AMSTERDAM STUDIES IN THE THEORY AND HISTORY OF LINGUISTIC SCIENCE

General Editor
E.F. KONRAD KOERNER
(Zentrum für Allgemeine Sprachwissenschaft, Typologie und Universalienforschung, Berlin)

Series IV – CURRENT ISSUES IN LINGUISTIC THEORY

Advisory Editorial Board

Lyle Campbell (Salt Lake City); Sheila Embleton (Toronto)
Brian D. Joseph (Columbus, Ohio); John E. Joseph (Edinburgh)
Manfred Křížka (Berlin); E. Wyn Roberts (Vancouver, B.C.)
Joseph C. Salmons (Madison, Wis.); Hans-Jürgen Sasse (Köln)

Volume 254

Philip Baldi and Pietro U. Dini (eds)

Studies in Baltic and Indo-European Linguistics.
In honor of William R. Schmalstieg.
"TO BE" OR "NOT TO BE" IN THE INDO-EUROPEAN LANGUAGES

XAVERIO BALLESTER
Universidad de Valencia

A verb not for all seasons

In most Indo-European languages we find at least two different roots for "be": one for the present tense ("I am": Gothic im, Greek εἰμί, Latin sum, Lithuanian esuėsmi, Old Church Slavonic jesmī, Sanskrit asmi, or Vedic ásmi ...) and another for the past tense ("I was": Latin fui, Lithuanian buvau) or for the non-present tenses ("I shall be": Lithuanian būsiu). For many forms, especially in non-present tenses, we can reconstruct a Proto-Indo-European root *bau- (or eventually *bau-), as we see in Gothic bauan "dwell" (and similar examples, Alinei 1996:598), Irish buith "be", Latin fuiisse "have been", Lithuanian būti "be", Old Church Slavonic byti "be, become". Originally the root had nothing to do with "being", since it is easy to recover the meaning "become, transform, change, be born, grow, bud" for this root (Greek φυω "I bring forth, beget", Sanskrit bhavāmi "I become"). Thus it seems obvious that many Indo-European languages used *bau- "become" independently as a suppletive form, especially in non-present tense, for "be". But this was not the only instance: other verbs, meaning "stand", "remain", or "sit", have likewise furnished historically complementary forms of "being" (Buck 1951:635).¹

No doubt the semantic similarity and the abstractness of "becoming" made it complementary to "being". Indeed, the semantic connection between "to be" and "to become" is not problematic at all. Ainu, for instance, does not distinguish between "being" and "becoming" (in the sense of changing into something else) and the meaning of ne can be static as well as dynamic (Refsing 1986:145); in Ainu, as an independent verb, an means "be, exist", but when suffixed to temporal expressions, it means "become (that time)" (Refsing 1986:147). With regard to the original (and therefore verosimmille, non-abstract) meaning for *bau-, it is an attractive hypothesis to connect the verb with the meaning "earth" that we find in Sanskrit bhūḥ (or Old Persian būmī-), a

¹ I am deeply thankful to Mila del Saz and Robert Quinn for reviewing my English.

noun that probably belongs to the same root *bau- (Alinei 1996:538). In such a case, the original meaning of *bau- could be that of "bud", a sense that soon was extended metonymically to "nature, change, transform(ation)", as is still clear in the Greek ψυχή "nature, birth", and φυτόν "germ, sprout". A parallel could be offered by the Lithuanian tapti "become", in case we accept a connection with Greek τόπος "place" (Buck 1951:637).

The suppletion strategies

As a logical result of all these circumstances, many languages use different roots, supplementing each other, for "be": one root being the true "be", which is defective, and the other root (or more than one), having contiguous senses, which is suppletive. The commonest suppletive roots are those meaning "become", "stand", "remain", "sit", or the like. Let us have a look at some examples from all five continents, presented here in the archaeologically recorded order for the diaspora of modern humans (Africa, Asia, Oceania, Europe, and America).

In Kisi, we find co and wa used as copula verbs, but while co is used only for the present tense and especially for the realis, wa is employed for the past, especially for the irrealis and some other distinctions (Childs 1995:120–122). In Bambara, "be" is covered by several words (Bailleul 1998:114–115). In Amharic, the copula has only non-past forms and the negative non-past copula is formed on a stem different from that of the affirmative, while existence is expressed in the past tense by another stem and in the future tense by hon "be, become" (Hudson 1997:475). Likewise, Argobba shows a similar situation expressing future with hon "be, become" (Hudson 1997:476). Ge'ez uses pronouns as copulas in the present tense, but the verbs konä and hålläwä for other tense or modal distinctions (Gragg 1997:260). In Tigre, besides the proper copula—i.e., with a merely identificational sense—there is an existential verb halla for present and ala for past (Campbell 2000:1655). In Harari, the copula is ta, but hal is the verb of existence, with na:ra for the past and e:l for negative (Wagner 1997:507). Indeed, as we shall see, many languages use a special form to express the negative copula. In Outer South Ethiopic, the copula and the locative-existential verb have a suppletive system (Hetzron 1997:548). In the Silte group, in the present tense the copula takes a suffix but the copula suffix occurs with naara in the past and the verb hoona "to become" is used for the future (Gutt 1997:532). In Tigrinya, the verbs näbärä and konä replace the copula in the past and in the future (Kogan 1997:444). The Berber dialect from Figuig uses four different constructions in order to express "be", depending especially on the time referent (Kossman 1997:360). In Uzbek, "be" is expressed by
means of copula suffixes, but with different roots for past tense, conditional, and
indirective moods (Boeschoten 1998:363). In Chuvash, pur means “there
is/are”, but the negative is suk (Clark 1998:449). In Chaghatai copula suffixes
are mostly attached to forms of tur- “stand” (Boeschoten & Vandamme
1998:171). In the Ob-Ugrian languages, the copula-like verb is phonologically
similar, but unrelated, to the verb “is” (Honti 1998:353). In Mari lij- “become”
provides a common suppletive stem for “be” (Kangasmäa 1998:230, 240). A
system of two verbs meaning “be” is discernible in most Finno-Ugric
languages, but the functions of the two verbs have become historically tangled
(Kangasmäa 1998:240). In the Turkic languages bol- “become” is a common
suppletive stem for the copula (Johanson 1998:42). In Mongolian there are dif-
ferent copulas both in positive and in negative (Peyró 2000:114–117). In Dumi
there are two verbs “be”, both covering the existential, attributive, and loca-
tional senses of English “be”: gin, used exclusively with inanimate referents,
and mini, used exclusively with animate referents; both verbs are sometimes
negated by forms that are suppletive as well (Van Driem 1993:168–170). In
Tibetan, the copula is yin in the I person and red in second and third persons,
while yod is the existential verb in I person and ’dug or yod in second and third
persons (Campbell 1995:539). In Ladakhi, the copula is dug or yod, but med for
negative (Campbell 2000:933). Duan Nu uses different roots for positive copula
and for negative copula, and a different form for the locational sense as well
(Campbell 2000:1126). In Brahui u[t “be” is supplanted by man “become” (Elfen-
bein 1998:400). In Middle Persian both būtan “become” (from *bau-) and
ēstātan “stand” are used as suppletives of “being”. The verb “stand” (sthā-) also
functions as a suppletive stem for “be” in the Dardic languages (Skalmowski
1986b:234), while Parachi uses “sit (down)” (Skalmowski 1986a:182). In Indo-
nesian, ada is used in existential (“there is”) and locational senses, and adalah
in attributive (and long) sentences (Kwee 1976:135–137). In Lezgian, ḵun
“become” furnishes a frequent suppletive form for “be”, especially in the future
tense (Haspelmath 1993:136–137, 312). In Portuguese and Spanish, ser (from
Latin esse “be”) denotes permanent or essential properties, while estar (from
Latin stare “stand”) denotes a temporary situation. In Portuguese and Spanish,
some forms of the subjunctive of “be” come from Latin sedere “to sit (down)”. The
paradigm of “be” was also typically suppletive in Cornish, with a present
yu, a preterit bu, and an imperfect olesa, all for the third person (Campbell
2000:419). The situation was very similar in many other Celtic languages. In
Pipil, nemi is the basic verb for “be” but there is also a variety of constructions
which express different notions of “being”, nemi being used in locational and
existential senses (Campbell 1985:110–112).

Thus, it is not so strange to find an analogous phenomenon in Proto-Indo-
European, where “be” was expressed by different roots, but with the important corollary that the true and proper meaning of “being” could not be in *bau-, which clearly meant “become” and was the suppletive—not the defective—form. Thus we can conclude that the meaning of “being” was probably in the root often used in the present tense.

The irregularity strategies: The zero copula

But true “be”—that is, the copula—displays not only some odd lexical peculiarities, such as the use of different stems, but also some odd morphological peculiarities as well, for example a clear tendency to irregularity even in the paradigms of one single stem. For example, Tswana has a cliticized form of the verb “to be” in the indicative present (Creissels 2000:255). In Tigrinya, the copula paradigm is very irregular in the present tense (Kogan 1997:444). In Somali, the main word for “be” is a somewhat irregular verb (Orwin 1995:206–208). In Sumerian, the copula me has neither aspect nor tense (Jiménez 1998:80). In Limbu, one of the verbs meaning “be”, the identity operator, has a defective conjugation and constitutes a unique defective, tenseless conjugation consisting solely of a set of adnominal suffixes (Van Driem 1987:55, 56). In Kashmiri, the copula distinguishes gender in all persons, both in singular and plural, and has independent forms for present, past, and future (Skalmowski 1986b:225). “To be” is also irregular in Agul with a present ile, a present negative davai, a past idi with variants, and a past negative dawadi (Campbell 2000:26). In Lithuanian, yra, the copula for third-person of present “be”, is clearly unconnected with the rest of the paradigm (esu, esi, esame, esate and proper third-person esti). The paradigm of “be” was typically irregular in Cornish with a present of, os, yu (sing.), on, ough, yns (plur.), and an impersonal form or (Campbell 2000:419). The situation was very similar in many other Celtic languages; in Breton, for instance, bezañ “be” is very irregular. The verb kaia “to be” is likewise irregular in Miskito (Campbell 2000:1135).

However, perhaps the most conspicuous—and common—irregularity that we can find in the verb “be” is what we may call the zero copula, an extreme case of the defective form, since we find no “be” at all, especially in present tense (especially in third person). Indeed, a word for “be” or “is” would appear to be less common in the world’s languages than it seems at first sight.

In Bisa, for instance, there is no copula and two things equated or otherwise identified are simply placed side by side (Kropp & Naden 1988:159). A rather typical situation in Turkic languages is that “be” in the copulative sense is expressed with (enclitic) personal markers attached to the predicate, but with ø marker for the third person singular, as, for example, in Turkish (Csató &
Johanson 1998:226), or with ø marker also for the third person plural in Kirghiz (Kirchner 1998b:349), Kazakh and Karakalpak (Kirchner 1998a:324), or in Yakut (Stachowski & Menz 1998:430). These Turkic constructions may also occur without a copula marker, as in Tatar and Bashkir (Árpád 1998:298) or Noghay (Csató & Karakoç 1998:341). In Old Turkic there were also, though not always, sentences without a copula (Erdal 1998:152). In Komi, in the present affirmative the copula is zero (Hausenberg 1998:320). In Mari, the copula is zero in the present indicative third person singular (Kangasmaa 1998:230, 240). In Limbu, the most used form of suffixal “be” (the so-called identity operator) in the third person singular is simply zero (Van Driem 1987:56). In Dumi, “be” as an identity operator—whereby two noun actants, whether animate or inanimate, are equated—is conveyed by zero (Van Driem 1993:173). Adjectives need no copula as predicates in Khmer (Sacher & Phan 1985:27) and in Korean (Chang 1996:47). In Chinese, the copula shì is only optional in positive sentences. In Mongolian, the copula is normally omitted (Peyró 2000:114). In the Dravidian languages, sentences with predicate nominals are normally found without a copula, and copular verbs are usually copied from neighbouring Indo-Aryan languages (Steever 1998:29), as, for example, in the Wardha dialect of Kolami (Subrahmanyan 1998:323, 326). In Sanskrit, the copula can be omitted in the present tense. In Indonesian, the copula can be left out in short sentences (Kwee 1976:8). There is no equivalent of the verb “to be” in the copulative function—and thus no copula—in Kilivila (Senft 1986:36). Kambera likewise has no separate copula verb (Klammer 1998:107). We find also zero copula in Russian in the present tense. South Saamic can also have purely nominal predicates (Sammallahti 1998:85). In Dakota, the copula is used to identify a definite referent, and is not used with descriptive adjectives nor with personal pronouns (Campbell 2000:455–456). In Nahua or Pipil, equational constructions are formed with affixes when a pronominal subject has a noun or adjective as complement, and nemi “be” is used in these constructions, in order to show more emphasis (Campbell 1985:55, 108, 111–112). In Guaraní, there is no proper copula in the attributive sense. Instead, there is only juxtaposition.

A flexible concept (if a concept) as a purely philosophical matter

Let us admit that “being” is a rather flexible concept. First at all, “to be” must include a very special notion, since it displays such a peculiar and bizarre morphology all over the world. If we include the meaning of “existing” in the main semantic field of “being”, as it seems logical and actually many languages do, we find a similar situation, although theoretically “to exist” should be less abstract a notion than “being”. Indeed the existential meaning “there is” may be
expressed by monolexemic verbs, such as hay (haber) in Spanish, but existence is more often expressed by the same verb which serves as a copula, or by verbs which can also mean “stand”, “live”, “give” (German es gibt, Spanish se da), or even “see” (Old Irish fi). In some languages, “there is” is expressed not by a verb at all but by a particle, or even by a determiner-like element (Goddard 2001:34). Ge’ez uses an expression meaning literally “in it” to signify “exists” (Gragg 1997:260). Maasai makes a distinction between ará “to be something” and áti “to be somewhere” (Campbell 2000:1021). As was mentioned above, in Harari, the copula is ta, but hal is the verb of existence (Wagner 1997:507). In the present tense especially, many languages employ a word different from the copula for the existential sense, such as allo in Tigrinya (Kogan 1997:444), la (negative eela) in the Silté group (Gutt 1997:533), or yeš (negative eyn) in Modern Hebrew. In Limbu, there are seven verbs which cover various senses (identificational, existential, negative existential, locational, adhesive, attributive, inchoative) of English “to be” (Van Driem 1987:55). In Guaraní, there are two verbs meaning “to be”: locational aime, and aiko, which is used with an adverb (Campbell 1995:207). Maya has no proper verb “to be”, but rather only an existential expression yan with locative and possessive senses (Raga 1995:23).

What then is the real meaning of “being”? Evidently, the subject constitutes a purely philosophical matter. Many philosophers have discussed the topic vehemently for many centuries, and in general they have not found much agreement. This circumstance is quite paradoxical for linguists, since we find the copula used from the first written records in many languages of the world, and this fact can only mean that normal humans have been using something equivalent to “be”, probably from the beginning of speech, although they might be not fully aware of its real meaning or etymology. Yet an abstract and philosophical meaning for “be” is very unlikely to be original and not only for the beginnings of speech, but also over the course of human history. This paradox is very attractive for all those linguists who, like Afinei (1996, 2000), Cavazza (2001), or Costa (2001), have supported the idea that Proto-Indo-European originated as a linguistic group in the Upper Palaeolithic. I personally believe that the Proto-Indo-European group, and most other linguistic groups or perhaps even all of the linguistic groups in the world, were already formed by the Upper Palaeolithic, and it seems obvious that such a purely philosophical meaning was not a priority for the needs of Palaeolithic mentality. In this primitive context, we cannot expect to find anything similar to “be”, at least in the sense used by Parmenides or Heidegger. So, where do all those “to be” forms come from? Maybe the problem does not concern pure ontology at all, but just mere morphology.
"TO BE" OR "NOT TO BE" IN THE INDO-EUROPEAN LANGUAGES

A morphological Proteus: demonstrative

An answer that could suit all the characteristics of "be" in Indo-European (and in many other linguistic groups) is: the only being in "to be" is a banal demonstrative. A demonstrative origin for the copula is historically well documented in many languages; therefore, theoretically this explanation could also be valid prehistorically for reconstructed languages or languages without written records. The fact is that "copulas often arise from demonstratives and third person pronouns" (Diessel 1999:33) and that we probably have few other places to look, since no other origins are as clearly evident for "be", and most speakers in the world do not behave like Parmenides, Spinoza, Leibniz, or Heidegger. But we must emphasize that demonstratives are the ultimate source of copulas, because third person pronouns very often arise from demonstratives. This shift from a demonstrative to an anaphoric (or personal) pronoun or eventually to an article is well documented in many languages. It would be enough to mention French il - elle, Portuguese ele - ela, Spanish él - ella, and so on, all from the Classical Latin demonstratives ille - illa, but let us include some other more exotic languages, such as Sayhadic, where the forms of the third person pronouns are identical to the nominative forms of remote deictics (Kogan & Korttayev 1997:224), or Chechen, Ossetic, and Ubykh, where the third person pronoun is a demonstrative (or vice versa), or P'urhépecha, where the demonstratives are directly used to express the third person singular pronouns (Monzón 1997:60). A good, brief survey of this topic for the third person pronoun is also offered by Diessel (1999:21). In any case, the use of demonstratives, pronouns, or eventually articles as copulas is well documented.

In Kilba, there are three nonverbal copulas, which have the same form as identificational demonstratives in nonverbal clauses, while they differ from demonstrative pronouns. The demonstrative pronouns are complex free forms, while the identificational demonstratives and non-verbal copulas are monosyllabic enclitics (Diessel 1999:36). In Ge'ez, the copula is expressed by independent pronouns for simple present tense predication, and sometimes the singular third person masculine pronoun can function as a default copula, even for other persons, both non-singular and non-masculine (Grigg 1997:260). In Tigre, the copula is the same as the independent personal pronouns, except in the third person, where the prefixed ha- is dropped, leaving tu, ta, tom and tan, according to grammatical gender and number (Campbell 2000:1655). In Harari, the copular verb ta shows a stem which is typical for deictic pronouns in many languages. In some Arabic dialects, pronouns of the third person usually serve as copulas or are the base for copulas in the present (Kaye & Rosenhouse 1997:303). In Modern Hebrew, the reanalysis of third person pronouns as copulas is a very recent development, but Modern Hebrew not only has copulas that
are derived from personal pronouns, but also has a set of non-verbal copulas that developed from the demonstratives. These demonstratives are still used as such (Diessel 1999:34). The Eastern Neo-Aramaic languages possess an inflected copula which is used with non-verbal predicates and harks back to personal pronouns partly supplemented with a particle (Jastrow 1997:372). In Chaghatay, the personal pronouns may, in cliticized form, be used as copula suffixes with present meaning, but no copula is required in the third person, because the use of the demonstrative pronoun ol as a copula became largely obsolete in the Kwarezmian Turkic period (Boeschoten & Vandamme 1998:171). In Pashto, the third person of present tense for “be” can be traced back to the old demonstrative *aita- (Skalmowski 1986a:188). In Ossetic, we can still trace the copulative elements for present tense a- back to old demonstratives (Skalmowski 1986b:204) and -γ in Yagnob (Skalmowski 1986b:209) as well. Since Ainu has an anaphoric root documented in ene or nean and an interrogative-indefinite root documented in nekon(ka) or nep(ka) (Refsing 1986:109), we cannot exclude that the copula ne is a cliticized form of the anaphoric demonstrative in Ainu (Refsing 1986:110). In Tolai, the existential meaning is expressed by means of the definite article a in a verbless sentence (Goddard 2001:34). In Tasmanian, the demonstrative wa “this” is very likely to have occurred as a copula too (Campbell 2000:1618). In Maltese a demonstrative followed by a noun can have the value of copula (Aquilina 1965:97). In Polish, the epideictic to is a kind of colloquial substitute for “be” in many constructions. Pietro U. Dini kindly reminds me of analogous uses of tai (“this, that”; cf. tai viskas “that’s all”) and čia (“here”) in Lithuanian. In Guarani, the equivalents to the verb “be”, aime and aiko, used in the non-attributive sense, could be related to the demonstrative root of aipo(ν)a “that (unseen and unknown)”.

The Indo-European case

Almost all of the typological data support the hypothesis that a demonstrative could be the origin of the Proto-Indo-European proper “be”. In order to express person and number and later aspect or tense, several affixes would be added to these demonstrative roots in Proto-Indo-European. Naturally, this demonstrative “be” could hardly cover as vast a paradigm as proper verbs did, and when writing was invented and regular grammar notions appeared in many linguists’ consciousness, some other roots—such as *bau- notoriously—were considered to be suppletives of “be” in different circumstances (aspect, tense, mood ...).

Elsewhere, we have attempted to show that a set of demonstrative roots like
*ha-, *ia-, *ta-, *ua- … is present in most languages of the world, and we have emphasized the importance of demonstratives for morphological development. The existence of such a common *ha-, *ia-, *ua- demonstrative series (mostly pronounced [ha ja waj]) could be explained because [h j w] represent the natural consonant partners of the cardinal vowels /a i u/ respectively, in the basic consonant-vowel pattern and preceding /a/, which as the vowel par excellence has no such unanimous consonant counterpart ([ʔ] and [x] and other velar consonants can function likewise as the consonant partner of the velar vowel /a/). All of this suggests a rather primitive situation that we can trace back for Proto-Indo-European, as well as for many other linguistic groups. In the Indo-European case, a root *ia- or especially *ha- could be the base for the present of "be" in most languages, but even *ua could be attested as another supplementary root for "being" in Germanic languages ("be" Gothic wisan, Old High German wesan, Old Norse vera; cf. Sanskrit vas- "abide, dwell"). The Proto-Indo-European copula could be formed by the addition of personal endings to a demonstrative root in -s (e.g. *ias-amî) or to demonstrative roots (e.g. *ia-sa-mî), since agglutination is common for demonstratives (cf. French celui-ci, celui-là, Spanish aquel, etc.), and *sa- is a typical demonstrative root for Indo-European animates. In any case, if we have to speak strictly, there was never a proper "to be" in Proto-Indo-European. That’s the question.

REFERENCES


