The construct form of nouns in African languages: a typological approach

Denis Creissels
Université Lumière (Lyon2)
e-mail: denis.creissels@univ-lyon2.fr

1. Introduction

The question addressed here is the possibility to extend the notion of *construct state* of nouns, traditional in Semitic linguistics, to the description of languages belonging to other families. In African linguistics, this term has been used mainly in descriptions of East African languages belonging to the Nilotic and Cushitic families (see for example *Andersen 2002* on Dinka, *Mous 1993* on Iraqw). The main aim of this paper is to propose some terminological clarifications in order to lay the foundations of a cross-linguistic study of morphosyntactic phenomena likely to be viewed as particular manifestations of the same type of mechanism as the construct state of Semitic languages.

2. Definition

In Semitic linguistics, the term of construct state applies to nouns immediately followed by another noun in the role of genitival modifier, or by a bound pronoun in possessive function. In Arabic, or in Hebrew, this context triggers the use of a special form of nouns.

For example, in Classical Arabic, the most general characteristic of nouns occurring in this context is the absence of definiteness marking, but a few nouns undergo additional modifications in the construct state. Note that nouns are NOT in the construct state, for example, when followed by a preposition phrase.

(1)   *Classical Arabic*[1]

   a. *daxal-a   kalb-u-n*
      enter.PF-3SM dog.SG-ABS-INDEF
      'A dog came in'

   b. *daxal-a   l-kalb-u*
      enter.PF-3SM DEF-dog.SG-ABS
      'The dog came in'

   c. **daxal-a   kalb-u*
      enter.PF-3SM   dog.SG-ABS

---

[1] ABS = absolute case ('nominative'), DEF = definite, GEN = genitive, INDEF = indefinite, PF = perfective, SG = singular, 3SM = 3rd person singular masculine.

d. *daxal-a   kalb-u   l-malik-i*   (*… kalb-u-n …, *…al-kalb-u…*)
   enter.PF-3SM dog.SG-ABS DEF-king-GEN
   'The dog of the king came in'

e. *daxal-a   kalb-u-hu*
   enter.PF-3SM dog.SG-ABS-3SM
   'His dog came in'

f. *daxal-a   kalb-u-n   li-l-malik-i*   (*… kalb-u …*)
   enter.PF-3SM dog.SG-ABS-INDEF to-DEF-king -GEN
   'A dog belonging to the king came in'

In Modern Arabic dialects, as a result of changes in the system of definiteness, the absence of definiteness marking is not relevant to the characterization of the construct state, but a subset of nouns take a special form when immediately followed by a genitival modifier or a bound pronoun, so that the notion of construct state itself is still relevant.

(2)   *Moroccan Arabic (Harrell 1962)*[2]

a. *xala*   '(paternal) aunt'

b. *xalt-i*   'my aunt'
   aunt.CSTR-1S

c. *xalt   l-bənt*   'the girl's aunt'
   aunt.CSTR DEF-girl

Cross-linguistically, the existence of a morphological marking of nouns fulfilling the role of head in a head-dependent construction is not very common, nevertheless it is not limited to the Semitic languages. The range of noun dependents triggering the choice of a special form of their head vary among the individual languages that have this phenomenon, but it has already been proposed by several authors to borrow the term of *construct state* from Semitic grammars as a general term for noun forms characterizing nouns assuming the role of head in combination with certain types of dependents.

My proposal is to apply this definition to *construct form* rather than *construct state*. The point is that it is not entirely clear wether, in the Arabic or Hebrew grammatical traditions, *construct state* primarily refers to the construction itself, or to the particular form taken by the head noun in this construction. But this ambiguity can easily be avoided by using the term of construct *form* rather than construct *state*.

3. Construct form and case

The morphological variations of nouns commonly designated as *case* have in common with construct forms that they are conditioned by the syntactic status of nouns. But case encodes the role of NPs as elements of broader constructions, irrespective of their internal structure, whereas construct forms encode information on the internal structure of NPs. In other words, case is a particular variety of dependent marking, whereas construct forms are an instance of head marking.

---

[2] CSTR = construct form, DEF = definite, 1S = 1st person singular.

Several authors have proposed to neglect this distinction and to consider construct forms as cases, which implies broadening the definition of case to any morphological variation of nouns carrying syntactic information.

It seems to me that, on the contrary, the distinction between head marking and dependent marking is a crucial methodological distinction, in describing individual languages as well as in a typological perspective. Consequently, I will not follow the proposal to consider construct forms as cases.

4. Terminological problems

When looking for documentation on construct forms of nouns in a cross-linguistic perspective, it must be kept in mind that, in the absence of a generally accepted label for morphological variations of nouns triggered by the presence of certain types of dependents, a variety of terms have been used in descriptions of individual languages:

> Such a form also occurs in some other Western Nilotic languages, and in descriptions of those languages it has been variously labelled "genitive" (Kohnen 1933:28 on Shilluk, Okoth-Okombo 1982:32 on DhoLuo), "appertentive" (Gregersen 1961:83 on DhoLuo), "status constructus" (Tucker and Bryan 1966:83), "antigenitive" (Andersen 1988:284 on Päri), and "modified noun form" (Reh 1996:116 on Anywa).
>
> Andersen 2002:13

Conversely, some descriptive traditions use terms that may suggest some analogy with the construct state of Semitic languages, but which in fact refer to very different phenomena. In this respect, descriptions of Berber languages use the term *annexion state* or even *construct state* (see *Penchoen 1973*) in a particularly misleading way. In Berber languages, the choice between the "free/normal state" and the "annexion/construct state" of nouns has nothing to do with the internal structure of the NP headed by the noun, and exclusively depends on the position of the NP in the constituent structure of the clause.

In the variety of Tamazight described by Penchoen, the "construct state" is used (a) when the noun is subject of the (verbal) utterance and is placed after the verb, (b) after preposition, (c) in noun complement constructions and after at least certain numerals, whereas the "normal state" is used in all other syntactic environments, and as the citation form of nouns.

In the variety of Kabyle described by Naït-Zerrad, the "annexion state" has a slightly different distribution, since in addition to the contexts listed above, all nouns heading an NP co-referent with a bound pronoun attached to a word somewhere in the preceding part of the utterance are in the "annexion state". It is however clear that, in this variety of Kabyle too, this distribution must be stated in terms of functions that NPs whose head is in the "annexion state" assume in a broader construction, and not in terms of their internal structure. In the following example, variations in "state" concern a noun (*aqcic* 'boy', ann. st. *weqcic*) uniformly devoid of any dependent.

(3)   *Kabyle (Naït-Zerrad 2001)* [3]

   a. *Aqcic   yettru*
      boy   3SM.cry.IPF
      'The boy is crying'

---

[3] ANN = annexion state, IPF = imperfective, PF = perfective, 3SF = 3rd person singular feminine, 3SM = 3rd person singular masculine.

b. *Yettru   weqcic*
   3SM.cry.IPF   ANN.boy
   'The boy is crying'

c. *Twala   teqcict aqcic*
   3SM.cry.PF   ANN.girl boy
   'The girl saw the boy'

d. *Twala-t   teqcict, weqcic*
   3SF.cry.PF-3SM   ANN.girl ANN.boy
   lit. 'The girl saw him, the boy'

e. *Yefka   aserwal i weqcic*
   3SM.give.PF   trousers to ANN.boy
   'He gave trousers to the boy'

f. *Isem n weqcic mechur*
   name of ANN.boy famous
   'The name of the boy is famous'

In other words, in a broad typological perspective, the two so-called "states" of Berber nouns are cases. It is true that their distribution does not fit into any cross-linguistically common and consequently well-identified configuration, and this is probably the reason why many specialists of Berber languages are reluctant to recognize them as cases. But if one agrees with the importance of the distinction between nominal head marking and nominal dependent marking, then one must at least recognize that the so-called "states" of Berber nouns are instances of nominal dependent marking, and are therefore functionally similar to cases rather than to the "states" of Semitic nouns.

Given the idiosyncracies of the distribution of the two cases of Berber nouns, none of the labels currently used in describing case systems can conveniently be applied to them. The labels traditionally used in description of Berber languages characterize this constrast as involving a form that can be used in the extra-syntactic function of designation (the "free state"), and a form (the "annexion/construct state") used exclusively when the NP occurs with a given role in certain syntactic configurations. Labels such as *absolute case* vs. *integrative case* would be equivalent in this respect, and would have the advantage of avoiding any confusion with the "states" of Semitic languages.

5. Some African illustrations

5.1. East African languages

As already mentioned, in African linguistics, construct forms of nouns have so far been identified mainly in East African languages belonging to the Nilotic and Cushitic families. I will not reproduce these illustrations here, and will rather draw the attention to the fact that the phenomenon is not limited to this particular area.

5.2. Hausa

In Hausa, the notion of construct form of nouns can apply to the form characterized by a suffix *-n* (singular masculine or plural) or *-r̃* (singular feminine), commonly called 'genitive linker'. This suffix occurs when the noun fulfills the role of head in the genitive construction – ex. (4a-b). It must also be used when the noun takes a possessive suffix other than 1st person

singular – ex. (4c). It results from the cliticization of a pronoun *na / ta* co-referent with the head noun in the synonymous construction illustrated by ex. (4a'-b').

(4) *Hausa*[4]

    a. *kàre-n*   *Daudà*   (cf. *kàree* 'dog')
       dog-CSTR.SGM   Dauda
       'Dauda's dog'

   a'. *kàree*   *na*   *Daudà*
       dog   that one (SGM) of   Dauda
       'Dauda's dog'

    b. *saanìya-r̃*   *Daudà*   (cf. *saanìyaa* 'cow')
       cow-CSTR.SGF   Dauda
       'Dauda's cow'

   b'. *saanìyaa*   *ta*   *Daudà*
       cow   that one (SGF) of   Dauda
       'Dauda's cow'

    c.

kàree-naa	'my dog'	*saanìyaa-taa*	'my cow'
kàre-n-kà	'your(SM) dog'	*saanìya-r̃-kà*	'your(SM) cow'
kàre-n-kì	'your(SF) dog'	*saanìya-r̃-kì*	'your(SF) cow'
kàre-n-sà	'his dog'	*saanìya-r̃-sà*	'his cow'
kàre-n-tà	'her dog'	*saanìya-r̃-tà*	'her cow'
kàre-n-mù	'our dog'	*saanìya-r̃-mù*	'our cow'
kàre-n-kù	'your(P) dog'	*saanìya-r̃-kù*	'your(P) cow'
kàre-n-sù	'their dog'	*saanìya-r̃-sù*	'their cow'

A difficulty in the analysis of *-n* ~ *-r̃* as the mark of a construct form of Hausa nouns is that the same suffix characterizes attributive adjectives preceding nouns in the construction illustrated by *fari-n kàree* 'white dog' / *fara-r̃ saanìyaa* 'white cow' (*fari / fara* are the masculine and feminine forms of the adjective 'white').[5] In this construction, if one accepts that the first term (the adjective) depends on the second one (the noun), *-n* ~ *-r̃* cannot be recognized as an instance of noun head marking. A possible solution is to consider that synchronically, when attributive adjectives precede nouns, they take an additional gender agreement mark homonymous with the suffix of the construct form of nouns.[6]

5.3. Wolof

In Wolof, a construct form of nouns characterized by the suffix *-u* (sg.) / *-i* (pl.) is used exclusively with nouns combined with another noun in the role of genitival dependent. It

---

[4] CSTR = construct form, SGF = singular feminine, SGF = singular masculine.

[5] In Hausa, attributive adjectives occur both before and after the head noun; postnominal adjectives are simply juxtaposed to the head noun: *fari-n kàree = kàree farii, fara-r̃ saanìyaa = saanìyaa faraa.*

[6] An alternative analysis would be to consider that *fari-n kàree* / *fara-r̃ saanìyaa* are *N of N* constructions meaning 'whiteness of dog / cow'. However, this analysis would imply a rule according to which, in the genitival construction, the gender of the head noun might be determined by the gender of the dependent noun. It is therefore preferable to consider that attributive adjectives can precede nouns in a construction resulting from the reanalysis of a former genitival construction, but synchronically distinct from it ("Prenominal adjectives in such phrases as *farin gidaa* 'white house' probably began as N of N constructions meaning 'whiteness of house'" – *Newman 2000:30*).

occurs with no other dependent, and, contrary to Semitic construct forms, it does not occur with possessives either.

The construct form of Wolof nouns shares with Semitic construct forms a constraint of strict contiguity with the dependent noun. This means that other dependents of the head noun in the construct form must follow the genitival dependent, and that, if the dependent noun itself has dependents that must precede it, they must be placed to the left of the head noun in the construct form, as illustrated by ex. (5).

(5)   *Wolof*[7]

    a. *fas   wu   ñuul*
        horse CL.LINK   be black
        'black horse'

    b. *suma   nijaay*
        POSS1S   maternal uncle
        'my uncle'

    c. *suma   fas-u   nijaay   wu   ñuul*
        POSS1S   horse-CSTR   maternal uncle   CL.LINK   be black
        'the black horse of my uncle' (lit. 'my horse of uncle black')

    d. * *fas-u   suma   nijaay*
         horse-CSTR POSS1S   maternal uncle

5.4. Tswana

Tswana nouns whose basic tonal contour ends with two successive H tones (which constitute an important proportion of Tswana nouns, perhaps the majority) show a tonal alternation …HH ~ …HL whose conditioning cannot be stated in phonological terms, which means that this alternation must be recognized as morphological.

In Tswana, …HH is automatically replaced by …HL before pause, but this phonological rule does not account for all cases of replacement of …HH by …HL. In particular, when nouns ending with …HH are in certain types of *head – dependent* relations with the following word, they show a … HL contour that must be analyzed as the characteristic mark of a construct form, since it disappears if the same word occurs in a phonologically identical configuration, but with a different syntactic relation with the following word. For example, in ex. (6a), *sìtswáná* is the head of the NP *sìtswánà sé bá-sì-búà-ŋ* 'the Tswana they speak → the way they speak Tswana', and consequently, the contact with the linker *sé* introducing the relative clause triggers the use of the construct form *sìtswánà*. By contrast, in *sìtílò sá sìtswáná sé bá-sì-rékílè-ŋ* 'the Tswana chair they bought', *sìtswáná* is in contact with the same linker *sé*, but the linker introduces a dependent of *sìtílò* 'chair'[8], not of *sìtswáná*; in this construction, *sìtswáná* has no dependent, and consequently the construct form would not be correct.

---

[7] CL = class marker, CSTR = construct form, LINK = linker, POSS1S = possessive, 1st person singular.
[8] The construct form *sìtílò* is triggered by the presence of the genitival dependent *sá-sìtswáná.*

(6) *Tswana* [9]

  a. *χà-kí-rátí sìtswánà   sé   bá-sì-búà-ŋ*
      NEG-S1S-like 7Tswana.CSTR   7LINK S3:2-O3:7-speak-REL
      'I do not like the Tswana they speak (the way they speak Tswana)'

  b. *χà-kí-rátí sìtílɔ̀   sá-sìtswáná   sé   bá-sì-rékílè-ŋ*
      NEG-S1S-like 7chair.CSTR 7GEN-7Tswana   7LINK S3:2-O3:7-buy.PF-REL
      'I do not like the Tswana chair they bought'

In Tswana, nouns whose basic tonal contour ends with …HH must take the construct form (characterized by a tonal contour …HL) when immediately followed by one of the following types of dependents:
  – a demonstrative,
  – a genitival dependent,
  – an adjectival dependent or a relative clause introduced by a linker homonymous with the demonstrative,
  – the interrogative determiner *-fí*,
  – the negative determiner *-pé*,
  – *-sìlí* 'other'.
By contrast, the construct form is not used when the noun is for example followed by *-ótɬʰé* 'all', or by the linker *-lí* introducing a numeral.

(7) *Tswana* [10]

  a. *bàsádí   bá-tsílè*
      2woman   S3:2-come.PF
      'The women came'

  b. *bàsádí   bótɬʰé   bá-tsílè*
      2woman   2all   S3:2-come.PF
      'All women came'

  c. *bàsádí   bálí   bábèdí   bá-tsílè*
      2woman   2LINK 2two   S3:2-come.PF
      'Two women came'

  d. *bàsádì   bá   bá-tsílè*
      2woman.CSTR   2DEM S3:2-come.PF
      'These women came'

  e. *bàsádì   bá   kí-bà-bídítsè-ŋ   bá-tsílè*
      2woman.CSTR   2LINK S1S-O3:2-call.PF-REL S3:2-come.PF
      'The women I called came'

---

[9] CSTR = construct form, GEN = genitive, LINK = linker, NEG = negation, O3:X = object marker, 3rd person, class X, PF = perfect, REL = relative verb form, S1S = subject marker, 1st person. Numbers at the beginning of glosses indicate noun classes.

[10] CSTR = construct form, DEM = demonstrative, GEN = genitive, LINK = linker, O3:X = object marker, 3rd person, class X, PF = perfect, REL = relative verb form, S1S = subject marker, 1st person, S3:X = subject marker, 3rd person, class X. Numbers at the beginning of glosses indicate noun classes.

f. *bàsádì*   *bá-mʊ́tsí*   *bá-tsílè*
   2woman.CSTR   2GEN-3village   s3:2-come.PF
   'The women of the village came'

g. *bàsádì*   *básʊ̀lí*   *bá-tsílè*
   2woman.CSTR   2other   s3:2-come.PF
   'Other women came'

6. Construct forms of nouns in Gur languages?

Elders 2003 proposes to use the term of *annexion state* to describe a type of syntactically conditioned alternation affecting nouns in Kulango and some other Gur languages. In the introduction of his paper, he explicitly states that he considers *annexion state* and *construct state* equivalents: "Cette distinction entre état libre et état d'annexion est empruntée à la linguistique berbère (Galand 1988:229); la linguistique sémitique fait une distinction pareille sous les désignations état construit (status constructus, construct state) et état indépendant."

In section 4, I have argued that *annexion/construct state* in Berber linguistics and *construct state* in Semitic linguistics refer to two very different phenomena, and that the "states" of Berber nouns are cases. Consequently, it remains to examine wether the morphological variations of nouns in Gur languages for which Elders proposes the term of annexion state are instances of nominal dependent marking (like the annexion state of nouns in Berber languages), or instances of nominal head marking (like the construct state of nouns in Semitic languages), or perhaps something that cannot be described in terms of either head or dependent marking.

In Kulango and other Gur languages, nouns in isolation have an obligatory class suffix (or number suffix, in languages in which the class system is no more active), but they lose this inflectional suffix in certain constructions. This alternation is conditioned by the role of the noun in the construction of a complex NP, and has nothing to do with the role of an NP in a broader construction.

Consequently, the hypothesis of an analogy with the annexion state of Berber nouns must be abandoned. By contrast, the possibility of similarities with the construct state of Semitic nouns deserves consideration.

However, unlike the construct state of Semitic nouns, which is a clear instance of head marking, in Gur languages, the omission of the inflectional suffix concerns nouns that are always in non-final position in the relevant construction, but that cannot be uniformly characterized as heads or dependents, since some constructions triggering the omission of the inflection of the first term have the order *head – dependent*, and some others have the order *dependent – head*. This is illustrated by ex. (8) from Dagara: 'goat' in the suffixless form is the head of an attributive construction in (8b), and the dependent of a genitive construction in (8c).

(8)   *Dagara (Delplanque 1997)*

   a. *bʊ-ɔ* 'goat' (pl. *bʊʊ-d*)

   b. *bʊ̀*   *bɛ́d-ʊ*   'big goat'
      goat big-SG

   c. *bʊ̀*   *zʊʊ-d*   'goat's tail'
      goat tail-SG

In Dagara (and in other Gur languages showing this kind of alternation), if one accepts that in (8b), 'goat' is the head of an attributive construction,[11] the use of the uninflected noun illustrated by ex. (8b) is functionally comparable to the construct state of Semitic languages, since it characterizes the noun fulfilling the role of head in a *head – dependent* construction. But this is not the case of the use illustrated by ex. (8c).

If one accepts my proposal to draw a strict distinction between cases (nominal dependent marking) and construct forms (nominal head marking), then the only possible conclusion is that this distinction is not sufficient to predict the distribution of bare nouns in Gur languages.

In addition to that, an important difference between Semitic nouns in the construct state and uninflected nouns in the attributive construction of Gur languages is that in the Gur attributive construction, the adjective bears a suffix providing the same information as the suffix attached to the noun in the absence of an adjectival dependent.

Consequently, in languages that have the kind of alternation illustrated by ex. (8), the behaviour of the class or number suffix must rather be described as follows:

(a) the class or number suffix is a *phrasal affix* attaching to the last word of the attributive construction,

(b) the absence of the class or number suffix characterizes genitive NPs.

References

Andersen, T. 2002. Case inflection and nominal head marking in Dinka. *Journal of African Languages and Linguistics* 23-1. 1-30.

Delplanque, A. 1997. La qualification en dagara. *Linguistique africaine* (N° hors-série). 41-73.

Elders, S. 2003. L'état d'annexion en kulango. *Gur papers / Cahiers voltaïques* 6. 67-81.

Harrell, R. S. 1962. *A short reference grammar of Moroccan Arabic*. Washington, D.C.: Georgetown University Press.

Mous, M. 1993. *A Grammar of Iraqw*. Hamburg: Helmut Buske.

Naït-Zerrad, K. 2001. *Grammaire moderne du kabyle*. Paris: Karthala.

Newman, P. 2000. *The Hausa language, an encyclopedic reference grammar*. New Haven / London: Yale University Press.

Penchoen, T. G. 1973. *Tamazight of the Ayt Ndhir*. Los Angeles: Undena Publications.

---

[11] In the same way as in Hausa, one might argue that this is in fact a genitive construction, lit. 'goat's bigness', but class agreement provides evidence against this analysis, at least in a strictly synchronic perspective. Note however that, within the frame of this alternative analysis, the recognition of the uniflected form of nouns as a construct form is excluded even more radically, since the use of the uninflected form of nouns would uniformly characterize nouns in the role of dependent in the genitive construction.