, taking in passive, reflexive, reciprocal, and stative senses, among others, I prefer the designation “nonactive” as the form itself is systematically opposed to active forms.

likely that it involves reference to present time in some way. Although co-opted into grammatical use to mark nonactive voice, nonetheless the *-h-* betrays its present-oriented original semantics by being restricted in Albanian to non-past forms; that is, while the present nonactive is as above, e.g. *lahem*, with *-h-* from **-sk-*, the past nonactive is formed rather by particle *u* (ultimately from PIE **swe*, an element with reflexive meaning) plus a verb form with active endings, e.g. *u lava* ‘I washed myself / I was washed’ (cf. *lava* ‘I washed (something)’).

Thus, persistence can be seen in change leading to new grammatical semantics, as in the case of **-ské/o-*, but it can be asked whether it can happen in meaning shifts that are of a purely lexical (i.e., nongrammatical) meaning. My answer is yes; persistence can indeed be observed in the semantic development of lexical items, and I discuss here several case studies from Albanian that demonstrate this. In particular, what these cases show is persistence — what I refer to here as “echoes” and “sub-echoes” — of Proto-Indo-European semantics in present-day Albanian.

Albanian is a separate branch of Indo-European that has often been “accused” of having borrowed a huge percentage of its lexicon;² while such a view may be due in part to the language’s relatively late attestation — the first Albanian text is a one-line baptismal formula from 1462 and the first substantial text is a book of missals, readings for the Catholic mass, that dates from 1555 — there may very well be more Indo-European elements in Albanian than is readily recognized. In a sense, then, these case studies offer testimony to the Indo-European legacy in Albanian.

Each one of the case studies presented here centers on the etymology of a different Albanian word or phrase and adds to that various details about how the lexical material in question is used in attested Albanian and other related languages. This combination allows for conclusions to then be drawn about how much of and what aspects of the original semantics are echoed in later usage, thereby addressing the matter of persistence.³

2. Case Study 1: Ideologies of “Us vs. Them” and Language Clarity

The issue in the first case study is the etymology of the Albanian ethnonym *shqip*. However, the background to this account is the marking within Indo-European, both the proto-language and the family at large, of the linguistic separation of ‘us’ from ‘them’ via the characterization **barbaro-*.

In particular, a PIE form **barbaro-* is reconstructible based on Greek *βάρβαρος* ‘unintelligible, non-Greek’ and Sanskrit *barbara-* ‘stammering’, and possibly Latin *baburrus* ‘foolish’.⁴ There are also likely related forms that show **-l-*, e.g. Lat. *balbus* ‘stammering, lisping’, Sanskrit *balbalā* (*kar-*) ‘(to make) a stammering’; with Slavic

² Baldi (1983: 88) speaks of Albanian’s “heavy lexical indebtedness” to Greek and Latin, and Fortson (2010: 446) notes the “massive overlaying of foreign vocabulary and concomitant loss of much of the native lexicon” within the history of the language.

³ The first two case studies are based on as-yet unpublished material presented, respectively, in Joseph 2005ac and in Joseph 2005b; the third case study is based on Joseph 2003, 2010. Other related cases are discussed in Joseph (To appear).

⁴ Admittedly, this form is late Latin, first found in the works of the 3rd century AD scholar Origenus; thus, it may not be relevant here.

cognates for both *-r-* and *-l-* forms (e.g. Serbian *brboljiti* ‘to chatter’). This appears to be an onomatopoeic formation, mimicking — or attempting to represent — the perception of what unintelligible speech sounds like. Interestingly, there is a degree of iconicity in the form of this word, in that there are two rare, and thus, in a certain sense, noncanonical and somewhat unintelligible Indo-European sounds, in particular **b* and **a*, in a word referring to unintelligibility and something linguistically unusual, as well as reduplication in a word that seems to indicate stammering. That is, one can interpret the unusual sounds in the word as lending it an ‘alien’ feel even in PIE and thus the meaning in PIE, in referring to garbled speech, may actually have been something like ‘alien -- or marginal -- to our linguistic norms’.

Impeded speech would have been one type of speech outside of PIE norms, but so too would the speech of outsiders; thus the non-impeded-speech meaning one finds in Greek, where *βάρβαρος* refers to language that is ‘non-Greek’ need not be a Greek innovation per se but could rather reflect the PIE meaning of ‘outside of our linguistic norms; outside of our usual speech’.⁵ There are also forms showing two language-specifying derivational patterns of Greek. One is a derived verb in *-ιζ-*, *βαρβαρίζω* ‘speak like a barbarian, speak broken Greek’, and the other is the derived adverb in *-(σ)τί*, *βαρβαριστί* ‘in barbarian or foreign language’; it should be noted that these derivatives also, and perhaps primarily, have non-language meaning, respectively ‘behave like a barbarian’ and ‘in barbarous fashion’.

In this view, then, this reconstructible lexical item for PIE has a distinctly ideological tinge to it, as takes a PIE-centric view of how to divide up the world linguistically, basically ‘us’ versus ‘non-us’.

It must be noted that not all linguistic differences were such as to lead to labels of *βαρβαρισμός*. For instance, there is a distinction seen in Homeric usage between “language of men” vs. “language of gods”, as seen in (1), from *Iliad* 1.403-4:

- (1) ὃν Βριαρέων καλέουσι θεοί, ἄνδρες δέ τε πάντες / Αἰγαίων’ ...
 ‘whom the gods call Briareus, but all men (call) Aegaeon’

This distinction, however, according to Colvin (1999: 44) was “an example of recognized linguistic diversity” for the Greeks. Moreover, it seems to have been an acceptable type of “variation” in PIE, as discussed by Watkins 1970a, who observes echoes of this distinction not only in Greek but also Sanskrit, Hittite, Old Norse, and Irish. In particular, Watkins suggests that since the Irish version of this distinction is embedded in a tradition that recognizes different types of *berla* ‘language’ “we have to deal here in Irish with a genuine inheritance from an Indo-European poetic doctrine, a doctrine of the nature of poetic language and its relation to ordinary language” (p. 16). Arguably, therefore, this is an element of IE (folk) linguistic ideology but not of the exclusionist *βαρβαριστί* type (presumably because the gods were part of “us” within the PIE world-view).

⁵ It must be admitted that the meaning ‘non-Greek’ may be somewhat late within Greek. That is, this adjective does not occur in Homeric Greek in just this form; *βαρβαρό-φωνος* ‘speaking in a foreign tongue’ does occur in *Iliad* 2.867, but one might argue that the added language-related specification (*-φωνος*) means that *βάρβαρο-* itself does not refer to language.

I have alluded here to *language/linguistic ideology*. I take this notion to refer to the collective beliefs that ordinary speakers hold about their language, its structure, its history, and its use, as they perceive it, and more broadly, about language in general.⁶ The linguistic reflection of an “us” versus “them” distinction, therefore, is an instance of linguistic ideology, since it has to do with speakers’ thinking about language in general and about its use to mark a distinction relevant to society.

This type of linguistic ideology was particularly well-developed in ancient India. Cardona (1990: 1) describes it thus:⁷

From early Vedic times, Indo-Aryans had an awareness of themselves as opposed to peoples with whom they came into contact and conflict. This awareness involved cultural and racial factors, including language ... Āryas ... against non-Āryas ... The contrast ultimately developed into one between idealized speakers of a language that was culturally and ritually pure (*samskr̥tam*) ... and barbaric speakers (*mlecchāḥ*)”.

Specifically regarding the language of humans and nonhumans, Cardona notes a highly relevant passage in the *Śatapathabrāhmaṇa* (3.2.1.22-24) referring to the use by demons (*asuras*) of *vācam* ... *upajijñāsyām* ‘indistinct speech’ (literally, “speech ... to-be-found-out/enigmatic”), speech that is identified as that of a *mlecchas* ‘barbarian’⁸. Moreover, anything involving *mleccha* is further identified as non-Brāhmanic behavior, behavior not befitting the highest caste of Hinduism, for it is said *na brāhmaṇo mlecchet* ‘a Brāhman is not to utter barbaric speech’.⁹ Thus, nonhuman language is equated with alien/muddled/unclear speech. As with PIE *barbaro-, there is an iconicity here, via the rarity — and thus “unclearity” — of an initial #ml- cluster in Sanskrit and the alien character of *-l-* in general in Sanskrit.¹⁰

Cardona (1990: 2) continues with an important observation: “If people thus refer to themselves and their language in a particular way, they certainly are aware of a contrast between themselves and others, both as a people and with respect to the languages they speak.” Thus, the ideology that emerges from all of this, an ideology that was presumably of PIE age, can be characterized as one of: ‘speak (as we do)’ vs. ‘speak

⁶ This definition is adapted from that of Silverstein 1979, where linguistic ideology is defined as “sets of beliefs about language articulated by users as a rationalization or justification of perceived language structure and use”.

⁷ The term *Arya* is the self-characterization (“autonym”) used by the Sanskritic speakers of India to label themselves.

⁸ The Sanskrit word *mlecchas* is defined in Monier-Williams (1899: s.v.) as “a foreigner, barbarian, non-Aryan, man of an outcast race, any person who does not speak Sanskr̥t and does not conform to the usual Hindu institutions”, thus clearly falling within an “us-versus-them” distinction.

⁹ This phrase is literally “not a-Brāhman should-‘mlech’”, with a negated optative verb (*mlecchet*) expressing the prohibitive modality; Monier-Williams (1899: s.v.) defines the verb *mlech-* as “to speak indistinctly (like a foreigner or barbarian who does not speak Sanskr̥t)”.

¹⁰ In the primary dialect represented in the earliest texts we have for Sanskrit, PIE *r and *l merged as /r/, so that the sound /l/ falls outside of this main Sanskritic dialect.

some other way', a characterization that is equivalent to 'speak clearly' vs. 'speak muddledly'.

At this point, and against this background, we can return to the matter of the etymology of *shqip*, the Albanian word for 'Albanian', for clarity of speech and "us-vs.-them" ideology figure in the etymology. Following Hamp 1999 (cf. also Lloshi 1999: 277-278, and Meyer 1891: 411), it can be posited that *shqip* ultimately derives from Latin *excipāre*, a derivative of *excipere* 'to pull or tease out', from *ex-* 'out' + *capere* 'take'; the semantics of this form was thus originally "take-out/take-in meaning".¹¹ There may have been some influence, both in meaning and maybe in form, from other Latin words with *ex-* *k..p-* or *ex-p...k-*, such as *explicāre* 'to explicate' or *excerptum* '(text) extract'. Particularly relevant here is the fact that *shqip* is used, among other things, as an adverb, as in *Flas shqip!* 'Speak Albanian!'. Interestingly, it always has this form, simply *shqip*, in this adverbial use, and is never **shqipisht*, with the adverbial *-isht*; in this way, it is unlike other language-adverbs, e.g. *anglisht* '(in) English' or *turqisht* '(in) Turkish'. Thus, *shqip* is at its roots means 'clearly', with reference to speech, thus 'clear speech' i.e. 'understandable/intelligible speech',¹² and, presumably since it came to be the autonym for the language, more specifically 'my/our language/speech'. It is thus not, *sensu stricto*, an ethnic label per se, but rather an ethnolinguistic one. Its etymology and original meaning is thus still echoed in its adverbial use in the collocation *flas shqip*; the aspect of its meaning referring to clarity of speech persists in this usage.

Relatedly, though as an aside, it is possible to point to other echoes in Balkan languages of this linguistic ideology. For instance, there is a parallel in Modern Greek. The language autonym *ρωμαίικα* '(Modern) Greek', based on associations between the Eastern Orthodox (or "Holy Roman") Empire and Hellenism, while perhaps not common nowadays, nonetheless is found in the phrase *(δεν) καταλαβαίνεις ρωμαίικα* 'Do(n't) you understand what I am saying?!' (literally, "do(n't) you understand Romaic (= Greek)?", that is, "Do(n't) you understand plain language?"), equating "our" language with clarity of understanding and intelligibility.

Similarly, there are various echoes in East South Slavic, where the usual autonymic language reference involves the use of the possessive pronoun *naš* 'our'; it is the usual linguistic autonym in Macedonian and on a more limited basis in Bulgarian. In Macedonian the substantival use of the possessive pronoun *naš*, i.e. simply *naš*, means 'Macedonian language' and so also with the definite neuter plural form *našata*. Moreover, the most usual form found is the adjectival *našinski* 'ours' (= 'our language'; *naški* also occurs) or the adverbials *po naše* (dialectal, cf. Hill 1991: 209) or *po našinski* (standard) 'in ours' (= 'in our language'), in each case thus with 'language' understood. The same is found in Bulgarian, where the expression *ponašenski* means 'in Bulgarian', though literally, "in ours".¹³

All of this taken together shows that embodied synchronically in the Albanian lexical item *shqip* are characteristics pointing to a long persistence, an echo, of a very old Indo-European linguistic ideology.

¹¹ Compare the use of *gather* in English for 'understand'.

¹² The verbs *shqipoj* 'to pronounce clearly' and *shqiptoj* 'to enunciate' are noteworthy here, as they refer to clarity of speech and are derivatives of *shqip*.

¹³ It usually refers to dialectal or regional speech.

3. PIE ‘Language’ as Body Part / Body Part as ‘Language’

In this second case study, the focus is on the Albanian word *gjuhë*, with the dialect form *gljuhë* (e.g., in Arvanitika) meaning ‘language’ but also ‘tongue’. This convergence of meaning is found in various Indo-European languages, as given below in (2), following Buck (1949: §18.24). Although the reconstruction of the PIE form for these words is tricky due to likely taboo deformations,¹⁴ the Latin, Germanic, Slavic, and Tocharian forms below (and others not shown here) are certainly true cognates amongst themselves in terms of form, and the Greek and the Albanian may well fit in here also; the Lithuanian form seems to belong with them too even if it has been re-shaped by association with the verb ‘lick’ (*liežiù*), as maybe also in Latin (cf. the verb *lingere*, cognate with Lithuanian *liežiù*), whereas the Hittite and Cuneiform and Hieroglyphic Luvian forms, while related, are clearly of a distinct origin.¹⁵

(2) Some Indo-European languages with ‘language’ = ‘tongue’

Greek	γλῶσσα
Latin	<i>lingua</i> (and all over Romance)
English	<i>tongue</i> (and elsewhere in Germanic)
OCS	<i>językŭ</i> (and elsewhere in South Slavic, with representation in West Slavic and at least some of East Slavic (e.g. Russian))
Toch A	<i>käntu</i>
Toch B	<i>kantwo</i>
Lith	<i>liežuvis</i>
Hittite	<i>lāla-</i> ‘tongue’ but also ‘speech’, especially ‘bad speech; slander’ but not ‘language’ per se
CLuvian	<i>lāla/i-</i> ‘tongue; gossip’
HLuvian	<i>lala(n)ti-</i> ‘language’ as a derivative of ‘tongue’ (<i>lali-</i>)

Languages like those in (2) can be referred to as “tongue”-languages because of the connection between ‘language’ and ‘tongue’.

Beyond the evidence of (2), there are Indo-European languages in which the word for ‘language’ is connected to something other than ‘tongue’; such “nontongue”-languages are given in (3), again, following Buck (idem):

(3) Some Indo-European languages with ‘language’ ≠ ‘tongue’

Old Irish	<i>berla</i> , from *bherH- ‘speak’, cf. Hamp 2004 on Slavic reflexes of this, e.g. Macedonian <i>zbor</i> ‘word’ (vs. Mod Irish <i>teanga</i> ‘tongue’), or possibly metathesis from <i>belra</i> ‘lips’ (Watkins 1970a)
-----------	---

¹⁴ This is not the place to rehash all of difficulties in the reconstruction of this word, see Hock (1991: 303-305) for some discussion. The PIE form may have begun simply with *d- and seems to have had a nasal and a velar or labiovelar in it, most likely with aspiration. But Hamp (p.c.) argued for initial *dl- with a sound change of *dl to *gl in Greek and Albanian, discussed briefly below. Beekes (2010: s.v.) does not connect the Greek word with the others listed here, for what it is worth.

¹⁵ Kloekhorst (2008: 516) states that the Anatolian forms cited here point to a Proto-Anatolian form that “is likely [to be] of onomatopoetic origin”.

Lithuanian	<i>kalba</i> , from the root seen in *kalbēti ‘speak’
Latvian	<i>valoda</i> , from the root seen in Polish <i>wolać</i> ‘call’
Sanskrit	<i>bhāṣā</i> , from the root *bhol- seen in Serbian <i>brbljiti</i> ‘chatter’)
Ukrainian	<i>movā</i> (with cognates in West Slavic, and maybe connected to Latvian ‘tongue’ <i>mēle</i>)

Given the data in (2) and (3), a question that naturally arises is what to reconstruct on the semantic side. In particular, should PIE be seen as a “tongue” or as a “nontongue” language? As far as Albanian is concerned, this question is asking whether Albanian is archaic in this regard or instead shares in an innovation with other languages.

The widespread occurrence of ‘language’ as ‘tongue’ as indicated in (2) makes reconstructing PIE as a “tongue” language seemingly quite straightforward. Slavic would then be retentive (conservative), and the ‘tongue’ Baltic forms would presumably be historically prior to the “nontongue” Baltic forms, even if the “nontongue” forms are ousting the “tongue” forms. If the Greek and Albanian forms fit with the others in (2), then Albanian also, as well as Greek, would be retentive in this regard.

However, one cannot always reconstruct based just on the most prevalent form; the dialectological distribution of the evidence needs to be taken in to consideration, and the linguistic geography of “nontongue” Indo-European languages might suggest PIE as “nontongue”. Important here is the fact that the “nontongue” pattern occurs in peripheral (lateral) and geographically somewhat isolated areas within the Indo-European family, namely Celtic and Indo-Iranian, so that that pattern may be archaic, in line with Bartoli’s principles of areal linguistics (see Bartoli 1925: Chapter 1). The “tongue” languages would then constitute an innovative core area (with Tocharian forming part of the core here, aligning with western languages, as in other respects). Baltic and Slavic ‘tongue’ forms would then be part of the innovative core, or else, it might be argued, they could be independent innovations based on a fairly natural sort of connection.

Still, taking Baltic ‘tongue’ forms as secondary, as Buck suggests, and Slavic ‘tongue’ forms also as innovative (on which see below), then the “nontongue” status of Balto-Slavic and Indo-Iranian could accord with other ways in which the two branches have been thought to form a larger grouping within IE, namely the so-called “satem” innovations of merging velar and labiovelar stops and the fricative outcomes of the palatal stops of PIE, and the *ruki* retraction; there would be no reason to privilege the “tongue” languages as a basis for a PIE reconstruction — they could all have developed the connection on their own, just as, in this interpretation, part of Baltic did, and maybe part of Slavic.

To push this somewhat further, it might be that PIE showed ‘language’ as deriving from a verb ‘speak’, as in Sanskrit, Lithuanian, and Irish; Slavic would then be innovative, moving away from an original (Balto-Slavic) situation seen still in Baltic.¹⁶ This situation would raise the question of how Slavic might have come to have “tongue” languages.

¹⁶ It is interesting that the verbs used in each of those languages have cognate forms in Slavic, allowing for speculation that Slavic might have started with a word for ‘language’ like one of them.

There are a few possibilities here. First, the linkage of ‘tongue’ with ‘language’ seems to be a somewhat natural connection; besides the many Indo-European languages that show it, it is found outside of Indo-European; it occurs in Semitic, as seen in Arabic *lisan* and Hebrew *lašon*, for instance. This means in principle that any Slavic language, and any of the other languages too for that matter, could have developed this semantic convergence on its own.

However, given that at least some early Slavs were in contact with two groups of speakers in the Balkans that *do* have the tongue/language connection, and that even have it with cognate material, namely Greek and Albanian with *glōssa* /g(l)juhě, respectively, one has to wonder if the tongue/language connection in Slavic is due to Greek or Albanian influence, maybe a South Slavicism that filtered into the rest of Slavic or else due to contact at an early enough stage to be part of Common Slavic. It is important to note here that these ‘tongue’ forms *could* be cognate, if from *dlng^{wh}- or the like, with the development of *dl to [gl] in both; the cognacy of Latin *dulcis* ‘sweet’ with Greek *γλυκύς* points to that outcome for Greek, and for Albanian, the word for ‘long’ offers important testimony, as it is *gjatë* (dialectally, *gljatë*), where *dl- is suggested by the Greek cognate *δολιχός*.¹⁷

In this interpretation, then, the Albanian polysemy of *gjuhë*, as both ‘tongue’ and ‘language’, would represent an archaism, and in particular, a very persistent and long-lived aspect of the semantics of this word, dating back to PIE.

By way of concluding this case study, there is another angle on “tongue”/“nontongue” languages that links the concept to the first case study; that is, whatever the formal reconstruction might be for this word, the tongue-language polysemy, or the lack of polysemy, reflects an element of a very early linguistic/language ideology. At issue here would be that viewing ‘language’ as ‘tongue’ would mean that PIE speakers characterized language in terms of its visible means of production, whereas viewing ‘language’ as ‘speech’ or something more abstract connected with talking would entail characterizing language in terms of its most obvious product or output modality. In either case, such a characterization would be a reflection of how PIE speakers conceived of what “language” or what “a language” is. This perspective on the Albanian present-day polysemy would mean that echoes of a PIE folk taxonomy would be reverberating still in Albanian *gjuhë* and possibly in other Indo-European languages with similar semantics.

4. PIE Concern for Source of Information (“Evidentiality”)

The final case-study concerns the Albanian expression *marr vesh* ‘to understand; to come to an understanding’. I argue that it must be seen against a backdrop of a concern within PIE for information-source, what is generally referred to as *evidentiality*. This notion refers to a linguistic indication of the source of a speaker’s information, the modality by which that information was gained, and/or the speaker’s stance (i.e., attitude) towards the truth of the information.

¹⁷ Presumably the *dl- cluster in ‘long’ was split up in Greek before the *dl > γλ change, by whatever sound or sequence of sounds — perhaps involving a laryngeal — would have led to the separation of the *d and the *l.

It must be admitted from the start that there is no basis for positing a grammatical evidentiality-marking system for PIE. While there are Indo-European languages with well-developed systems for marking evidentiality, such as Albanian and Macedonian, it is clear that these are relatively recent developments within those languages. Nonetheless, there are various indications amongst the Indo-European languages of information-source mattering in some way or another.

For one thing, there is a relationship between the meanings ‘know’ and ‘see’ within and across various languages. For instance, the root **k^weyt-*, glossed in Rix (2001) as ‘bemerken, erkennen’ (‘notice, recognize’), shows the meanings in (4):

(4) Meanings of **k^weyt-*

- ‘recognize, realize’ (as in Avestan *cit-*)
- ‘count, honor, read’ (as in Slavic *čit-*)
- ‘see’ (as in the Sanskrit derivative *citana-* ‘coming into appearance’)

Thus there are elements of both seeing and knowing in the cognate forms for this root across the family, so that one might speculate that this root was originally ‘know’ through visual evidence, which would make it especially appropriate for the later meaning ‘read’ in Slavic since sight is involved in reading. And, in the case of the root **weyd-*, the meaning ‘see’ is evident in the Greek aorist *εἶδον* ‘saw’ and the meaning ‘know’ in the perfect *οἶδα* ‘I know’, a form that has present reference, as also in the exactly equatable Gothic *wait* ‘I know’ and Sanskrit *veda* ‘I know’), a grammatically based distribution which is usually explained as ‘to have seen’ something (perfect) implying that one ‘knows’ it (in the present). Moreover, though, ‘know’ and ‘see’ are distributed across the languages with this root in ways that are not always connected to verbal tense, as in Greek. Bearing in mind that other metaphors apparently were involved in the development of ‘know’ meanings within Indo-European — e.g. Latin *sciō* / Hittite *sekk-/sakk-* ‘know’ are possibly from a root **sekH-* (Rix 2001: 524), originally meaning ‘cut; distinguish’ (seen in Latin *secō* ‘cut; mow’) — the connections between ‘see’ and ‘know’, even if grammatically motivated in some cases, are suggestive of an interest on the part of PIE speakers in the modality by which knowledge is gained.

Furthermore, there is evidence of a cultural concern for knowledge-source and veracity in various forms that the root **H₁es-* ‘be’ takes. For instance, within the context of Indo-European legal language (cf. Watkins 1967, 1970b, 1987), derivatives of **H₁es-* come to have meanings tied to matters of evidence (in a legal sense), especially Sanskrit *satya-* ‘true; truth’, Latin *sons* ‘guilty’, Old Icelandic *sannr* ‘true; guilty’, Hittite *asan* ‘(it) is (so)’ (in public confession); these forms can be taken to mean that ‘be’, at least in legal language, could mean ‘must be’ or ‘be evident’, with ‘truth’ as one side of what the evidence shows things to be and ‘guilty’ as the other (cf. Benveniste 1960 on ‘be’ as originally ‘really, actually be, exist’).

In this regard, an Indic distinction seems relevant here, namely the difference between *śruti-* and *smṛti-* as two types of knowledge, gained in different ways. *śruti-* is a derivative of the root *śru-* ‘to hear’ and *smṛti-* is a derivative of the root *smṛ-* ‘to remember’. The distinction is described by Monier-Williams (1899) in his account of *śruti* (s.v.) and he defines *śruti* as “sacred knowledge orally [and thus aurally/BDJ] transmitted by the Brāhmins from generation to generation ... and so differing from *smṛti-* or what is remembered and handed down in writing by human authors”. Thus the

distinction is roughly ‘authoritative knowledge that has been heard and taken in that way’ versus ‘authoritative knowledge that has been handed down in other ways’. This distinction can therefore be seen as a further indication that source of knowledge was culturally significant, and if so for the ancient Indo-Aryans, perhaps so too for PIE speakers.

Finally, then, within the scope of a concern for evidentiality and knowledge source, we can examine the Albanian idiom *marr vesh*, meaning ‘understand’. It is a combination of *marr* ‘take’ and *vesh*, the word for ‘ear’, so that it is literally “take ear”. As such, against the backdrop of suggestive evidence for a concern for information-source in PIE, *marr vesh* would seem to show a modern-day echo — a persistence — of a more ancient importance placed on ‘hearing’ as the source for knowledge, in line with the Indic distinction between *śruti*- and *smṛti*-, meaning in essence that heard knowledge in this cultural context is the key to (true) understanding.

5. Concluding Remarks

From a methodological standpoint, the fact that aspects of meaning can persist in traces and only in certain uses means that it is not enough in studying semantic change to just go on dictionary definitions; rather, it is important to consider as full a range of uses as possible for any item under consideration. Moreover, from an Indo-European and Albanological perspective, these echoes confirm that there is likely to be a greater Indo-European component in Albanian than is usually thought, even if evident just in traces and detectable only with some degree of digging. The nuggets one finds when digging can make the exercise particularly worthwhile.

References

- Baldi, P. 1983. *An Introduction to the Indo-European Languages*. Carbondale: Southern Illinois University Press.
- Bartoli, M. 1925. *Introduzione alla Neolinguistica. Principi—Scopi—Metodi*. Modena: L.S. Olschki.
- Beekes, R. 2010. *Etymological Dictionary of Greek*. Leiden: Brill.
- Benveniste, É. 1960. *Etre et avoir dans leur fonction linguistique*. *Bulletin de la Société de Linguistique de Paris* 55.114ff. (reprinted in *Problèmes de linguistique générale* I: 187-207 (Paris: Gallimard)).
- Buck, C.D. 1949. *A dictionary of selected synonyms in the principal Indo-European languages: A Contribution to the History of Ideas*. Chicago: University of Chicago Press (reprinted in paperback, 1987).
- Cardona, G. 1990. *On Attitudes towards Language in Ancient India*. (Sino-Platonic Papers 15.) Philadelphia: University of Pennsylvania.
- Colvin, S. 1999. *Dialect in Aristophanes: the Politics of Language in Ancient Greek Literature*. Oxford: Oxford University Press.
- Fortson, B.W. IV. 2010. *Indo-European Language and Culture. An Introduction* (2nd ed.). Chichester: Wiley-Blackwell.
- Hamp, E.P. 1999. *Lectures on the Albanian Language: History and Dialectology*. The Ohio State University, November 29-December 4.
- Hamp, E.P. 2004. 360p ‘word’: An Un/Misrecognised Indo-European Root in Slavic. Macedonian Studies. In B.D. Joseph & M.A. Johnson (eds.), *Papers from the 5th International Macedonian-North American Conference on Macedonian Studies 1-4 May*

- 2003 at The Ohio State University (*Ohio State Working Papers in Slavic Studies* 4). Columbus: The Ohio State University Department of Slavic and East European Languages and Literatures, 103-106.
- Hill, P. 1991. *The Dialect of Gorno Kalenik*. Columbus: Slavica.
- Hock, H.H. 1991. *Principles of Historical Linguistics* (2nd ed.). Berlin: Mouton de Gruyter.
- Hopper, P. 1991. On some principles of grammaticization. In E.C. Traugott & B. Heine (eds.), *Approaches to Grammaticalization*. Amsterdam: John Benjamins, Vol. 1: 17–36.
- Joseph, B.D. 2003. Evidentiality in Proto-Indo-European? Building a Case. In K. Jones-Bley, M. Huld, A. Della Volpe, & M. Robbins Dexter (eds.) (*Journal of Indo-European Studies Monograph Series*, No. 47) *Proceedings of the Fourteenth Annual UCLA Indo-European Conference (Los Angeles, November 8-9, 2002)*. Washington, DC: Institute for the Study of Man, 96-111.
- Joseph, B.D. 2005a. Preliminary thoughts on Proto-Indo-European language ideology. *Paper presented at 2nd Martin Luther King Day Symposium on Linguistics: Studies in Ancient Indo-European Languages*, Ohio State University, Columbus OH, 17 January 2005.
- Joseph, B.D. 2005b. 'Tongue' and 'nontongue' Indo-European Languages: Where Slavic Fits in. *Paper presented at Midwest Slavic Conference*, Columbus OH, 5 March 2005.
- Joseph, B.D. 2005c. More Balkan Archaisms: Grammar, Phraseology, and Proto-Indo-European Language Ideology. *Paper presented at 24th Annual East Coast Indo-European Conference (ECIEC)*, University of California, Berkeley, 3 June 2005.
- Joseph, B.D. 2010. Arkaizmat semantikë të Shqipërisë ['Semantic Archaisms of Albanian']. In *Seminari ndërkombëtar për gjuhën, letërsinë dhe kulturën shqiptare* 28/2.159-160 (Universiteti i Prishtinës, Fakulteti i Filologjisë).
- Joseph, B.D. To appear. Here's to a Long Life! Further Echoes of Indo-European Semantics in Albanian. To appear in *Variation and Contact in the Ancient Indo-European languages: between Linguistics and Philology*.
- Kloekhorst, A. 2008. *Etymological Dictionary of the Hittite Inherited Lexicon*. Leiden: Brill.
- Lloshi, X. 1999. Albanian. In U. Hinrichs (eds.), *Handbuch der Südosteuropa-Linguistik*. Wiesbaden: Harrassowitz Verlag, 277-300.
- Meyer, G. 1891. *Etymologisches Wörterbuch der albanesischen Sprache*. Strassburg: Trübner.
- Monier-Williams, Sir Monier. 1899. *Sanskrit-English Dictionary*. Reprinted (1976), New Delhi: Munshiram Manoharlal Publishers.
- Rix, H. 2001. *LIV. Lexikon der indogermanischen Verben*. Wiesbaden: Dr. Ludwig Reichert Verlag.
- Silverstein, M. 1979. Language Structure and Linguistic Ideology. In P.R. Clyne, W.F. Hanks & C.L. Hofbauer (eds.), *The Elements: A Parasession on Linguistic Units*. Chicago: Chicago Linguistic Society, 193–247.
- Watkins, C. 1967. Latin sons. In W.W. Arndt, P.W. Jr. Brosman, F.E. Coenen & W.P. Friedrich (eds.), *Studies in Historical Linguistics in Honor of George Sherman Lane*. Chapel Hill: University of North Carolina Press, 186-194.
- Watkins, C. 1970a. Language of Gods and Language of Men: Remarks on Some Indo-European Metalinguistic Traditions. In J. Puhvel (eds.), *Myth and Law among the Indo-Europeans. Studies in Indo-European Comparative Mythology*. Berkeley: University of California Press, 1-17.
- Watkins, C. 1970b. Studies in Indo-European legal language, institutions, and mythology. In G. Cardona, H. Hoenigswald & A. Senn (eds.), *Indo-European and Indo-Europeans*. Philadelphia: University of Pennsylvania Press, 321-354.
- Watkins, C. 1987. 'In the interstices of procedure': Indo-European legal language and comparative law. In W. Meid (eds.), *Studien zum indogermanischen Wortschatz*. Innsbruck: Innsbrucker Beiträge zur Sprachwissenschaft, 305-314

