The Grammaticalization of "Say" and "Do": An Areal Phenomenon in the Horn of Africa
David Cohen, Marie-Claude Simeone-Senelle, Martine Vanhove

To cite this version:

HAL Id: halshs-00009729
https://halshs.archives-ouvertes.fr/halshs-00009729
Submitted on 22 Mar 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
in Güldemann, Tom and von Roncador, Manfred (eds.), 2002. Reported Speech : 
A Meeting Ground for Different Linguistic Domains. Amsterdam, Philadelphia: 
John Benjamins, 227-251.

The grammaticalization of ‘say’ and ‘do’: An areal 
phenomenon in East Africa

David COHEN
Professeur Emérite 
Marie-Claude SIMEONE-SENELLE
CNRS - LLACAN 
Martine VANHOVE 
CNRS - LLACAN

1. Introduction
In this volume devoted to reported discourse and quotatives, we would like to 
deal with a very particular development of the use of main verbs meaning ‘say’.
Our purpose is to show how, in some languages, the quotative ‘say’ has come to 
be used both as an auxiliary verb and as a formative of verbs derived from 
various word categories, leading ultimately to a reorganization of the verb 
system. We have had to include the verb ‘do’ as well, given its similarity and 
complementarity of its behavior in the languages studied.

In several language families, Egyptian, Cushitic, Omotic, Semitic, and 
nilo-Saharan, spoken in Eastern Africa (i.e. Egypt, Sudan, Ethiopia, Eritrea, 
Somalia, Djibouti, Northern Kenya and Northern Tanzania), full verbs meaning 
‘say’ and ‘do’ (and a few others with related meanings such as ‘think, have in 
mind, intend, consider, shout’ and ‘put, place’) are frequently found to have uses 
as auxiliary verbs and as formatives for new verbs. M. Cohen (1936), following 
Junod (1933), called the resulting forms ‘descriptive compounds’. This has 
become the conventional designation in Hamito-Semitic studies and will be used 
in this paper. The base of the new verbs may be a noun, an adjective, an adverb, 
an onomatopoetic / ideophonic form, a verb or even a phrase. Descriptive 
composition thus involves both transcategorial and intracategorial processes. 
The antiquity of this process of verb creation is attested by its presence in 
ancient Egyptian.

As auxiliaries, ‘say’, and ‘do’ may have other modal values such as
future’, ‘injunctive’, ‘inchoative’, ‘volitive’, ‘assertive’, ‘authenticative’. They may also be used as focalizers. These values are usually linked to different syntactic constructions.

In this paper, we will focus on the process of verb creation and its morphological consequence, the renewal of verb inflection. Indeed, in some languages, such as most Cushitic languages, in Egyptian, and, to some extent, in Dongolese Nubian, auxiliary verbs have been reduced to the status of verb morphology.

The role of ‘say’ and ‘do’ as auxiliaries and in the creation of new verb inflection has been recognized for more than a century (see Isenberg 1842, Praetorius 1879, Afevork 1905, 1911, Armbuster 1908-1920, Walker 1928, Baeteman 1929 for Amharic, Reinisch 1878 for Cushitic, and Praetorius 1894 for Hamito-Semitic). The work of these scholars was elaborated in France by M. Cohen (1936, 1939).

As far as we know, there are no comparative linguistic studies describing this areal phenomenon in detail. We therefore provide below a comparative and typological survey. We will see that, in some languages, verbs meaning ‘say’, and ‘do’ work in this way, while in others, only one of these verbs may be used.

Illustrations will be provided from a variety of languages in different families, and we will give particular attention to Afar, a Cushitic language for which we can use data from our own fieldwork. We will also show that the East African data can be related to similar phenomena, hitherto unrecognized as such, in Aiki, a Nilo-Saharan language spoken further west in the Central African Republic. Kanuri, another Nilo-Saharan language, and Hausa, a Chadic language, also show the same tendency. We will also take into account languages where the process of grammaticalization has gone to completion, i.e. Ancient Egyptian and Cushitic for ‘say’, and Coptic for ‘do’.

Proposals will be made, within the framework of D. Cohen’s general theory of verb auxiliarization (see D. Cohen 1983-1985, Vanhove 1993: 101-105, Simeone-Senelle & Vanhove 1997, and Cohen, Simeone-Senelle & Vanhove in prep.) to explain how ‘say’ and ‘do’ become auxiliaries and to define which semantic features of these two verbs account for their comparable behavior, setting them apart from other categories of verbs.
2. The function of ‘say’ and ‘do’ in East Africa

2.1. Cushitic

2.1.1. The case of Afar

The following study is based on data collected by Simeone-Senelle & Vanhove in Eritrea in 1996, and in Paris between 1996 and 1998 (Djibouti dialect). We have also consulted Parker and Hayward (1985), Morin (1995), and Hayward (1996). We will first review the morphology of descriptive compounds and then discuss their syntactic and semantic features.

The descriptive compounds of Afar are formed with one of two auxiliaries, edhe ‘say’, an irregular verb based on two different roots, and hee ‘put, place’, which is regularly inflected with suffixes. The periphrastic word order is always base verb - auxiliary verb. Compounds may coexist with a corresponding ‘ordinary’ verb, or be created from any other word class.

Semantically, edhe forms intransitive or middle verbs, while hee is used to construct transitive ones, but a few onomatopoeic bases may use either auxiliary (though edhe is more frequent by far).

(1a) c`m÷m g``g'rrh,gdd or (1b) g``g'rrh,hxxd
donkey bray-put.3 M.SG.PFT bray-say.3 M.SG.PFT
the donkey brayed (cf. g``ghr· ‘it brayed’)

Outside certain specific contexts such as negative or coordinated sentences, the compound constitutes a prosodic unit with main stress on the penultimate syllable of the stem. Moreover, the auxiliary may undergo phonetic reduction.

The morphology of the stems varies. The most common ones in our data have a geminated final consonant, followed by a vowel harmonizing with the preceding one. Some dialectal differences in vowel quality exist: the data in Parker & Hayward (1985) and Bliese (1981: 146) show no vowel harmony, the final vowel is always a. With this pattern, the stem is most often derived from a verb (in base or derived form).

The two examples below illustrate intransitive compounds with edhe:

(2) kobódq˘u-iyyaannah
gather-3PL.PFT.say
They gather themselves (cf. kobódq˘n ‘they gather themselves’)

(3) way-timtí-itte awká baça
lack-PAS.-3F.SG.PFT.say girl daughter
The little girl could not be found (cf. way-timtí ‘she was not found’)

The following are examples of transitive compounds with hee:

(4) wáyya-heeni
lack-put.3PL.PFT.
An areal phenomenon in Eastern Africa 4

They did not find (waytén ‘they lacked, did not find’)

(5)  hulú-ssu-hee
      enter-CAUS.-put.3M.SG.PFT.

He made enter (hulu-se ‘he made enter’)

The same pattern may sometimes be used with nouns (ex. 6) or ideophones
(which cannot be used in isolation, ex. 7, and 8):

(6)  tibbi-iyye
      silence-3M.SG.PFT.-say

He fell silent (cf. tibba ‘silence’)

Examples with ideophones

(7)  kábba-idqéh
      IDEO-say.IMP.SG

come near!

(8)  bir  ḍongoló sugtéh  ḍh  yí
      yesterday evening  noise  stay.3F.SG.PFT  this mine

huggaaney dábbu  itta  haak  sugté
      neighbors  IDEO  REC.  put.PART.  stay.3F.SG.PFT

Yesterday evening there was a dull thudding sound, my neighbors
were fighting

Monosyllabic CVC ideophones may also be made into verbs by a
derivational morpheme -y. The stem is then lengthened to CVVC: думbуй ‘he
beat’. The origin of this -y could have been a verb meaning ‘say’. This
possibility of verb derivation varies with the dialect.

In the data we elicited from our Tadjoura (Djibouti) informant, and in
Parker & Hayward (1985), we find other stem patterns for the base of
descriptive compounds. Monosyllabic ideophones (CVC) may keep their stem
unchanged:

(9)  dub  má-dh-in
      IDEO  NEG-say.IMP.SG-NEG

Do not kick up a row!

or may have it expanded by full or partial reduplication (C₁V(C₂)C₁V(C₂)),
or by vowel lengthening conveying a different meaning:

(10a)  dubdúb-edße  and  dubdúb-edße

are variants of

(10b)  díbbá-edße
      IDEO-1SG.PFT.say

I made a dull thudding sound (Parker & Hayward)

as opposed to:

(10c)  díub-edße  ‘I collapsed slowly to the ground’

(10d)  díub-hee
Vowel lengthening also applies to verb stems:

(11) \[ \text{dáaf-iyye} \]
     sit down-3M.SG.PFT.-say
     he hardly sat down (cp. \text{dáffa-iyye} ‘he sat down’)

(12) \[ \text{ḥulúu-s-hee} \]
     enter-CAUS-put.3M.SG.PFT
     he hardly made (someone) enter (cp. \text{ḥulússu-hee}, ex. 5)

The lengthening of the final stem vowel of verbs and ideophones thus conveys an ‘attenuative’ sense. According to our Tadjoura informant, this device is used only in literary language. His observation seems to be corroborated by the fact that these forms are unattested in our spontaneous data.

When the base is an adverb or a noun that can be used in an adverbial function, the form of the compound is stem + -h ‘towards’ + ‘say’ or ‘put’:

(13) \[ \text{bihó-ḥ idqeh} \]
     far or forward-DIR say.IMP.SG
     Step back; go ahead; push; progress!

(14) \[ \text{inni addá-ḥ ḳee} \]
     or \[ \text{addá-ḥ inni ḳee} \]
     myself depth-DIR put.1SG.PFT
     As for myself, I humiliated myself

(15) \[ \text{addá-ḥ edje} \]
     depth-DIR 1SG.PFT.say
     I was humble

As stated above, descriptive compounds can undergo morphological derivation, like any other verb in the language. In addition to the process found in our spontaneous data, i.e. derived stem + ‘say’ or ‘put’ (see. ex. 3, 5, and 12), the auxiliary may bear the derivational morpheme and the stem be in base form:

(16) \[ \text{galbó kaa-k ḳálla-hay-stéeh} \]
     skin him-from take off-put-AUTOBEN.3M.SG.PFT.
     He skinned for his own benefit (cf. \text{kalsitti-hee}) (Morin 1995: 100)

(17) \[ \text{tiḥbi-yedjehe} \]
     silence-3M.SG.PFT.say.AUTOBEN.
     He held his tongue for his own benefit

Another process involves the derivation of both the base and the auxiliary verb:

(18) \[ \text{ḥulú-ssu hay-site} \]
     enter-FACT. put-AUTOBEN.3M.SG.PFT.
     He caused (someone) to enter for his own benefit (cf. \text{ḥulsitó}).
Syntactically, the stem and the auxiliary of descriptive compounds may be separated by two word classes: the adverbial particles eddé and ellé, and the personal pronouns (with or without a postposition). Their prosodic unity is then broken:

(19) aká akat aqúyu-haynah giitná wa’diina úsk other rope fasten-put.1PL.IPFT pull.1PL.IPFT when he tifbi edde iyáa akwi baqí tifbi edde silence in 3M.SG.IPFT say boy son silence in iyáa 3M.SG.IPFT say

We fasten another rope (to the cradle) and when we pull it, he quiets down, the small boy becomes quiet

(20) tifbiqí kaa hee
catch him put.3M.SG.PFT

He caught him

(21) dib ko-t hée-yyo
IDEO you-on put-1SG.FUT.

I’ll beat you

Another important syntactic feature is that, in a series of coordinated stems, a single auxiliary, the last one, may apply to all of them:

(22) hangoyíeena baaqó-l háyya-ay, gíli baaqó-l háyya-ay, forefinger floor-on put-and thumb floor-on put-and ánnak booha baqá takké-m gidaháa, wóo little hole daughter 3F.SG.IPFT become-REL so that this guri gabaatáa hangoyíeena kee gíli baaqó-l left hand forefinger and thumb floor-on háyya-han put-put.3PL.IPFT.

They put the forefinger on the floor, they put the thumb on the floor so that it will form a small hole, they put the forefinger and the thumb of the left hand that way on the floor

If the verb stems do not take the same auxiliary, the one that applies to the last verb predominates:

(23) dini-k ugú-ttu-y, fóóhu ka’l-íssi-y, af sleep-from wake-PAS.-and face wash-FACT.-and mouth sadássa-y, duuli híllu-y šanta akúcci-y, afá-k brush-and toilets go-to-and bag carry-and door-through ewécci-iyye
go out-3M.SG.PFT say
He got up, washed his face, brushed his teeth, went to the toilets, took his bag, and went out through the door.

Descriptive compounding of verbs is very productive in Afar. Longacre (1990: 18-19), following Bliese, considers such verbs as marking “pivotal storyline action/ events”. Indeed, they are often used in narratives where they may alternate with their corresponding uncompounded verbs. They are also frequent in discourse. Their use depends on the speaker’s intention and expresses his attitude towards the predicate. The reported facts and processes are not indifferent to the speaker; rather, the compound verbs are used to convey his emotions such as astonishment, admiration, disapproval, etc. In the following example, indignation is perceptible in the answer (where use of the simple form yo-l usuuléh would be a simple statement of fact):

(24) - ka-t daawitëh - hâyyee? - usuulë lu yo-l bee
- him-at scold.1SG.PFT - so what? - laugh me-at put.3M.SG.PFT
- I scolded him - So what? - (Well) he had laughed at me!

The aim is to narrow down the range of possible interpretations of the utterance by the other discourse participant. In the following example, the old woman leaving the recording session early lets us know that she has a good reason for going:

(25) anú dôhri abïtti-hee
I noon prayer make.AUTOBEN.-put.1SG.PFT
As for me, I am going to pray

The speaker’s involvement also accounts for the use of descriptive compounds when giving orders or instructions:

(26) ifissirri kaa hays
translate him put.2SG.IMP
Translate for him!

To some extent, then, the use of descriptive compounds falls within the domain of discourse-participant interplay. This is confirmed by the fact that these compounds can be used with any verb inflection (simple or compound tenses) conveying values of tense, aspect, or epistemic or deontic modal categories.

2.1.2. Other modern Cushitic languages

Other modern Cushitic languages use descriptive compounds. Below are a few examples from languages which, unlike Afar, only use the verb ‘say’ as auxiliary.

In Oromo, a Lowland East Cushitic language, the use of the verb ‘say’ appears in two examples in Stroomer (1987). ‘Say’ is yed’a, jed’a or yaa-. It would seem that ‘do’ is not used to create verbs (Stroomer p.c.). The base can be
either a noun or a verb with slightly modified stem (reduction of final long vowel or vowel harmony):

(27)  
 c‘alla yed’a
 only say
 Be silent, remain silent

(Three derived verb forms of the above base may also be used: c‘alleed’d’a ‘be silent, remain silent’, c‘alleesa ‘be quiet, silent’, and c‘allised’d’a ‘be silent’.)

(28)  
 odoo innii mina sceene yed’u
 when he house enter say.3SG.PFT
 When he entered the house

To judge from the few examples given by Roper (1929: 84), Beja, a Northern Cushitic language, also uses only ‘say’ (di); the base component may be from any word class:

(29)  
 mici di
 click meaning ‘no’ say
 To make the click meaning ‘no’

(30)  
 fidig indi
 to depart 3SG.IPFT.say
 He is going

(31)  
 jar di
 shoo! say
 To shoo (hens) away

2.1.3. Early Cushitic
From a diachronic viewpoint, it is important to note that most Cushitic languages have suffixal conjugations for both perfect and imperfect. It is generally believed that these conjugations are recent and developed from the prefixal conjugation of variable stems. It has been recognized since Praetorius (1894) that the suffixes go back etymologically to a very short auxiliary verb conjugated with prefixes. Reconstruction based on comparative studies has shown that this auxiliary is a verb meaning ‘say, be’, *Vn, *a, *y, etc. (see for instance Zaborski 1975, Cohen 1984, 1988). Moreover this reconstruction is supported by the fact that some Cushitic languages (Beja, Afar, Somali, South-Agaw) have remnants of earlier prefixal inflection, which is still used for a limited number of verbs (the proportion varies from 60% of all verbs to five verbs only).

The following is the conjugation of a Somali verb (Digil variety) rog ‘stay’ in the imperfect (see D. Cohen 1984: 89):

<table>
<thead>
<tr>
<th></th>
<th>Sg.</th>
<th>Pl.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>rog-ä</td>
<td>1. rog-n-ä</td>
</tr>
</tbody>
</table>
The stem *rog-* is thus inflected with the auxiliary *a*, conjugated with the Hamito-Semitic personal prefixes.

2. **Ethio-Semitic**

Modern Ethio-Semitic languages make widespread and colorful use of descriptive compounds. Typologically, they can be distributed into five groups:

1. languages using only a verb meaning ‘say’ in base form
2. languages using only ‘say’ in base and derived form
3. languages using both ‘say’ and ‘do’ in complementary distribution
4. languages using both ‘say’ and ‘do’ in free variation
5. languages using both ‘say’ and ‘do’ with their derived forms

2.2.1. **Languages using only a verb meaning ‘say’ in base form**

From the examples given by Leslau (1956: 145-6) for Ethio-Semitic, it seems that only Zway, out of a list of 15 languages, has a system with only one auxiliary, *balo* ‘say’: *sam balo* ‘be quiet’.

2.2.2. **Languages using ‘say’ and a derived form of ‘say’**

Leslau (1956: 145-6) mentions the use of *balä* ‘say’ and its causative derived form *a-balä* ‘make say’ to create verbs in Gafat.

The invariable stem used with *balä* expresses mainly intransitive action; transitive action is expressed by *a-balä*. To judge from the examples provided, the stems of descriptive compounds are restricted to this particular usage. This is why no word-for-word translation is given hereafter. Leslau recorded only bi-consonantal stems: *qaT balä* ‘be straight’, *Ìnga balä* ‘disobey’; as a causative form, he gives *käf a-balä* ‘raise, make higher’.

Leslau also mentions Tigrinya, Gogot, and Aymellel as following the same pattern of verb creation, the only differences being a larger syllabic range of bases and the fact that the compounds may alternate with an uncompounded verb.

2.2.3. **Languages using both ‘say’ and ‘do’ in free variation**

Some languages can use both ‘say’ and ‘do’ in free variation with no difference in meaning. This usage is only marginal in Amharic (ex 32) or in Tigre (ex. 33) where it is limited to ideophonic bases:

(32) *tuss åli* or *tuss adǝrrǝggǝ*
2.2.4. Languages using both ‘say’ and ‘do’ in complementary distribution

In many of the languages using both ‘say’ and ‘do’, their syntactic roles generally turn out to be complementary: ‘say’ forms intransitive verbs while ‘do’ forms transitive ones, whatever class the base belongs to.

The Gunnän-Gurage languages belong to this typological category. Verbs can be formed from interjections and onomatopoeias or derived from verbs. ‘Say’ is balä / barä and ‘do’ is amännä. Here are some examples provided by Hetzron (1977: 109-110) in various dialects:

(35) bäy balä
    all right say
    Agree

(36) źaa amännä
    IDEO do

(37) daqq barä
    laugh say
    Laugh

Leslau (1956: 145) attributes the same pattern to Argobba.

2.2.4. Languages using both ‘say’ and ‘do’ with their derived forms

This typological category shows several patterns of distribution of the base and derived forms.

According to M. Cohen (1936: 262-275), Amharic makes considerable use of descriptive compounds with expressive value. They are formed from an uninflected base (final consonants are most often geminated) derived from any word class, and an auxiliary verb. ḫül ‘say’ forms verbs with a “neutral meaning” (i.e. middle verbs), and is the most frequent auxiliary. The less frequent adarraga ‘do’ forms active verbs. In addition, derived forms of ‘do’ (asdaarraga ‘make do’) and ‘say’ (tabüla ‘be said’) may also be used:
An areal phenomenon in Eastern Africa 11

(38) yaqør tæ-bæla
stay.3MSGJUS PAS-say.3SG.PFT
He was forgiven

(39) yohan nagar lobb asdarragaañ
DEM.A thing intelligence FACT-do.3SG.PFT-ME
He pointed this thing out to me

Onomatopoeic stems, occasionally used as interjections, can only appear in the form of descriptive compounds; no regular verb exists alongside the compound:

(40) da āla or dada āla
ONO say RDP.ONO say
Be slow to speak, walk

(41) sagg or ṣagg or ṣaggg adarraga
ONO ONO RDP.ONO do
Cool down

Verbs can also be converted into descriptive compounds by special stems. The result is considered to be more expressive. Compounds have intensive or attenuative value, according to stem morphology. We thus have an intracategorial process of verbal creation giving rise to semantic derivation:

(42) səbabət āla - səbabətər āla
be broken say - be smashed say
Be somewhat broken - be somewhat smashed

(43) səbabət āla - səbabətər āla
be broken say - be smashed say
Be totally broken - be totally smashed
(ta-səbabəbəra ‘be somewhat broken, smashed’)

(44) sakakk adarraga - sokakk adarraga
put on partially do - put on in large quantities do
Put on partially - put on in large quantities

A transcategorial process allows a few descriptive compounds with idiomatic structure to be formed from nominal bases and short expressions:

(45) pələtən dəl adarraga
enemy victory make.3SG.PFT
He beat the enemy

(46) hullun əmbi āla-ñ
all no say.3SG.PFT-ME
He refused everything to me
In Southern Ethio-Semitic, the examples provided by Leslau (1956: 145-6) show that some languages such as Ennermor use ‘say’, ‘do’, and derived forms of ‘do’: the causative of *tis barä* (lit. ‘say »tis »’) ‘drip’ is formed with *epa* ‘make’, and the factitive, with the derived factitive form of *epa, atepä*. Leslau gives other examples from Selti and Wolan. In Čaha and Muher, the passive derived form of ‘do’ is semantically equivalent to ‘say’: *tis baräm* (lit. ‘say tis’) ‘drip’ is equivalent to *tis mänäm* (lit. ‘be made tis’), and their causative form is formed with the derived causative form of ‘be made’, *amänäm*.

Tigre uses still another pattern: *bela* ‘say’, its derived causative form *abala*, and *wada* ‘do’, are in complementary distribution. Raz (1983: 66) specifies that ‘most *bela* compounds happen to be intransitive. They may be rendered transitive or become causative of the intransitive by the use of the verb *abala* ‘be said’: *kaf bela* ‘sit’ - *kaf *abala* ‘make sit’, *bah bela* ‘rejoice’ (intr.) - *bah *abala* ‘rejoice’ (tr.).

As in Amharic, the stem may be limited to its use in the compound:

(47)  qa\c bela  
    IDEO.sharp noise say  
    Make a sharp noise

or may be used along with a regular verb:

(48)  ba\h\a bela  
    pardon say  
    Pardon, excuse (*ba\hala* ‘to pardon’)

The preceding example shows no semantic difference between the regular and the compound verb. The example below, however, show semantic derivation:

(49)  waswas bela  
    move say  
    move fast/to and fro (*waswasa* ‘move’)

According to Raz (1983: 67), the compound form may, as in Amharic, convey further information concerning the aspect of action or state of happening – as compared with the semantic content of a coexisting verb of the same origin as the first element of the compound. The addition in meaning can be specified in terms of intensity or manner of the activity, such as: augmentative, attenuative or iterative.

Some stems may only be used with *wada* ‘do’, and not with *bela*, to form intransitive verb compounds: *yak wada* ‘hurry’, *aːqal wada* ‘be patient’.
An areal phenomenon in Eastern Africa

2.3. Egyptian

2.3.1. Ancient / Middle Egyptian

In Ancient Egyptian, the verb system was based on several patterns involving a stem + inflection whose origin is, in some cases, auxiliary + subject pronoun.

One such inflection, the so-called pseudo-participial, stem-\(j/w\), was probably a vestigial form lying outside the regular verb system. One also finds the basic patterns of suffixal conjugation, stem-SM (\(\text{\(s\)m.f} = \text{hear-SM}\)) and stem-n-SM (\(\text{\(s\)m.n.f} = \text{hear-n-SM}\)), together with three other conjugations stem-\(k3\).SM, stem-\(hr\).SM, stem-\(j.n\).SM which later superseded the basic ones in Middle Egyptian, thereby renewing the tense-aspect system.

The most ancient evidence of grammaticalization of a verb ‘say’ as a verb inflection is provided by the pseudo-participial inflection whose final morpheme \(j\) has been suggested to come from a verb \(j\) ‘say’ conjugated with prefixes (perhaps related to Cushitic *\(y\) ‘say’, see above 2.1.3.) (see M. Cohen 1922, Lefèvre 1955, Diakonoff 1965, Gardiner 1969, and D. Cohen 1984: 117, 123).

The three conjugations, stem-\(k3\).SM, stem-\(hr\).SM, and stem-\(j.n\).SM, have been diachronically analyzed as: stem + auxiliary + subject pronoun (Cohen 1984: 130-6 and Vernus 1988: 172). -\(j\)- ‘say’ (giving -\(j.n\)- when conjugated according to the stem-n-SM pattern) has been thought a possible source for the stem-\(j.n\).SM conjugation. The other two conjugations, stem-\(k3\).SM and stem-\(hr\).SM, are also formed with auxiliaries (also used as main verbs in Middle Egyptian) with meanings akin to ‘say’: \(k3\) ‘think, have in mind, intend’, and \(hr\) ‘shout’. All three auxiliaries became grammaticalized as conjunctural elements. They were first suffixed to the verb stem before subject pronoun. Two of them, \(k3\) and \(hr\), even gained enough autonomy to precede the verb stem, with or without the subject attached to them. They became “enunciative auxiliaries” (Vernus 1988: 172, 180), i.e. auxiliaries indicating the speaker’s attitude towards the sentence: \(k3\).SM-stem.SM or \(k3\).stem.SM, \(hr\).SM-stem.SM or \(hr\)-stem.SM.

The 3000-year history of Egyptian provides the best available evidence of all the steps to full grammaticalization by a transcategorial process, starting from a free syntactic construction (stem + auxiliary), through morphologization (verbal inflection), and on to the ultimate formation of a grammatical morpheme with a new semantic value (modal particle).

These etymological hypotheses, while highly probable, are as yet unproven. They are nevertheless significant in view of the parallels between Egyptian and the Cushitic verb system, where verbs meaning ‘say’ have also renewed the conjugation (see above 2.1.3.).
2.3.2. Neo-Egyptian / Coptic

The constructions which gave rise to the Coptic verb system are attested in earlier stages of Egyptian. One of the most productive is etymologically a conjugated auxiliary ‘do’ + an infinitive (see Cohen 1984: 126 sq.).

Already in Neo-Egyptian, the stem-SM and stem-n.SM conjugations mentioned above were vestigial, and when the verb root had more than three consonants, the stem-SM form was no longer used; it was replaced by a periphrasis formed with the auxiliary irj ‘do’:

\[(50) \text{irj.i smtj} \]
\[\text{do-1SG.PFT examine} \]
I examined (the documents)

This construction was generalized in later stages. By the end of a process which began as early as the Middle Empire, the structure of the verb system in Coptic, the latest attested stage of Egyptian, was radically changed. But by a process reminiscent of Cushitic, the conjugation of some kinds of the so-called ‘form I’ are historically based on an older verb conjugated in the original way, i.e. by suffixing possessive personal pronouns. The affixes of the new Coptic conjugation are a which comes from i.irj.f, the emphatic form of irj ‘do’, and re from irj.f, the subjunctive of the same verb.

The tenses involved are (with the verb sōtm ‘hear’):

<table>
<thead>
<tr>
<th>Tense</th>
<th>Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perfect I</td>
<td>a-f-sōtm</td>
</tr>
<tr>
<td>Perfect II</td>
<td>nt-a-f-sōtm</td>
</tr>
<tr>
<td>Habitual</td>
<td>ū-a-f-sōtm</td>
</tr>
<tr>
<td>Optative</td>
<td>ma-re-f-sōtm</td>
</tr>
<tr>
<td>Final</td>
<td>ta-re-f-sōtm</td>
</tr>
</tbody>
</table>

The structural evolution here shows the same kind of renewal as in Egyptian and Cushitic, except that the verb ‘do’ is used instead of ‘say’. The descriptions that follow show that the same process involving both ‘say’ and ‘do’ is at work today in a large variety of modern languages and language families in Africa. The semantic developments are also the same.

2.4. Other language groups

2.4.1. Dongolese (Nubian)

In Dongolese Nubian, a Nilo-Saharan language spoken in the Sudan, there is a compound conjugational form: invariable stem + conjugated verb ‘say’ (= -e / -ē / -ê). The stems can be interjections, onomatopoeias, or any word of Nubian or Arabic origin. Their stem is specific to the construction, and cannot be used independently (Armbruster 1960: 192). Armbruster (p. 32) notes that a “large [...] number of verbs in -ê should have been formed in the last 1000 years”, since
Old Nubian times.

The verb ‘say’ is clitic to the stem, but the stress is placed differently depending on the linguistic origin of the stem: 99% of Nubian bases are stressed, and -ē ‘say’ is unstressed, while the opposite occurs with Arabic bases. This could be indicative of differing degrees of grammaticalization, Arabic bases being less well integrated into the verb system, so that the compound form keeps the stress on the verb ‘say’ and its long vowel as when used as a main verb. This is supported by the fact that almost all Arabic loanwords (99%) are conjugated in this way. Here are some examples of borrowings taken from Armbruster (1960: 246):

51) ǧámm-ē
   assembly-say
   Come together (one to another)

52) tingu ḍādl-ēran
   These stay-say,3PL.IPFT
   These are staying

Armbruster classifies other stems into two different categories, which may also reflect differences in the degree of grammaticalization: (1) Nubian onomatopoeic stems with (C)VCC pattern, which he says are mostly intransitive (Armbruster 1960: 192), cannot be used outside this construction; there are, however, a very few transitive verbs:

53) kītt-ē ‘be silent’
54) gūrr-ē ‘rejoice’
55) úff-ē ‘blow’ (intr. & tr.)

(2) Other transitive or intransitive bases (whose meanings outside this construction are not provided) have diverse morphological patterns:

56) bagāš-ē ‘step over’ (tr.)
57) gūl-ē ‘think, remember’ (tr.)
58) nēw-ē ‘breathe’ (intr.).

Dongolese Nubian uses another verb meaning ‘say’, an, to create compound conjugational forms. According to Armbruster (1960: 271) it “appears much less often than -ē”. Contrary to -ē, an is subject to combinatoric and semantic restrictions. It can only be suffixed to some word classes: adverbs, adjectives, participles, cardinal numerals, nouns indicating condition, and can only create verbs of “becoming”:

59) dūl-an
   great / aged-say
An areal phenomenon in Eastern Africa

become great, increase, grow old

(60) \textit{missi-wéran} \\
\hspace{1cm} \text{eye-one-say} \\
\hspace{1cm} \text{Become of fixed opinion (lit. become one-eyed)}

2.4.2. Omotic

In Kafa, an Omotic language, Cerulli (1951) mentions the use of a verb \textit{y ‘say’} to derive verbs from nouns or verbs:

(61) qa\ddot{g}\ddot{g}o \textit{y ‘refuse’} (= qa\ddot{g})

(62) \textit{cu\ddot{f}u\ddot{f}u} \textit{y ‘fumigate’} (\textit{\ddot{c}u\ddot{f}o} ‘smoke’).

3. Languages outside East Africa

The use of ‘say’ and ‘do’ as auxiliary verbs for verb creation is not limited to East Africa. Discussions with our colleagues at the unit ‘Langage, Langues et Cultures d’Afrique Noire’ (LLACAN) have led us to conclude that other language families, spoken further to west, are affected by this process.

Although the present distribution of the use of the descriptive compounds with ‘say’ and ‘do’ and their tendency to develop into verb inflection is without doubt largely due to areal diffusion, the extent of the phenomenon outside East Africa and the languages mentioned above is remarkable. Before concluding, we would like to cite three other languages: Kanuri (a Nilo-Saharan language spoken in Northern Nigeria), Aiki (a Nilo-Saharan language spoken in the south of Chad and north of the Central African Republic), and Hausa (a Chadic language spoken in northern Nigeria and southeastern Niger).

3.1. Kanuri (Saharan)

In Kanuri, where there are two verb classes, verb class 2 is inflected by a morpheme which is reconstructed as a verb ‘say’ (\textit{ngin}), still in use as a main verb today. According to Hutchison (1981: 225), “verb class 2 is virtually unlimited in size since any lexical item of Kanuri or of a contact language can today be inflected [in such a way]”. Cyffer (1997: 34) states that “more than 95% of verbs belong to this group. It is highly productive and all verbs introduced through innovation or borrowing belong to class 2”.

Moreover, ideophones can be inflected in class 2 or be introduced by the full form of the verb \textit{ngin} ‘say’ (Hutchison, p. 229), a process clearly parallel to the one found in Afar and Amharic for instance:

(63) kàrugè-nzà (bàdàk) bàdàkcin \\
\hspace{1cm} \text{heart-POS.3SG is beating}
His heart is beating (loudly)

(64) kårógé-nz’ó bádák (bádák) shin
heart-POS.3SG IDEO say.3MSG.IMPFT

His heart is going badak badak

Hutchison assumes that class of ideophones was the first to be inflected in verb class 2.

3.2 Aiki (Maban)

In Aiki, Nougayrol (1989: 141) mentions the existence of verbal periphrases formed with a stem + a verb ir. He translates this verb as ‘*do’, where the asterisk indicates this sense cannot be found outside this particular context (*’do’ is otherwise ees). The data presented here have led him to revise his translation, particularly in view of the fact that the paradigm of ir is, with the exception of the imperative singular form, strictly identical to that of a verb meaning ‘say’ (Nougayrol p.c.). His choice to translate ir as ‘do’ was dictated only by the irregularity of the imperative singular form and by the existence of a French equivalent using faire (‘do’ in French), as in faire vite, faire risette, etc.

The stems used with ir may be Aiki bases, but unlike the situation in Afar or Amharic, no Aiki verb can be used as the base of a compound form. This also constitutes the only way to integrate loanwords into the verb system. Overall, 30% of the verbs recorded in the lexicon are formed solely by this procedure.

Compound verbs can be both transitive and intransitive:

(65) lëmbá dáráp ir
fabric dáráp say

Weave (< Chad Arabic darab čāqa ‘weave’)

(66) jär ir ‘breathe one’s last (human being or animal)’

(67) ajép ir ‘astonish’ (< Arabic)

Furthermore, the verb ir can itself be in base form or in derived forms, including the reciprocal/reflexive form with an -o suffix (used only in this compound form): törık ir-o ‘collide’ (cf. törık ir ‘knock, run against, into’), the factitive form with prefix nind-: bāŋ nind-ir ‘warm up’ (cf. bāŋ ir ‘be warm’), and some rare derived forms with a suffix -e having no definable meaning: jím ir-e ‘rot, decay, turn septic’.

Although rarely used, the verb ees ‘do’ may also form compound verbs. Like ir, ees is used for both transitive and intransitive verbs: kóní ees ‘be ashamed’, gënsít ees ‘hunt’ (cf. Ar. gannis ‘hunting’; also gánás ir ‘hunt’, cf. Ar. ganas ‘hunt’). The latter is the sole example of a doublet with ‘say’ and ‘do’, but the stems are different, the former being a verb, the latter a verbal noun.
3.3. **Hausa (Chadic)**

In Hausa, the situation is similar, though compounding is less widespread and limited in the verb system to some conjugations where the verb yi ‘do’ is used as an auxiliary or a copula with nominals (Abraham 1962 (1946), and Caron p.c.).

In the perfect form, the stem, which is a *nomen actionis*, can only be introduced by the verb yi:

(68) sun yi kuukàa
they.PFT do crying

They cried

In the continuous, one finds the regularly formed inflection with the verbal noun:

(69) yanàa shàafa-n bangòo
he.CONT rubbing-of wall
He is whitewashing the wall

or its counterpart in which the verbal noun of the verb yi is added:

(70) yanàa yi-n shàafa-n bangòo
he.CONT doing-of rubbing-of wall
He is whitewashing the wall

The latter construction with yi is, according to Caron (p.c.), very productive as the marked emphatic counterpart of the former, and can be freely used.

This is similar to the case cited above, where a regularly inflected verb has a counterpart formed with a stem + ‘say’ or ‘do’ (cf. Afar, and Ethio-Semitic).

In addition, yi is a copula in predicative sentences:

(71) yà yi tsàadà
he.PFT do expensiveness
It is (too) expensive

4. **Discussion and conclusion**

This typological study has linked geographically distant and genetically unrelated African languages having three common features: the use of ‘say’ and/or ‘do’ as a means of verb creation, the ability to grammaticalize the resulting compounds into verb inflection, and the absence of restrictions on the choice of the base component.

The involvement of verbs ‘say’ and ‘do’ in the process of renewal of the verb system has been remarked at various synchronic stages. This is an areal linguistic feature: note that the process is unknown in Semitic languages other than those spoken today in East Africa (Ethio-Semitic). This feature may thus be attributed to Cushitic influence.
The auxiliary verbs ‘say’ and ‘do’, used to create new verbs, are integrated into various typological systems which vary with the language. We have thus far identified six such systems. We must, of course, recall that our tentative classification depends mainly on second-hand data not expressly collected to match our framework. Some languages may therefore have to be reclassified in the light of further research, but the sixfold classification below should not be affected:

- the only auxiliary is a verb meaning ‘say’, never ‘do’ (Cushitic: Oromo, Beja; Ethio-Semitic: Zway; Omotic: Kafa; Ancient and Middle Egyptian; Dongolese Nubian; Kanuri);
- the only auxiliary is a verb meaning ‘do’, never ‘say’ (Neo-Egyptian, Coptic; Hausa);
- the verbs ‘say’ and ‘do’ are in complementary distribution (Cushitic: Afar; Ethio-Semitic: Gunnän-Gurage, Argobba, Amharic, Ennemor, Selti, Wolan, Čaha, Muher, Tigre);
- the verbs ‘say’ and ‘do’ vary freely (Cushitic: Afar; Ethio-Semitic: Amharic, Tigre, Harari; Aiki);
- verb compounds with ‘say’ and ‘do’ may alternate with their simple equivalent (Cushitic: Afar, Oromo, Beja; Ethio-Semitic: Amharic, Gunnän-Gurage, Argobba, Tigre; Omotic: Kafa; Hausa);
- ‘say’ and/or ‘do’ undergo morphological de rivation (Cushitic: Afar; Ethio-Semitic: Gafat, Tigrinya, Gogot, Aymellel, Amharic, Ennemor, Selti, Wolan, Čaha, Muher, Tigre; Aiki).

The languages of East Africa bear witness to a recurrent process of grammaticalization, i.e. the use of quotative verbs meaning ‘say’ and active verbs meaning ‘do’ as auxiliaries in order to create new verbs of which they may ultimately become inflectional markers. This phenomenon has been recurring cyclically over more than 5000 years in Hamito-Semitic languages. It is also attested in varying stages of advancement in languages from other areas of Africa. Furthermore, the process exists elsewhere in the world, for instance in Chinese (Waley & Armbruster 1934), in Kambera (Klamer, this volume) for ‘say’, and in pidgins for ‘do’ (Mühlhäusler 1986: 173).

Such an extension in space and time forces the linguist to consider the problem in terms of general linguistics, and not only in terms of areal diffusion, as for Ethio-Semitic and Cushitic. What is at stake here is the morphogenesis of the verb systems insofar as it can be envisaged from the standpoint of a general theory of the syntax and semantics of auxiliarization.

We will now consider how verb periphrases develop and change, then look at their values and functions. We will conclude with a tentative analysis of what
An areal phenomenon in Eastern Africa makes ‘say’ and ‘do’, and verbs of related meanings, suitable for verb creation.

Our survey has shown an absence of restrictions on the terms used with the auxiliaries ‘say’ and/or ‘do’, which differs from languages such as those of the Bantu family (Samarin 1971, Creissels 1997, Güldemann p.c.) or Cushitic Iraqw (Mous 1993: 227-8, and p.c.) where the periphrastic construction is limited to a specific lexical class, viz., ideophones. In the languages considered here, the stem may belong to a range of grammatical classes: bases may be verbs, nouns, adjectives, adverbs, phrases, ideophones, onomatopoeias. These differences reflect semantic restrictions on this periphrasis, or equivalently, restrictions on the ability of the two verbs to act as auxiliaries.

Languages which make use of both auxiliaries may also impose syntactic restrictions whereby the distribution of ‘say’ and ‘do’ may be free or complementary, according to whether intransitive or transitive verbs are to be formed. The data here suggest that only ideophones can be auxiliarized by either ‘say’ or ‘do’ without distinction. Further research is nevertheless needed to confirm this observation and clarify its semantic entailments.

Be that as it may, it would seem that the construction may originate not only with ideophones, as prior literature has remarked, but also with elements such as loanwords which initially have a peculiar status in the language (cf. Aiki and Dongolese Nubian).

Any item, whether its choice be categorically restricted or not, can be constructed with ‘say’ or ‘do’, but when the base component is not morphologically modified, sometimes only the context or the intonation may clarify whether ‘say’ is used as a main quotative verb or as an auxiliary. In Tigré for instance, ከ часа በላ (ex. 48) may either mean ‘to pardon, to excuse’ or ‘he said: ‘pardon!’ In both cases, the syntactic construction is asyndetic but in the latter, both elements keep their own meaning. The difference between a solely asyndetic structure and an auxiliary structure is semantic, not formal: asyndeton involves the sum of two meanings, while the components of an auxiliary structure form a semantic unit.

Such constructions may also lead to the emergence of special verb paradigms where the two components of the periphrasis can still be identified; this is the case in all the contemporary languages studied here. In some of these, some verb inflections could themselves be derived from similar structures whose constituents can only be identified through reconstruction (cf. Egyptian, Coptic, early Cushitic, Kanuri). The range of languages considered thus enables us to envisage the grammaticalization process diachronically at various stages, the last of which is full morphologization.

Synchronic degrees of auxiliarity correspond to diachronic stages of
auxiliarization. By degrees of auxiliarity, we mean the syntactico-semantic constraints that govern the distribution of the auxiliary (restrictions on the class of the base constituent, or on the subject and/or object of the periphrasis). Stages of auxiliarization are the steps leading from main to auxiliary verb status, and perhaps ultimately to full morphologization. Note that this process does not prevent a verb from continuing to function as a main verb. Within our framework, one may consider that the synchronic restriction of the periphrasis to ideophones represents a first stage in the auxiliarization process. A given language may not go beyond this first stage, or any subsequent stage prior to full morphologization.

Thus, starting from what we may call a free creation, insofar as it depends on the speaker’s personal strategy, the scope of auxiliarization may widen progressively until the construction becomes automatic or generalized and enters the language as an idiomatic construction.

Once integrated into the verb system, the periphrasis will be subjected to the internal pressure of this system, and to processes such as derivation, inflection, etc., like any other verb. Likewise, contextual phonetic attrition may affect the auxiliary structure as its frequency of use progresses, opening the way to full morphologization. Thus, in Afar, the juncture between the two components of the periphrasis is rarely audible and it is the unstressed clitic auxiliary that undergoes various assimilatory processes: shortening of geminated consonants, loss of vowels. Only under unusual speech conditions (emphasis, deliberate speech, stopping to look for the next word...) or syntactic constraints (insertion of pronouns and adverbs) will the two components recover clearly separate identities. Yet, the transition to inflectional status is never inevitable and can be thwarted by various factors which modify the organization of the emerging system.

Auxiliarity has thus to be dealt with from a dynamic perspective, not as an invariant phenomenon, but as the result of a process that can be captured at different stages at different moments in the history of a language. ‘Say’ and ‘do’ are not alone in this.

The use of descriptive compounds has a very peculiar and characteristic function in the verb systems of the languages under consideration here. It is well known that tense, aspect, voice, mood, and modal notions in general are values conveyed by auxiliaries (cf. Guillaume 1938, Tesnière 1939, Benveniste 1965, and Heine 1993), but the specific values of ‘say’ and ‘do’ have received little attention in the literature. These are directly linked to a need for expressivity, often a driving force of change affecting forms with a high frequency of use. It is not by chance that ‘say’ and ‘do’ can be used with all verb inflections in
languages of East Africa and elsewhere, and that they have a tendency to be used with all kinds of base constituents, resulting in a doubling of the whole conjugation, or even its replacement. Independently of any specific epistemic modal sense (cf. the attenuative and intensive values in Amharic for instance) linked to particular patterns of auxiliaries or base constituents, descriptive compounds are created to express the relationship between the discourse participants, as we saw in Afar.

The question thus arises why this particular intersubjective modal function happens to be expressed by structures using precisely ‘say’ and ‘do’, and verbs of related meanings including ‘be’, as in some languages such as Saho (cf. Reinisch 1878), ‘say’ and ‘be’ are the same term. ‘Be’ is, of course, the basis of historical change in the verb systems of many of the world’s languages.

From a communication viewpoint, the expressive nature of an ideophone or a loanword may help to explain why a speaker selects a construction with ‘say’: the item is, as it were, “quoted”, hence actualized in the language. Quoting is equivalent to posing an item as existing in the language. To say something is to make it be, to make it exist. A construction with ‘say’ reproduces an external sign standing for an event, be it phonic (as in the case of ideophones) or expressed by any item of the language (for instance a verb). This is parallel to the role of ‘say’ in reported discourse (see Klamer this volume). Moreover uttering a process may be equivalent to achieving it; so, like any performative verb, ‘saying’ can be ‘doing’ any action. At a certain level of abstraction, the three notions of ‘be’, ‘say’, and ‘do’ are identical. This semantic continuum is, as a matter of fact, also found when ‘be’, ‘say’, and ‘do’ are used as main verbs in a certain number of languages. Apart from Saho mentioned above for ‘be’ and ‘say’, other languages may use either ‘say’ or ‘do’ to express either notion. In colloquial French both faire (‘do’) and dire (‘say’) can introduce direct speech. Such is also the case for Maltese ghamel/’do’. On the other hand, in Yemeni dialectal Arabic, it is the verb qāl ‘say’ which may be used in the sense of ‘do, make’. In Soqotri (a modern South Arabian language), the same verb Communist expresses both ‘say’ and ‘do’. As far as periphrases with ‘say’ and ‘do’ are concerned, the possibility of using one or both verbs within a single language is syntactically motivated: in some languages, the valency of the base component is the determining criterion for the distribution of the auxiliaries; in others, this criterion plays no part, and only one of them, most often ‘say’, is used.

Understanding how descriptive compounds are formed not only leads to a better understanding of morphogenesis in a language or group of languages, but also to a more general awareness of how human language works. These features,
An areal phenomenon in Eastern Africa 23

which are common to languages with no genetic connection or areal contact, seem to derive from a process inherent in human language in general.

Notes

* We are indebted to Raymond Boyd, Bernard Caron and Pierre Nougayrol for fruitful discussions, and to Manfred von Roncador and Tom Güldemann for providing us with a rich bibliography, and for valuable comments on this paper. Last but not least, we would like to take this opportunity to thank the Eritrean authorities for facilitating our fieldwork, and all our language assistants, particularly Makki Houmedgaba, Kadiga Abdallah and Addaawa Hasan Ali.


2 The authors’ transcriptions have been retained throughout this paper, except for Egyptian which has been harmonized with Vernus (1988).

3 Compare with Afar (ex. 10c, 10d, 11, 12) which does not have the intensive meaning.

4 Vernus (1988: 172) gives them as all meaning ‘say’.

5 Actually, some Nubian compound forms also attest this process of grammaticalization: some words may have alternative stress on either the base or the verb ‘say’: welêse or welêse ‘leave, let be, abandon’ (Armbruster 1960: 192).

Bibliography


An areal phenomenon in Eastern Africa 24


An areal phenomenon in Eastern Africa 26


Abbreviations

AUTOBEN autobenefactive CAUS causative
An areal phenomenon in Eastern Africa 27

CONT  continuous
DEM  demonstrative
DIR  directional
FACT  factitive
F  feminine
FUT  future
IDEO  ideophone
IMP  imperative
INTR  intransitive verb
IPFT  imperfect
JUS  jussive
lit.  literally
M  masculine
NEG  negation
ONO  onomatopoeia
PART  participial
PAS  passive
PL  plural
PFT  perfect
POS  possessive
RDP  reduplication
REC  reciprocal
REL  relater
SG  singular
SM  subject marker
TR  transitive verb

Index of Languages

Afar  Gogot
Aiki  Gunnän-Gurage
Amharic  Harari
Argobba  Hausa Muher
Aymellel  Kafa
Beja  Kanuri
Čaha  Oromo
Coptic  Seltu
Dongolese Nubian  Tigre
Egyptian  Tigrinya
Ennemor  Wolan
Gafat  Zway