



Cleft sentences and beyond: identification, specification and clause structures in Zaar

Bernard Caron

► To cite this version:

Bernard Caron. Cleft sentences and beyond: identification, specification and clause structures in Zaar: Sans ‘it’, sans ‘COP’, sans ‘REL’, sans everything. 2016. hal-01370125

HAL Id: hal-01370125

<https://hal.science/hal-01370125>

Preprint submitted on 22 Sep 2016

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.

Copyright

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.

CLEFT SENTENCES AND BEYOND: IDENTIFICATION, SPECIFICATION AND CLAUSE STRUCTURES IN ZAAR

SANS 'IT', SANS 'COP', SANS 'REL', SANS EVERYTHING¹

BERNARD CARON

LLACAN (UNIVERSITE SORBONNE PARIS CITE, CNRS)

TABLE OF CONTENTS

1. Introduction	2
1.1. Stating the problem	2
1.2 The Zaar language	3
2. Clefts-Definition (1)	4
3. Predication and Specification	5
3.1 Intuitive contrasting properties	5
3.2. formal properties of ascriptive and specification nominal clauses	7
3.3. Clefts and presentational relative constructions	7
4. Clefts and nominal clauses without copula	9
4.1. Nominal clauses without copula	10
4.2. Clefts without copula	10
5. Universal dependencies	10
5.1. Nominal clauses in UD	11
5.2. Locative nominal clauses	12
Clefts in UD	13
5.3. Presentational relatives in UD	13
6. Clefts – Definition (2)	14
7. Zaar identifying utterances	15
7.1. Unmarked identifying utterance	16
7.2. Marked identifying utterance	17
8. Zaar copulas and identifying utterances	18
8.1. Morphology	18
8.2. Semantics	19

¹ "Sans teeth, sans eyes, sans taste, sans everything". As you like it. William Shakespeare.

This paper was read at the GD1 workshop of the Labex EFL: "*The typology and corpus annotation of information structure and grammatical relations*", Villejuif, September 20, 2016]

1. INTRODUCTION

1.1. STATING THE PROBLEM

The term cleft is commonly used to describe a syntactic pattern which serves to separate a discourse prominent constituent structurally from the rest of the clause. It is formed by dividing a more elementary clause into two parts. One of the two parts is foregrounded, and the other, backgrounded. (Huddleston & Pullum 2008; Hartmann & Veenstra 2013). The following sentences are plain examples of what is called 'cleft' structures in English and French.

- (1.) *It was CHICKEN WINGS that Peter ordered for lunch.* [English]
(2.) *La plupart du temps, c'est L'UTILISATEUR qui a fait une fausse manœuvre.* [French]

The sentence (1) 'It was CHICKEN WINGS that Peter ordered for lunch.' can thus be analysed as:

(3.) *It was CHICKEN WINGS (that) Peter ordered for lunch*
(Cleft PRO) (COP) CLEFTED CONSTITUENT Cleft Clause

The structure is thus characterized by the presence of a proleptic pronoun (it), a copula (was), and a relative clause (that Peter ordered for lunch). This process (foregrounding through cleaving) is not limited to Indo-European languages and can be observed in other languages, e.g. Zaar:

- (4.) a *ka bəl tərtín //*
ka bəl tərti -ín fá:
2SG.FUT dig root PROX indeed
you will dig up this root indeed //
- b *nə təRTÍN >+ ka bəl //*
nə tərti -ín ka bəl fá:
cop1 root prox 2sg.fut dig indeed
(It) is THIS ROOT >+ (that) you will dig > indeed. // (Moral_Har_069)

The non-cleft sentence in (4a) is divided into 2 segments and the second one *tərtín* 'this root', corresponding to the direct object of the verb *bəl* 'dig' in (4a) is foregrounded through left-dislocation and identification with the copula *nə* 'it is'. However, since copulas do not require a subject in Zaar, there is no proleptic pronoun in (4b). Moreover, in (4b) there is no morphological exponent of relativization in the cleft clause *ka bəl* 'you will dig'. A further morphological reduction of the structure is observed in (5) where the left-dislocation of the foregrounded element *gí*: 'this' is not accompanied by a copula.

- (5.) *tô: gí: >+ tətāyá: fū:mí tʃí:* [Zaar]
tô: gí: tətāyá: fū: =mí tʃí: -í:
DM DIST 3PL.REM.ICPL tell 1PL.OBJ thus DIST
'Well it is THIS that they used to tell us like that.'
(lit. 'THIS they used to tell us like that') (Moral_Har_088)

The same construction is observed in (6) in classical Latin :

- (6.) *IN CAUDA >+ venenum //* [Latin]
'(C'est) DANS LA QUEUE (que se trouve) le venin.'

Finally, clefts are to be distinguished from apparently similar constructions whose meaning is different, as in (7) below, to be compared to (2).

- (7.) *C'est un outrage que nous n'acceptons pas.* [French]

Although (7) has the same pronoun, copula, and relative clause as (2), the result is a presentational sentence with an ordinary restrictive relative clause. The difference appears in (8) when cancelling the left-dislocation which is accepted for (8a) but not for (8b):

- (8.) a L'utilisateur a fait une fausse manœuvre.
 b * Nous n'acceptons pas un outrage

I propose to examine what characterises these foregrounding structures beyond the formal components defining them in e.g. English or French, and to find a unifying definition that sets it apart from presentational constructions illustrated in (7). In the process I will argue that this type of syntactic structure is best accounted for within the framework of Universal Dependency Grammar (UD) which only considers content words as governors in dependency relations, thus accounting for the absence of copula in (5 & 6). Finally, I will present a brief description of copulas in Zaar.

1.2 THE ZAAR LANGUAGE

Zaar, also known as Saya, is spoken by about 150 000 speakers in the South of Bauchi State (Nigeria), in the Tafawa Balewa and Bogoro Local Government Areas. Together with 30 or so other related languages first identified by Shimizu (1978), Zaar forms a sub-branch of West Chadic languages named the South-Bauchi languages². Apart from the dominant languages, i.e. English (official national language) and Hausa (dominant all over Northern half of Nigeria), South Bauchi languages are surrounded by Niger-Congo languages in the West (*Izere, Birom*); in the East (*Jarawan Bantu*³); in the South (*Tapshin, Fyem, Kwanka*) and further South-East (*Tarok*). Two isolates inside South-Bauchi languages are *Bankal* in the North and *Boi* in the South.

Most Zaar people of the younger generation are Hausa-Zaar bilinguals. They are schooled in Hausa in primary school, before learning English. The Zaar are Christians and use a Hausa translation of the Bible. The older generation are not fluent in Hausa, whereas the younger educated elite, who often hold positions in the Nigerian administration, police and education, switch comfortably between Zaar, Hausa and English.

From a typological point of view, Zaar shares with its Hausa ‘big brother’ the main characteristics of most Chadic languages: it is a SVO head-first language where TAM is conflated with the exponent of the subject function into a pre-verbal pronominal clitic. Contrary to Hausa, this pre-verbal complex does not include the expression of focus. This same portmanteau morpheme can be omitted in sequential clauses – a phenomenon different from subordination, and appearing in narration to indicate consecutive events – and in Serial Verb Constructions. Zaar uses prepositions and the genitival modifier follows the noun it modifies. There is no case marking of object and subject. Zaar does not use relative pronouns, but has a relative subordinator **DAN**, different from interrogative pronouns. (Caron 2005; 2015)

The 90 min annotated corpus used for this paper was collected in the 1990’s in the village of Tudun Wada, (Bauchi State, Bogoro LGA) where the author had been working for over 10 years and has become part of the social life. The 11 files have been selected to balance genres (3 traditional animal tales; 3 free conversations; 5

² (Newman 1990) classified South-Bauchi languages as the B3 sub-branch of West Chadic. (Newman 2006; 2015) now treat these languages as a third sub-branch (West-C) within West Chadic.

³ The name Jar, or Jarawa is misleading since it refers to different populations, speaking different languages: the Jarawan Dutse (Mountain Jars) speak Zarek (Zere, Zarek, Afizere, Ifizere), a Benue-Congo language, and the Jarawan Kogi (Plain Jars), speaking Jààr (Zhar), a Bantu language, commonly called Jarawan Bantu. Finally, the Jerawa are another population, speaking Zele, a Benue-Congo language from the Kainji group (Shimizu 1975).

extracts from an interview about Zaar history and culture), gender (5 men and 5 women), and age (from 20 to 75). They have been transcribed, using a phonological orthography marking tone and vowel length, and translated into Hausa by M.S. Davan a trained and highly competent native speaker.⁴ The translation into English, alignment and glossing has been done by the author using the Elan-Cortypo programme. (Chanard 2014)

2. CLEFTS-DEFINITION (1)

Clefts are commonly divided into four types, e.g. (Huddleston & Pullum 2008).

- | | | |
|-------|---|-------------------------|
| [9.] | <i>I bought a red wool sweater.</i> | [non-cleft] |
| | <i>It was a red wool sweater I bought.</i> | [it-cleft] |
| | <i>What I bought was a red wool sweater.</i> | [basic pseudo-cleft] |
| | <i>A red wool sweater was what I bought.</i> | [reversed pseudo-cleft] |
| [10.] | <i>The wording of the question confused me.</i> | [non-cleft] |
| | <i>It was the wording of the question that confused me.</i> | [it-cleft] |
| | <i>What confused me was the wording of the question.</i> | [basic pseudo-cleft] |
| | <i>The wording of the question was what confused me.</i> | [reversed pseudo-cleft] |

'Cleft' is a process term: the idea behind it is that a cleft clause is formed by dividing a more elementary clause into two parts. [...] One of the two parts [...] is foregrounded, and the other, backgrounded. Syntactically, the foregrounded element is made a complement of the verb be in its specifying sense – an internal complement in the it-cleft and basic pseudo-cleft, a subject in the reversed pseudo-cleft." (Huddleston & Pullum 2008:1414)

In terms of information structure, as a result of the cleaving, the backgrounded element (the relative clause) is interpreted as 'presupposed', i.e. as "a proposition whose truth is taken for granted or not at issue" (ibid. :1415)⁵

Over the years, this definition, directly inspired by generative syntax studies of the English language has to take into account variations due to languages that don't have a copula or a cleft pronoun, as some languages lack expletive subjects or a copula, or both. (Gundel 2008:70). Ironically, if the cleft pronoun is absent, one ends up with an "it-Cleft" structure with no 'it' in it. This is the case in Zaar where no expletive pronoun, but two copulas (*nə* and *kən*) are used for clefting, with the meaning '(it) is X': *nə X* (cop1); and *X kən* (cop2).

- (11.) *nə tərtín >+ ka bəl > fá: //*
 nə tərti -ín ka bəl fá:
 cop1 root prox 2sg.fut dig indeed
 (It) is THIS ROOT >+ (that) you will dig > indeed. // (Moral_Har_069)

- (12.) "*tô:*" < *tá yísáŋjáy tu [kyá:ni n̩] >+ mbwá:tə //]//*
 tô: té yísáŋj -í: tu kyá:ni kən mbwa: =tə
 dm 3pl.aor know res comp 2s.idp cop2 shoot 3s.obj

⁴ Mr Davan was able to use the competence and knowledge acquired in the process to publish 'Bup Dzanyi Gwaa (Davan 2010), a book about Zaar history, religion and culture, entirely written in Zaar.

⁵ Even if I would not use the terms 'presupposed' but 'preasserted' and 'truth' but 'illocutionary value', on the whole, I am quite comfortable with this analysis.

"Well" they know that [(it) is **you** >+ (who) shot it. //]// (Hunt_Har_047a)

But examples without pronoun or copula are regularly found in the corpus: In (13), no copula is used for the cleft structure

- (13.) "â:" < dzàŋ lá:dì má: <+ **kakáp** >+ má ge:wàyé //
â: dzàŋ lá:dì má: kakáp má ge:wayé -í:
ah day Sunday even everywhere 1pl.fut walk_around res
"Ah" < on Sunday indeed <+ (it is) **EVERYWHERE** >+ (that) we will stroll. // (Girls_A_010)

Zaar also possesses wh-Clefts, also called pseudo-Clefts, where the cleft clause is a free relative clause, which appears in sentence initial position: '*What Peter ordered for lunch was CHICKEN WINGS.*' Example (14) below illustrates the structure in Zaar with the *nə* (COP1) copula:

- (14.) ^àmá: mán yó:dan tʃá: fi <+ **nə mán mársəŋ** //
àmá: mán yó:dan tʃá: fi nə mán mársəŋ
but people which 3pl.icpl do cop1 people Lusa
But the people who did it <+ were the people of Lusa. // (Cal_Har_010)

NB: The it-Cleft equivalent of (14) would be 'But it was THE PEOPLE OF LUSA who did it.'

Cleft structures in Zaar correspond to a single intonation constituent with no internal prosodic break. This is paralleled by a close monosentential syntactic integration of cleft structures. The dependency relationship of the clefted constituent is preserved and no clitic or lexical duplication is needed. In (12), the cleft clause *mbwá:tə*, 'shoot it' has no subject clitic standing for the clefted element *kyâ:n*, 'you' nor does any adverb or lexical equivalent stand for *kakáp*, 'all' in (13). In (5), no COD clitic stands for the clefted element *gí:*, 'this'.

3. PREDICATION AND SPECIFICATION

The semantics of copulas is generally described along a two-way split between predicative (or ascriptive) and specifying uses. (Halliday 1967a; Akmajian 1970; Higgins 1973), illustrated in (15):

- (15.) a The current French president is François Hollande.
b The current French president is short and fat.

3.1 INTUITIVE CONTRASTING PROPERTIES

The contrastive properties of these two sentences are characterized by (Mikkelsen 2006) as:

3.1.1. ABOUTNESS (AKMAJIAN 1979:162–165)

- Like (16), (15b) tells us something about the referent of its syntactic subject:

- (16.) Chris ran a marathon in 3 hours and 8 minutes.
- (15b) doesn't tell us something about its syntactic subject, it tells us who it is.

3.1.2. VARIABLES AND VALUES⁶ (HIGGINS 1979:153FF, 234FF)

- specificational subject introduces variable:

x such that x is the president of France.

- post copular expression provides value for variable: *François Hollande*

3.1.3. FILLING OUT FORMS METAPHOR

The form in (17) filled conventionally in (18) corresponds to specificational copular clauses, whereas unconventional responses in (19) correspond to ascriptive copular clauses.

- (17.) a. Name:
b. Address:
c. Shoe size:
d. Height:
e. Marital status:

- (18.) a. Name: *Bob McPearson*
b. Address: *1 Easy Street*
c. Shoe size: *44*
d. Height: *2 meters*
e. Marital status: *single*

- (19.) a. Name: *difficult to spell*
b. Address: *easy to remember*
c. Shoe size: *a problem*
d. Height: *my advantage*
e. Marital status: *irrelevant*

3.1.4. AScriptive NOMINAL CLAUSES AS LINKS BETWEEN A THEME AND A PROPERTY

A THEME is a semantic role used to describe non-affected direct objects (20 a & b), an entity that is displaced in space or whose position is described in relation to a spatial reference (20c) or a subject of predicative clause (20d).

- (20.) a Zoé aime **les maths**
b. Il faut absolument lire **ce livre**.
c. Mettez **le plat** dans un four préchauffé à 180°.
d. **Paul** est blond

The predicate complement expressing the property is in close semantic relationship with the theme and expresses one of its aspects or qualities.

⁶ NB : the notion of *variable* (noted *x, y*) is borrowed from mathematics where it has several meanings: it can be an arbitrary or value, a role in a predicate, an indeterminate value. In its simplest meaning, a variable is not totally specified, and its value can be unknown, but it must belong to a set. Through the specifying process, an element from that set is chosen that will provide one and only one value, THE value taken by that variable. In the sentence: 'consider an integer x such that x=5', x is the *variable*, and 5 is the *value*.

3.2. FORMAL PROPERTIES OF AScriptive AND SPECIFICATIONAL NOMINAL CLAUSES

Let us see two of the three criteria identified by Den Dikken (2013:36 ff.) for the distinction between ascriptive and specificational copular sentences.⁷ The first one is word-order flexibility, as shown in (21) & (22) below.

Whereas the two major constituents of the specificational sentence in (21) can be inverted, this is not the case in (22).

- [21.] a The winner of the chess tournament is Brian.
b Brian is the winner of the chess tournament.

- [22.] a George is a good chess player.
b *A good chess player is George.

The second criteria is control. In (23a), murderer controls PRO, whereas Ryan doesn't in (23b)

- [23.] a. the murderer_i was insane, besides PRO_i being a bad guy. [ascriptive]
b *the murderer_i is Ryan, besides PRO_i being a bad guy. [specificational]

3.3. CLEFTS AND PRESENTATIONAL RELATIVE CONSTRUCTIONS

To this difference between specificational and ascriptive copular constructions corresponds a contrast between wh-clauses in clefts and presentational relative constructions

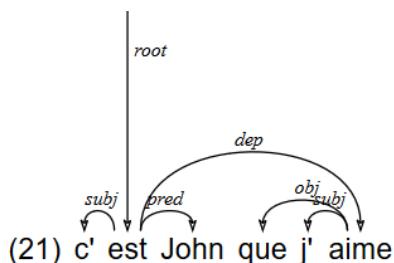
3.3.1. EUROPEAN LANGUAGES

- [24.] *C'est quelqu'un que j'aime beaucoup* [presentational sentence]
[25.] *C'est John que j'aime.* [it-cleft]

In (23), '*que j'aime beaucoup*' is a relative adjectival clause modifying '*quelqu'un*' and depending on it. The '*ce*', subject of the sentence, is referential and not merely expletive. (den Dikken 2013:43). The sentence (weak) focus is on '*beaucoup*'.

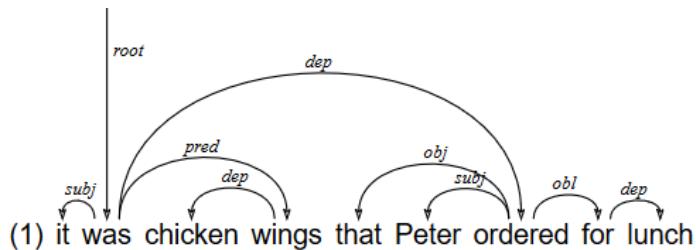
In (24), '*John*' specifies a value for the variable in the wh-clause '*que j'aime*'. The '*ce*' is an expletive subject of the copula. The sentence focus is on the value ('*John*'), while wh-clause containing the variable is backgrounded.

This contrast between (24) & (25) can be represented using the microsyntactic annotation scheme introduced in the Rhapsodie project (Kahane et al. 2013:25 ff.). In (24). In this representation, the wh-clause of the 'it-cleft' is a dependent of the copula, marked dep.

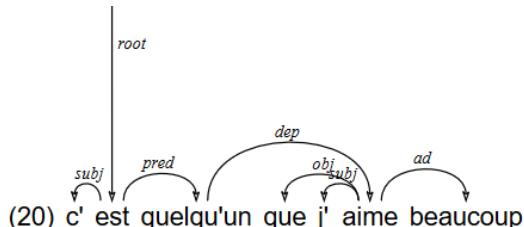


Likewise, the English it-cleft in (1) can be represented as follows:

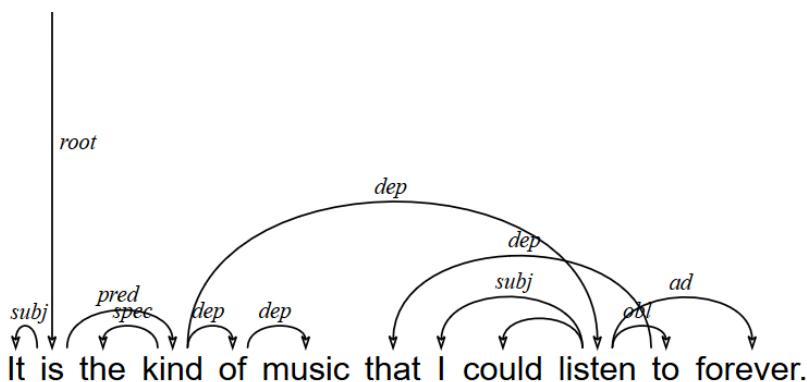
⁷ The third one is too formal and theory-dependent for us to use in this work.



In presentational relative clauses, the wh-clause is an adjectival clause (*dep*) dependent of the predicate complement (*pred*).



[26.] It's the kind of music that I could listen to forever.



3.3.2. ZAA'R

The same contrast between basic cleft and presentational constructions is found in Zaar.

A. ZAA'R BASIC CLEFT CONSTRUCTION

[27] exemplifies a basic cleft construction, in a conversation where the speaker lists the various masquerades and the clothes that give them their identity. It opposes thatch that characterises the clothes of the masquerade Murgi to various tree leaves that define other masquerades.

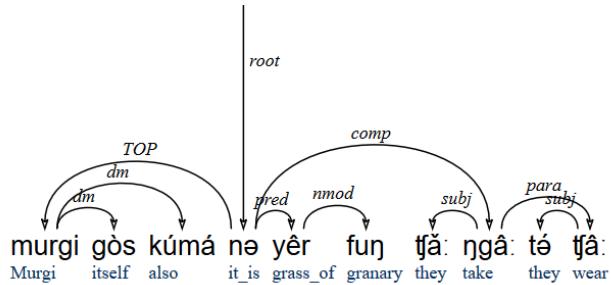
[27.]	<i>murgi gòs kúmá [...]⁸ nə yêr FUN >+ tʃā: ngâ: tá tʃā: } //</i>
	murgi gòs kúmá ngétn -én e: tún
	disease_sp 3SG.CTR and thing PROX er since
	nə yêr fuŋ tʃā: nga: té tʃā:
	COP1 thatch granary 3PL.ICPL take 3PL.AOR wear

⁸

We have omitted the passage where the old speaker stutters and fumbles for words.

'As for Murgi, [...] . it is THATCH that they take and wear.' (Rel_Har_079)⁹

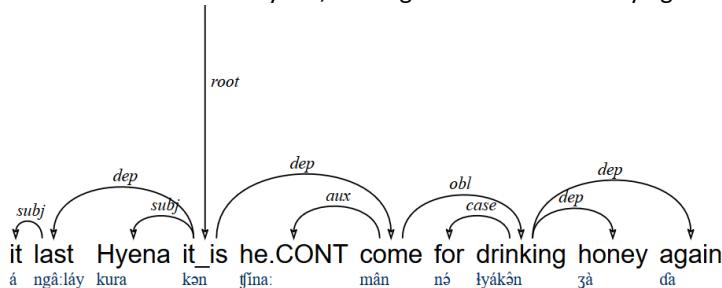
The copula (*nə*, 'it is') links a value (*yér funj*, 'thatch') to a variable in a clause (*tʃá: ngâ: tâ tʃá:*, '(that)' they take (and) they wear'. The value 'thatch' is selected in the set of values that belong to the various grasses and leaves used to dress up the masquerades. (26) can be paraphrased as: In the function $y = f(x)$, 'thatch' is the value y ; for the function $\text{take}(x)$; $\text{wear}(x)$, i.e. 'what they take' & 'what they wear'. Structurally speaking the copula *nə* has two complements : the predicative complement *yér funj* and the specifical complement *tʃá: ngâ: tâ tʃá:*.



B. ZAAR PRESENTATIONAL RELATIVES

(28) exemplifies construction where the copula *kən* has a presentational value, and the property (*tʃína: mân ná t'yáyâñ ʒà da* 'he is coming to drink honey again') is conveyed by a clause that modifies the predicative complement *kura* ('Hyena').

- [28.] á ngâ:láy <+ kura yən tʃína: mân ná t'yáyâñ ʒà da //
 á nga:lá -í: kura kən tʃína: mán né t'ya -kénì ʒà da
 3SG.AOR spend_time -RES hyena COP2 3SG.REM.CTF come for drink -NMLZ water again
 after a while there was Hyena, coming to drink some honey again. (Mbrt_S1_123)



4. CLEFTS AND NOMINAL CLAUSES WITHOUT COPULA

In French, English etc. this relation between theme and property is established with a verb of the 'be' type, i.e. either the verb *be* itself that bears little semantic component except for this linking function, or copula, or verbs that add to the link extra temporal or aspectual modalities, e.g. become, stay, seem, etc. However, many (if not most) languages do not need a verb for such clauses, e.g. signed French (Nève de Mévergnies 1996:316) or classical Latin (ibid :317).

⁹

'Murgi' is the name of a masquerade clothes are made of thatch.

4.1. NOMINAL CLAUSES WITHOUT COPULA

- [29.] *Homo homini lupus*
man-NOM.SG wolf-DAT.SG wolf-NOM.SG
'Man is wolf to man'
- Some Zaar nominal clauses are similar to Latin, in that they lack a copula in their predication.
- [30.] *ləpm za:r <"m:::"' gistá //*
ləpm za:r m: gisté
festival Zaar fill harvest_season
'The Zaar festival, er... (it's in) the dry season.' (Cal_Har_068)

- [31.] *gí: < ña: la: bastə //*
gí: ña: la: bas =tə
dist small work at =3sg.obj
'That (is) small work for him.' (i.e. 'not a problem'). (Girls_B_094)

- [32.] *myà:ní guði < mì gòpm < ngá jayndí gip dàn //*
myà:ní guði mì gòpm ká jakndí gip dàn
1pl woman.PL 1pl.sbj 1pl.pos at cooking_pot inside house
'We women we are by the cooking pot inside the compound.' (Wom_B_214)

4.2. CLEFTS WITHOUT COPULA

In sentences where nominal clauses without copulas exist, the same is found in clefts. It is observed in Latin but also in Zaar.

- [33.] *in cauda >+ venenum //* [Latin]
in cauda venenum
dans queue-NOM.SG venin-NOM.SG
'(It's) in the tail (that) the venom (is situated).'

- [34.] *limêš má: < myâ:m >+ matá tya //* [Zaar]
lim -êš má: myá:ni matá tya
six -DEF even 1SG 1SG.REM drink
'The sixth even I am the one who scored.' (Boys-A_407)

- [35.] *gerjí < kusun kawéy >+à: dak jí //* [Zaar]
gèri-sə kusun kawéy à: dak =jí
chicken-pl hunger merely 3SG.CPL refuse =3PL.OBJ
'The hens, (it's) only hunger (that) is disturbing them.' (Wom_A_96)

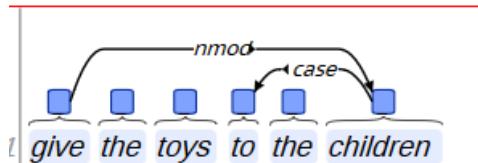
- [36.] *"wókè:" kyâ:n >+ kyá: tfa:tayá > njá:n //* [Zaar]
wókè: kyá:ni kyá: tfa: =tə káy njá:n
OK 2S 2SG.COND put 3SG.OBJ LOC QUEST
'"OK, it's you who made him do it, isn't it?"' (Hyena_S1_319)

5. UNIVERSAL DEPENDENCIES

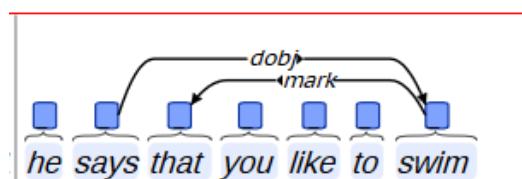
Universal Dependencies provide the guidelines for a unified syntactic analysis of predicative constructions that accounts for the absence of copula in languages like Latin, Russian, or Zaar.

In dependency grammars, which derive from Tesnière's initial work (Tesnière 1934; 1959), syntactic annotation consists of typed dependency relations between words instead of constituents as is the case in the Xbar theory

(Jackendoff 1977) and other derived syntactic analyses. In the resulting tree, each word is either the dependent of another word in the sentence or of a notional root of the sentence (Kahane 1997)¹⁰. The goal of the typed dependency relations is a set of broadly observed “universal dependencies” that work across languages. In the *Universal Dependencies* version which is becoming the standard for formal dependency grammars (de Marneffe, Dozat, et al. 2014; de Marneffe, Ginter, et al. 2014; Nivre et al. 2016; Gerdes & Kahane 2016). Dependency relations hold primarily between content words, rather than being indirect relations mediated by function words. The primacy of content words implies that function words (e.g. prepositions, conjunctions) normally do not have dependents of their own, e.g. ‘to’ depends on ‘the toys’ in the ex. below



And ‘that’ is dependent on ‘swim’ in the ex. below.



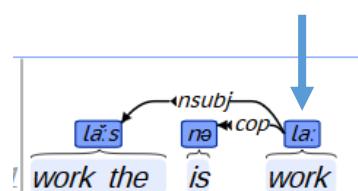
5.1. NOMINAL CLAUSES IN UD

In the UD frame, in nominal clauses, the predicate (and root) is a noun or an adjective, which takes a single argument with the nsubj (nominal subject) relation. The copula verb (if present) attaches to the predicate with the cop relation. (de Marneffe, Ginter, et al. 2014)



The same representation holds for Zaar nominal clauses with a copula:

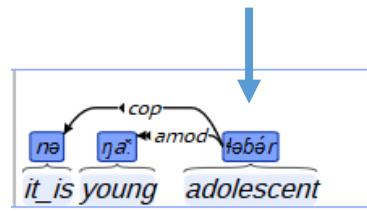
[37.]	<i>lă:s nə la: //</i>
	la: -ēs nə la:
	work -DEF COP1 work
	this work is serious.(Boys-B_065)



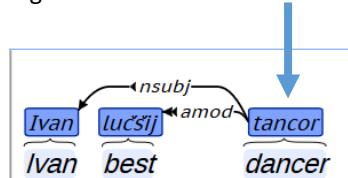
However, in my corpus, most Zaar ascriptive nominal clauses have no subject.

¹⁰ See (Kahane 2001) for the developments of formal dependency grammars, and (Hays; Hudson 1991; Mel'čuk & Pertsov 1987; Mel'čuk 1988; lordanskaja & Mel'čuk 2009) for implementations of these theories in various languages.

- [38.] *nə ñă: təbér //*
 nə ñă: təbér
 cop1 small young_man
 It's a young man. (Bury_Har_214)

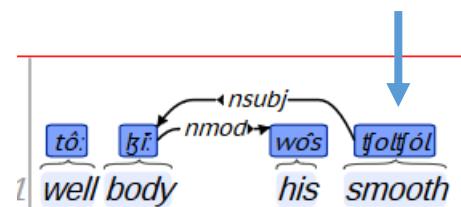


This analysis accounts as easily and naturally for nominal clauses without a copula. This is the case in Russian, e.g.



Zaar nominal clauses without copula are represented in the same way.

- [39.] *'tô:' b̥i:wôs tʃoltfól //*
 tô: b̥i: =wôs tʃoltfól
 dm body =3sg.pos very_smooth
 'Well, his body is very smooth.' (Wom_A_090)



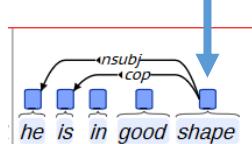
5.2. LOCATIVE NOMINAL CLAUSES

The current version of the UD frame gives a special treatment of locative nominal clauses.

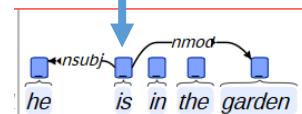
This analysis of copula constructions extends to adpositional phrases and oblique nominals as long as they have a predicative function. By contrast, temporal and locative modifiers are treated as dependent on the existential verb "be". (de Marneffe, Ginter, et al. 2014: Specific constructions)

See (40) where 'in good shape' is the predicate of the clause, whereas 'in the garden' depends on 'be' in (41)¹¹.

- [40.] *he is in good shape*



- [41.] *he is in the garden*



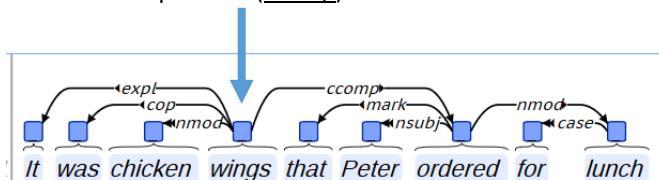
¹¹ Although this analysis is not supported by any argument by (de Marneffe, Ginter, et al. 2014), Zaar would corroborate this special treatment, since locative clauses are not nominal but use the lexical verb *yi*, 'be'. The only problem comes from *da*, a particle that can be used on its own to mean 'exist', e.g. *tʃokn da* 'God exists'. As in this case, the particle *da* bears the illocutionary stress, (whereas in the clause *kádi kán* 'it is a dog', the stress falls on *kádi* 'dog'), I would say *da* is the predicate/root of the clause, with *tʃokn* as its nominal subject; and *kádi* is the predicate/root with *kán* as its copula. The same applies for other nominal locative clauses, e.g. with the prepositions *bas*, etc. (cf. ex. (30 & 31) above).

CLEFTS IN UD

Consider example (42).

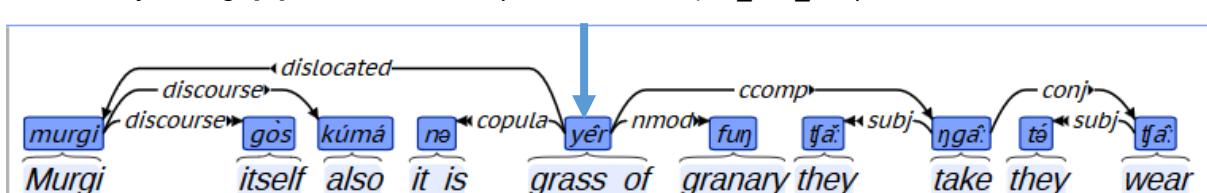
- [42.] *It was chicken wings that Peter ordered for lunch.*

The clefted constituent ('chicken wings') corresponds to the illocutionary nucleus, and is the root of the clause. 'it' is an expletive pronoun (expl) linked to the root, just like the copula. The cleft clause ('*that Peter ordered for lunch*') is a clausal complement (ccomp) of the root.



The same analysis applies to Zaar in (43):

- [43.] *murgi gòs kúmá [...] nə yér funj >+ tʃā: ɳgā: tā tʃā: } //*
 murgi gòs kúmá ngétn -én e: tún
 disease_sp 3SG.CTR and thing PROX er since
 nə yér funj tʃā: nga: té tʃā:
 COP1 grass granary 3PL.ICPL take 3PL.SBJV put
'As for Murgi, [...] it is thatch that they take and wear.' (Rel_Har_079)¹²

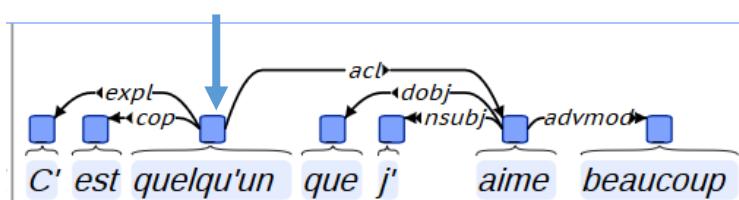


5.3. PRESENTATIONAL RELATIVES IN UD

Consider the presentational sentence in (44):

- [44.] *C'est quelqu'un que j'aime beaucoup.*

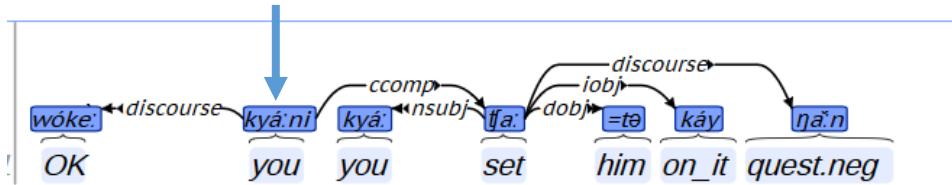
The presented constituent 'quelqu'un' is the nucleus with its dependent copula and expletive pronoun 'c'est'. the difference is the wh-clause 'que j'aime beaucoup' which is a "garden variety restrictive relative clause" (den Dikken 2013), which functions as an adjectival clause (acl) depending on the root.



¹² We have omitted the passage where the old speaker stutters and fumbles for words.

¹³ 'Murgi' is the name of a masquerade clothes are made of thatch.

- [45.] "wókè:" kyâ:n >+ kyá: tfa:təyáy > ηă:n //
 wókè: kyá:ni kyá: tfa: =tə káy ηă:n
 OK 2S 2SG.COND put 3SG.OBJ LOC QUEST
 "OK, it's you who made him do it, isn't it?" (Hyena_S1_319)



As can be seen, as a result of the fact that the root of the sentences is the nominal predicate in all cases, the only difference in the analysis of (44) (=presentational relative) and (42, 43) (=cleft) is in the type of dependency that links the dependent (wh-) clauses to the nominal predicate: it is marked as acl (=adjectival clause) in (43) and a ccomp (clausal complement of the predicate) in (41, 42).

6. CLEFTS – DEFINITION (2)

A better characterization needs to be found for clefts that does not rely so heavily on the morphology and syntax of European languages. It can be found Halliday's concept of IDENTIFICATION, and the term 'IDENTIFYING CLAUSE' (Halliday 1967a:223ff), where the only morphological component that is retained is that of nominalization.

Halliday explains that any clause such as *John saw the play* can be organised into a 'cleft sentence' with equative form (i.e. of the form 'x equals y' as in *the leader is John*) through the nominalisation of one set of its elements, e.g. *what John saw was the play*. The former, without the nominalisation is non-identifying and the second is identifying. The identifying clause adds the further information that one of the participants is definable by participation in the process. In an identifying clause, it is always the nominalization which is 'to be identified'. (*op.cit.* 224)

Halliday's 'identification' and 'equation' is what he calls the 'class 2 be' which means '*identifies or is identifiable as, can be equated with*' and which was characterized in section 3 above as 'specification'. The assertive predication is described under 'class 0' by Halliday while 'class 1' covers spatial and temporal predictions. (Halliday 1967b:66)

Halliday continues by establishing the systematic equations:

- *what John saw* = nominalisation = identified = given = variable;
- *the play* = identifier = new¹⁴ = value:

There is thus an association of variable – value with theme – rheme similar to that of identified – identifier with given – new: in the unmarked case, the identified is given, the identifier new, and the variable is theme, the value rheme. [...] in a sense a theme is a variable to which a value is to be

¹⁴ Instead of new/given, I prefer to say that the identifier is the illocutionary nucleus of the sentence, and that the identified is pre-asserted. Moreover, the fact that the identifier is 'new' as opposed to the identified which is 'given' needs to be qualified. If it applies to *contrastive*, or *stressed-focus it-clefts* (e.g. *what got you interested in clefts? – it was Brian's book that got me interested in clefts*) it is not true in the case of so-called '*continuous-topic it-clefts*' (*do you know Brian's book? – yes in fact, it was Brian's book that got me interested in clefts.*) (den Dikken 2013:62).

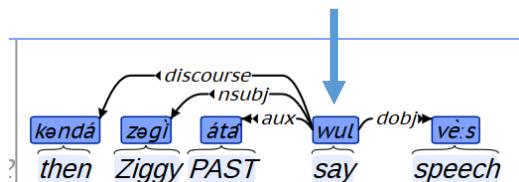
assigned. But as always the speaker may exploit the contrastive possibility of not mapping the variable on to the theme; hence to the unmarked, operative [Type(1) what John saw was the play] corresponds a marked, receptive form [Type(2) it was the play that John saw]. (op.cit. 228)

I propose to name type (i) UNMARKED IDENTIFYING CLAUSE (aka *wh*-cleft,) ; and type (ii) MARKED IDENTIFYING CLAUSE (aka *contrastive it*-cleft). The two types are illustrated below in section 7 for Zaar.

7. ZAAR IDENTIFYING UTTERANCES

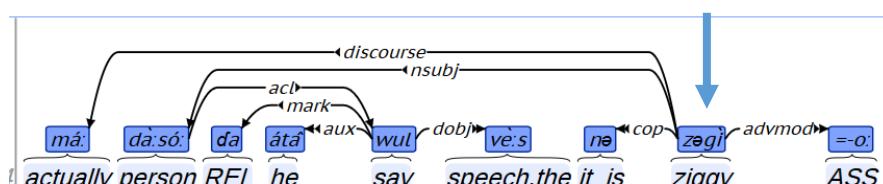
An identifying clause (IC) is defined in section 6 as equating a variable in a nominalised clause (called the identified IDed) to a value given by an NP (the identifier IDer). The typical ICs in Zaar are exemplified below

Starting from the non-IC (46) where the root of the utterance is the verb *wul* 'say':

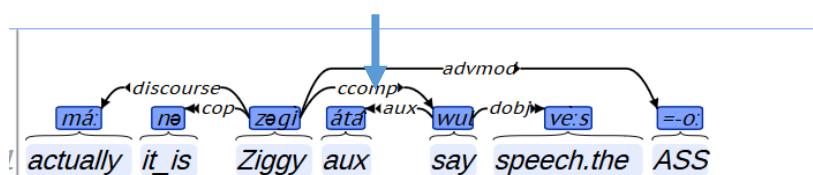


- | | | | | | | | |
|-------|--|---|-------|---------|-----|-------|------|
| [46.] | <i>^kəndá zəgì átâ wul vè:s // / /</i> | kəndá | zəgì | átâ | wul | vì: | -êš |
| | | then | Ziggy | 3SG.REM | say | mouth | -DEF |
| | | 'Then, Ziggy spoke.' (lit. 'said the speech') | | | | | |

In the corresponding unmarked IC in (47) the root is the IDer, i.e. the nominal predicate *zägi*, ‘Ziggy’, and the IDed is the nominal subject (*nsubj*) *dà:só*: ‘the man’, and its adjectival clause modifier (*acl*) *dă:tâ wul vè:s*, ‘who spoke’, *lit.* ‘the person who said the speech’. The link identified – identifier is done through the copula (*cop*) *na* ‘it is’.



The corresponding marked IC is (48) where the root is still the IDer 'Ziggy' but this nominal predicate is now subjectless, and the identified is now its clausal complement (ccomp) *átâ wul yè:s*. '(who) spoke'.



- | | | | | | | | |
|-------|--|-------|---------|-----|-------|------|------|
| [48.] | "má:" nə zəgì >+ átâ wul vè:sò:_//] // | | | | | | |
| má: | nə | zəgì | átâ | wul | vì: | -êš | -o: |
| even | COP1 | Ziggy | 3SG.REM | say | mouth | -DEF | -FCT |

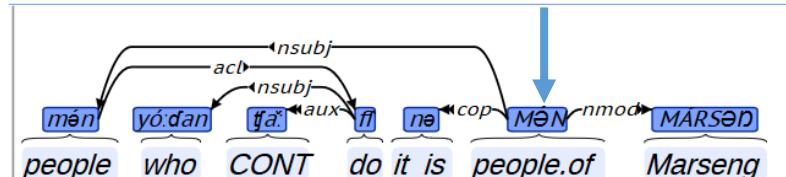
Actually, it is ZIGGY who spoke. (Boys-A_455)

Sections 7 will examine the various types of marked and unmarked ICs in Zaar.

7.1. UNMARKED IDENTIFYING UTTERANCE

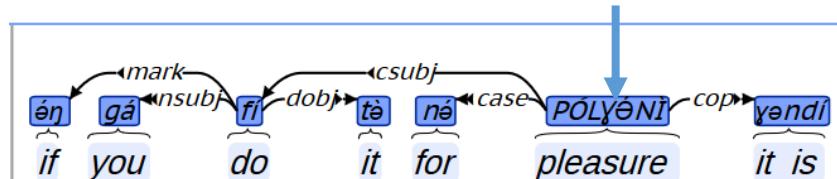
Two (or no) copulas appear with ICs in Zaar: *nə*, *kən* (var.: *kəndá/kəndí/kənín*) and \emptyset .

7.1.1. NӘ



- [49.] ^àmá: mán yó:dan tďă: fi <+ nə MÁN MÁRSƏY //
àmá: mén yó:dan tďă: fi nə mén mársəy
but people which 3PL.ICPL do COP1 people Lusa
^But the people who did it <+ were THE PEOPLE OF LUSA. (Cal_Har_010)

7.1.2. KƏN



- [50.] [“è:’ tō:’ lăpm za:r < (’yâ:n t̪i tu::: á vònì)] éŋ gá fítè <+ ná PÓLYÉNÌ yəndí //
 ín ká fi tə ná pol -káni kändí
 if 2PL.FUT do 3S.OBJ for please NMLZ COP2
 [“Yes” “Well” the Zaar festival < (that is to say... every year)] if you do it <+ it is for PLEASURE //
 (Cal Har 051)

Compare the paradigm of possible ICs :

- [51.] a *ká fítà ná pólýáni* [Non-IC]
'you will do it for pleasure'

b *ín ká fí tå <+ ná PÓLKÁNÌ kandí* [Unmarked IC]
'if you do it <+ it is for pleasure indeed'

c *ná PÓLKÁNÌ >+ ká fítà* [Marked IC]
'(it is) for pleasure (that) you do it'

7.1.3. ZERO

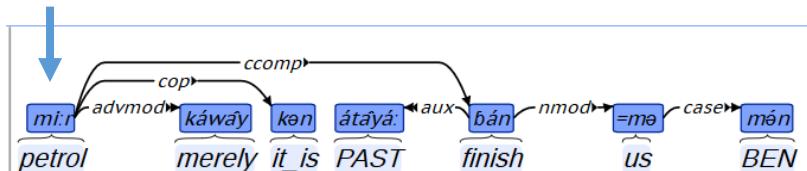


- [52.] [háŋkalí <+ à: mbí: tʃigòskáy tu] kówây tè tú: <+ tó: BÁBÍ: KÚRÛM TÈ DÁ:DÍ: //

merely 3SG.SBJV get DM falling -RES merely 3SG.SBJV swallow -RES
 [Gradually, he started thinking that] (what) he only needed to do <+ (was) TO OVERTHROW HIM AND SWALLOW HIM. (Hyena_S1_366)

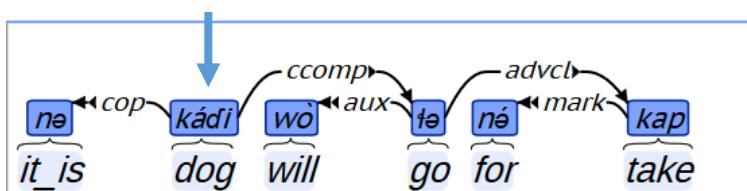
7.2. MARKED IDENTIFYING UTTERANCE

7.2.1. KƏN



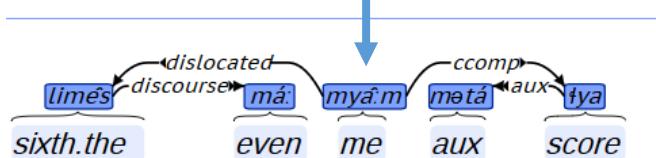
- [53.] MI:R káwây yəŋ >+ átâyá: báñəm mán //
 mi:r káwây kən átâyá: báñ =mə mán
 petrol merely =COP2 3SG.REM.ICPL finish =1SG.OBJ BEN
 It's only THE PETROL >+ (that) got finished. (Boys-A_255)

7.2.2. NƏ



- [54.] nə KÁDI >+ wò tág::: kap //
 nə kádi wò tə né kap
 COP1 dog 3SG.FUT go for take
 'It is DOG >+ (who) will go to ... take (them).' (Hyena_S1_051)

7.2.3. Ø

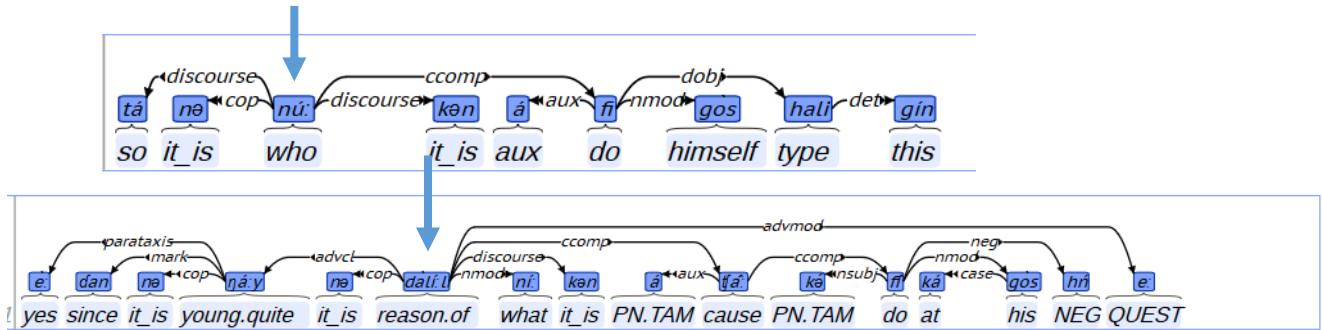


- [55.] limés "má:" < MYÁ:M >+ métá týa //
 lim -é:s má: myá:ni métá týa
 six -DEF even 1SG 1SG.REM drink
 'The sixth even < I (am the one who) > scored.' (Boys-A_407)

7.2.4. NƏ + KƏN

The only two examples in the corpus are found in wh-questions:

- [56.] [^tá nə NÚ:ŋ >+á fi gós hali gín //= [...]]
 tá nə nú: kən á fi gós hali gín
 then COP1 who =COP2 3SG.AOR do 3SG.POS character PROX
 Well, WHO is it >+ (who) did this thing //=[and ...] (Hyena_S1_414)



[57.] 'è:' dān nə ñá:y <+ nə DÀLÍ:LI Níŋ >+ á tʃá: { yá::: /ká } fí yá gòs hñé: //

è:	dān	nə	ñá:	-i	nə	dàlí:li	ni:	kən	á	tʃá:
er	as	COP1	child	INDEF	COP1	reason	what	COP2	3SG.AOR	put
									2PL.AOR	2PL.AOR
									do	at
									3SG.POS	NEG2
									PN.TAM	Q

Yes, since he is young, what is the reason why (lit. the REASON OF WHAT is it?) you... didn't do (it) at his (burial)?

8. ZAAR COPULAS AND IDENTIFYING UTTERANCES

(Site under construction)

Beside the absence of any verb, as seen above, there are several words in Zaar that can stand for the Indo-European ‘be’ verb. : yi, nə, kən, and their variants and combinations. Are they specialised in any function/uses? Do they have other uses than copular? Can they shed any light on what we have seen up to now?

8.1. MORPHOLOGY

Morphologically, yi is a defective verb (a quasi-auxiliary) while kən and nə are syntactic particles.

- Yi combines with TAM & Person ‘subject pronouns’ (typed nsubj or aux in the UD trees).
- Yi has a derived (“causative”) form yir meaning ‘have, possess’.
- Yi combines with many prepositions : yi da, yi káy ‘exist’ ; yi té ‘have’
- Yi+ ká [jíyá] + Verbal Noun, lit. ‘be at --ING’ has formed the continuous aspect.
- The combination átāyi ‘3SG+remote past+be’ has been lexicalised into a temporal adverb meaning ‘formerly’
- Kən & nə are invariable particles that are insensitive to TAM and person. When the clause has TAM and Person modalities, these are borne by yi, yielding the combination: PN.TAM + yi + na/kən. The frequent combination yi+nə has produced the variant yin, whose meaning is hardly different from yi or nə on their own.
- Kən has 4 variants : kən, kəní, kəndí, kəndá
- Kən can be combined with deictic suffixes yielding more variants (kənín, kəntʃín, kəndí:, kəntʃí:, kəndídí:, kəntʃídí:)

8.2. SEMANTICS

8.2.1. *YI* = LEXICAL (DEFECTIVE) ‘BE’ VERB

- The basic meaning of *yi* is ascriptive.
- When modified by a preposition, it becomes locative (*yi ká*) possessive (*yi tá*) or existential (*yi da*)¹⁵.

8.2.2. *KƏN* = COP2

- The basic meaning of *kən* is specifical, and is often used in ICs (cleft clauses).
- It is used as an emphasis particle at the end of an intonation unit (e.g. Topic) or of an utterance. In the latter case it emphasises the whole utterance.
- The *kəndí* variant can be used as an interjection with the meaning ‘That’s it!', Voilà!'.

8.2.3. *NƏ=COP1*

- *Nə* has both ascriptive and specifical meanings, and it combines the uses of *yi* and *kən*
- It is the default copula for attributive adjectival constructions
- It appears as well in ICs, but not as often as *kən*.

8.2.4. *NƏ + KƏN*

The combination is quite common. In this case *nə* assumes the copulative function and *kən* the emphatic one.

REFERENCES

- Akmajian, Adrian. 1970. Aspects of the grammar of focus in English. Massachusetts Institute of Technology Thesis.
<http://dspace.mit.edu/handle/1721.1/15655> (14 September, 2016).
- Caron, Bernard. 2005. *Za:r (Dictionary, grammar, texts)*. Ibadan (Nigeria): IFRA.
- Caron, Bernard. 2015. Zaar Grammatical Sketch. In Amina Mettouchi, Martine Vanhove & Dominique Caubet (eds.), *The CorpAfroAs Corpus. ANR CorpAfroAs: a Corpus for Afro-Asiatic languages*.
- Chanard, Christian. 2014. *ELAN-CorpA-V4.7.3*. http://llacan.vjf.cnrs.fr/res_ELAN-CorpA.php.
- Davan, M. S. 2010. *Bup Dzanyi Gwaa*. Ibadan (Nigeria): Ifra-Nigéria.
- Dikken, Marcel den. 2013. Predication and specification in the syntax of cleft sentences. In Katharina Hartmann & Tonjes Veenstra (eds.), *Cleft structures*, 35–70. Amsterdam: J. Benjamins.
- Gerdes, Kim & Sylvain Kahane. 2016. Dependency Annotation Choices. Assessing Theoretical and Practical Issues of Universal Dependencies. Berlin (Germany).
- Gundel, Jeanette K. 2008. Contrasting perspectives on cleft sentences. In María de los Ángeles Gómez-González, J. Lachlan Mackenzie & Elsa González Álvarez (eds.), *Languages and cultures in contrast and comparison*, 69–8. Amsterdam, Pays-Bas.
- Halliday, M. A. K. 1967a. Notes on Transitivity and Theme in English: Part 2. *Journal of Linguistics* 3(2). 199–244.
- Halliday, M. A. K. 1967b. Notes on Transitivity and Theme in English: Part 1. *Journal of Linguistics* 3(1). 37–81.
- Hartmann, Katharina & Tonjes Veenstra (eds.). 2013. *Cleft structures*. Amsterdam: J. Benjamins.
- Hays, David. Dependency theory: A formalism and some observations. *Language* 40(4). 511–525.
- Higgins, Francis Roger. 1973. The pseudo-cleft construction in English. Massachusetts Institute of Technology Thesis.
<http://dspace.mit.edu/handle/1721.1/12988> (14 April, 2016).

¹⁵ Pronominal subjects are fused with the ‘be’ verb *yi* yielding a pseudo-paradigm of subject pronouns of nominal clauses: sg: *mi* (< mə-yi), *ki* (kə-yi), *tʃi* (< tə-yi); pl: *mì* (mə-yi), *kì* (< kə-yi), *tʃì* (< tə-yi), e.g. *mi/mì tá kádī*, ‘I/we have a dog’, lit. ‘I/we-be with dog’.

- Huddleston, Rodney & Geoffrey K. Pullum. 2008. *The Cambridge Grammar of the English Language*. 2nd ed. Cambridge: Cambridge University Press.
- Hudson, Richard A. 1991. *English word grammar*. Oxford, UK; Cambridge, Mass., USA: B. Blackwell.
- Iordanskaja, Lidija & Igor Mel'čuk. 2009. Establishing an Inventory of Surface-Syntactic Relations: Valence-controlled Surface-Syntactic Dependents of the Verb in French. In Alain Polguère & Igor Aleksandrovič Mel'čuk (eds.), *Dependency in linguistic description*, 151–234. Amsterdam, Pays-Bas, Etats-Unis d'Amérique.
- Jackendoff, Ray. 1977. *X Syntax: A Study of Phrase Structure*. Cambridge, Mass: MIT Press.
- Kahane, Sylvain. 1997. Bubble trees and syntactic representations, in Becker & Krieger (eds), Proc. 5th Meeting of the Mathematics of Language (MOL5), : , 70–76. In T. Becker & H.U. Krieger (eds.), *Proceedings of the 5th Meeting of the Mathematics of Language (MOL5)*, 70–76. Saarbrücken: Deutsches Forschungszentrum für Künstliche Intelligenz.
- Kahane, Sylvain. 2001. Grammaires de dépendance formelles et théorie Sens-Texte. *Actes, TALN 2001, Tours*, 2–5 juillet, 63. Tours.
- Kahane, Sylvain, Kim Gerdes, Paola Pietrandrea & Christophe Benoitoun. 2013. *Protocol for micro -- syntactic coding*. http://www.projet-rhapsodie.fr/plus/tutoriels/doc_download/201-protocol-for-micro-syntactic-coding.html.
- Marneffe, Marie-Catherine de, Timothy Dozat, Natalia Silveira, Katri Haverinen, Filip Ginter, Joachim Nivre & Christopher D. Manning. 2014. Universal Stanford Dependencies: a Cross-Linguistic Typology. *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*. Reykjavik, Iceland: European Language Resources Association (ELRA).
- Marneffe, Marie-Catherine de, Filip Ginter, Yoav Goldberg, Jan Hajíč, Christopher D. Manning, Ryan McDonald, Joakim Nivre, et al. 2014. Universal Dependencies. Online documentation (Version 1). <http://universaldependencies.org/> (16 June, 2016).
- Mel'čuk, Igor' A. & Nikolaj V. Pertsov. 1987. *Surface syntax of English: a formal model within the meaning-text framework*. (Ed.) Richard Kittredge. (Linguistic & Literary Studies in Eastern Europe v. 13). Amsterdam ; Philadelphia: John Benjamins Pub. Co.
- Mel'čuk, Igor A. 1988. *Dependency syntax : theory and practice*. Albany: State University of New York Press.
- Mikkelsen, Line. 2006. Specificational copular clauses. Paper presented at the Kobe Area Circle of Linguistics, Kobe Shoin Women's University. http://linguistics.berkeley.edu/~mikkelsen/talks/Shoin_ho.pdf (13 September, 2016).
- Nède de Mévergnies, François-Xavier. 1996. *Essai de grammaire de la langue des signes française*. Liège, Belgique, Suisse: Bibliothèque de la Faculté de philosophie et lettres de l'Université de Liège.
- Newman, Paul. 1990. *Nominal and Verbal Plurality in Chadic*. Berlin: Walter de Gruyter.
- Newman, Paul. 2006. Comparative Chadic revisited. In Paul Newman & Larry M. Hyman (eds.), *West African linguistics: papers in honor of Russell G. Schuh*, 188–202. (Studies in African Linguistics: Supplements 11). Columbus, Ohio: Published by the Department of Linguistics and the Center for African Studies, Ohio State University.
- Newman, Paul. 2015. *Comprehensive Bibliography of Chadic and Hausa Linguistics, Third Edition*. Paul Newman. <https://scholarworks.iu.edu/dspace/handle/2022/20576> (21 January, 2016).
- Nivre, Joakim, Marie-Catherine de Marneffe, Filip Ginter, Yoav Goldberg, Jan Hajic, Christopher D. Manning, Ryan McDonald, et al. 2016. Universal Dependencies v1: A Multilingual Treebank Collection. In Nicoletta Calzolari (Conference Chair), Khalid Choukri, Thierry Declerck, Sara Goggi, Marko Grobelnik, Bente Maegaard, Joseph Mariani, et al. (eds.), *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*. Paris, France: European Language Resources Association (ELRA).
- Shimizu, Kiyoshi. 1975. A Lexicostatistical Study of Plateau Languages and Jukun. *Anthropological Linguistics* 17. 413–418.
- Shimizu, Kiyoshi. 1978. *The Southern Bauchi Group of Chadic Languages: A Survey Report*. (Africana Marburgensia Special Issue 2). Marburg.
- Tesnière, Lucien. 1934. *Grammaire de la langue Bulgare*. Paris, France.
- Tesnière, Lucien. 1959. *Eléments de syntaxe structurale*. Paris: Klincksieck.