



HAL
open science

Naija grammatical classes and constructions

Bernard Caron, Sylvain Kahane

► **To cite this version:**

Bernard Caron, Sylvain Kahane. Naija grammatical classes and constructions. Society for Pidgin and Creole Linguistics 2021 Summer Meeting, SLPC, Jun 2021, Paris, France. halshs-03983525

HAL Id: halshs-03983525

<https://shs.hal.science/halshs-03983525>

Submitted on 11 Feb 2023

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.

Naija grammatical classes and constructions

SYLVAIN KAHANE (MODYCO, UNIVERSITÉ PARIS NANTERRE)

BERNARD CARON (LLACAN, CNRS)

SLPC PARIS 2021, JUNE 29

Aim of the presentation

- ▶ Extract a corpus-based grammar of Naija
- ▶ Grammar rules contain quantitative information
 - ▶ this rule occurs in $x\%$ of the situations
- ▶ Methodology
 - ▶ The grammar is extracted from a treebank, that is, a corpus annotated with syntactic dependency trees

NaijaSynCor treebank

- ▶ 150 000 words annotated manually + 350 000 parsed
- ▶ The treebank contains:
 - ▶ POS (parts of speech)
 - ▶ morphosyntactic features (not very useful for creole)
 - ▶ syntactic relations
- ▶ The annotation is a starting point for a distributional study of the items
 - ▶ The analysis evolves during the annotation process, and annotating decisions can be changed

Requests on the treebank

Grew-match
developed by
Bruno Guillaume
(INRIA Nancy)

available online:
grew-match.fr

Example: distribution of POS

5

Hide corpora list

SUD_Naija-NSC@latest [9242 trees, 140836 tokens] last

update: 2021/06/25

Relation tables

Validation

```
1 % Search for a token of a given upos tag
2 % Available tags: ADJ, ADP, ADV, AUX, CONJ, DET, INTJ, NOUN, NUM, PART,
3
4 pattern { N [upos<>PUNCT] }
```

Clustering 1: No Key Whether

N.upos

lemma upos xpos features textform/wordform ? sentences order:

initial

context

Search

Save

Export

Example: distribution of POS

6

More than 1000 results found in 1.27% of the corpus [0.01s]

192 VERB

190 PRON

127 NOUN

114 AUX

72 DET

65 PART

56 ADP

48 ADV

43 INTJ

41 SCONJ

24 ADJ

10 PROPN

10 NUM

6 X

2 CCONJ

distribution of POS of subjects

7

Hide corpora list

SUD_Naija-NSC@latest [9242 trees, 140836 tokens] last update:

2021/06/25



Relation tables

Validation

```
1 search for a token of a given upos tag
2 available tags: ADJ, ADP, ADV, AUX, CONJ, DET, INTJ, NOUN, NUM, PART, PRON, PROPN, PUNCT, VERB
3
4 tern { N [upos<>PUNCT] ; H -[subj]-> N }
```

Clustering 1: No Key Whether

N.upos

lemma upos xpos features textform/wordform  sentences order: 

context

Search

Save

Export

distribution of POS of subjects

8

More than 1000 results found in 10.84% of the corpus [0.19s]

826 PRON

150 NOUN

9 PROP

5 ADJ

4 NUM

2 DET

2 X

1 ADV

1 VERB

More results

Metadata

CoNLL

SVG

e get one lady < hm she just enter inside shop o //

←

←

1 / 10

→

→

KAD_22_Chatting-At-The-Restaurant_DG_1

KAD_22_Chatting-At-The-Restaurant_DG_2 [1/2]

KAD_22_Chatting-At-The-Restaurant_DG_2 [2/2]

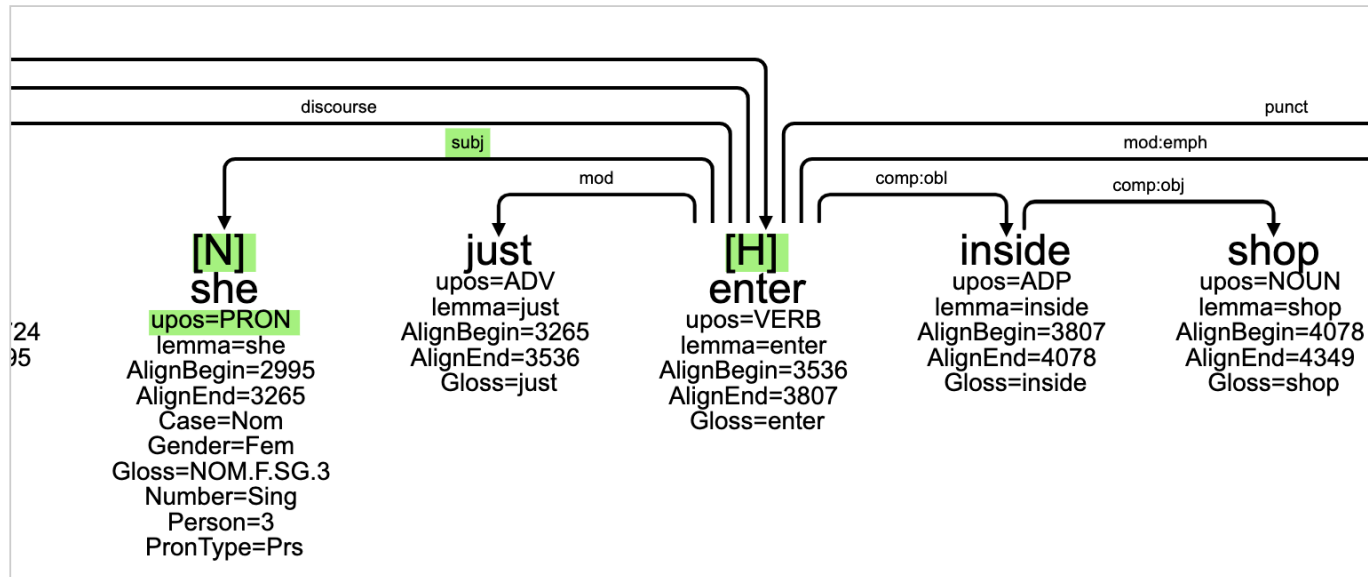
KAD_22_Chatting-At-The-Restaurant_DG_3 [1/2]

KAD_22_Chatting-At-The-Restaurant_DG_3 [2/2]

KAD_22_Chatting-At-The-Restaurant_DG_4

KAD_22_Chatting-At-The-Restaurant_DG_5

KAD_22_Chatting-At-The-Restaurant_DG_7



governors of subjects

```
11 pattern { N [upos<>PUNCT] ; H -[subj]-> N }
```

Clustering 1: No Key Whether

H.upos

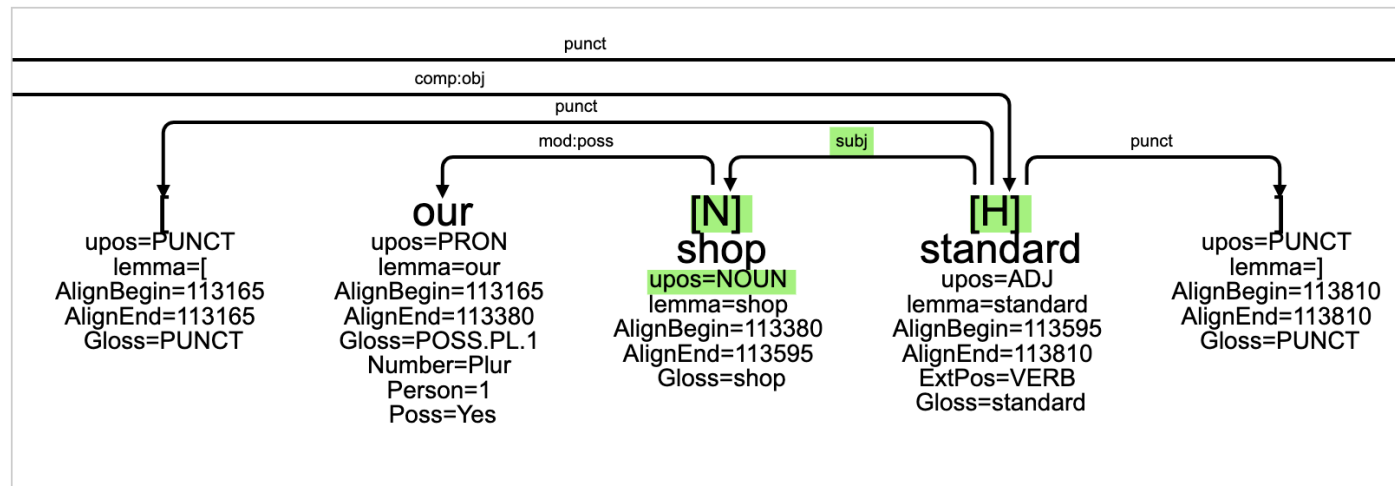
More than 1000 results found in 10.84% of the corpus [0.01s]

549 AUX 407 VERB **20 ADJ** 18 PART 4 X 2 ADP

- More results
- 1 / 10
- KAD_22_Chatting-At-The-Restaurant_DG_71 [2/2]
 - JOS_21_Marriage-Talk-With-Oscar-1_DG_7 [2/2]
 - JOS_21_Marriage-Talk-With-Oscar-1_DG_43
 - JOS_21_Marriage-Talk-With-Oscar-1_DG_122 [1/2]
 - JOS_10_Mothers-Against-Mini-Skirts_DG_22
 - ENU_17_Buying-Grocery_DG_9
 - ENU_17_Buying-Grocery_DG_12
 - ENU_17_Buying-Grocery_DG_40
 - ENU_17_Buying-Grocery_DG_41

Metadata > CoNLL SVG

you know sey [our shop standard] //



Focus on AUX and ADV

- ▶ Can we characterize auxiliaries and distinguish them from adverbs?
 - ▶ difficult question, because:
 - ▶ no inflection
 - ▶ same position (preverbal)

More than 1000 results found in 12.15% of the corpus [0.41s]

256 dey

237 go

189 no

89 con

73 don

66 make

30 fit

12 bin

10 never

9 be

7 will

5 can

5 do

4 gats

2 for

2 must

1 should

1 may

AUX

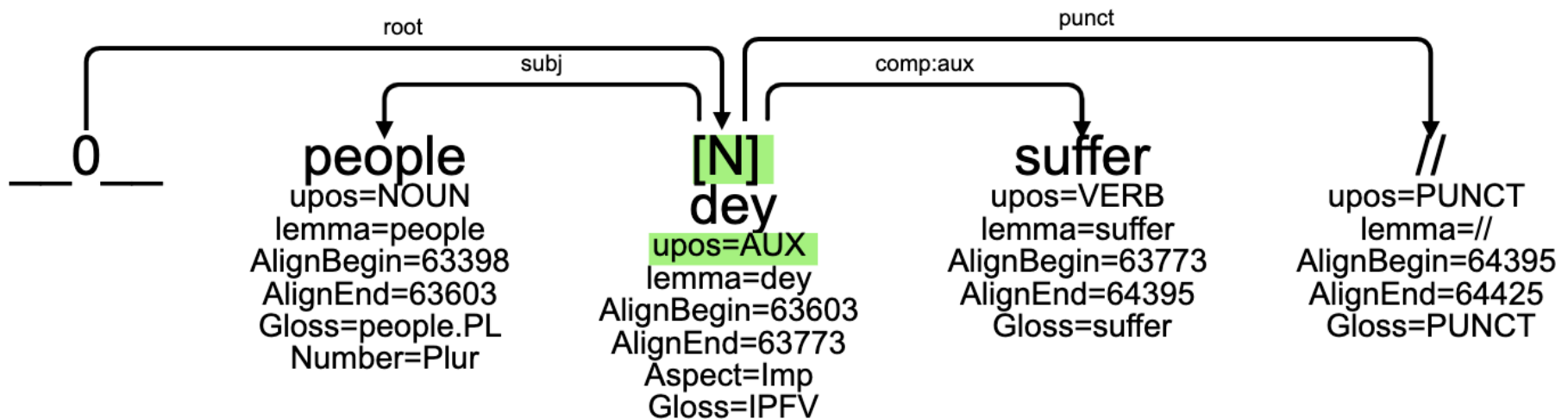
12

- ▶ dey = IMPERFECTIVE ≈ 'be Ving' 26%
- ▶ go = FUTUR ≈ 'will' 24%
- ▶ no = NEGATION ≈ 'don't' 19%
- ▶ con = CONSECUTIVE ≈ 'then' 9%
 - ▶ I con gast carry her go di spare room //
'Then I took her to the spare room.'
- ▶ don = PERFECTIVE ≈ 'have Ved' 7%
- ▶ make = JUSSIVE/HORTATIVE ≈ 'let X' 6%

AUX

- ▶ AUX is preverbal

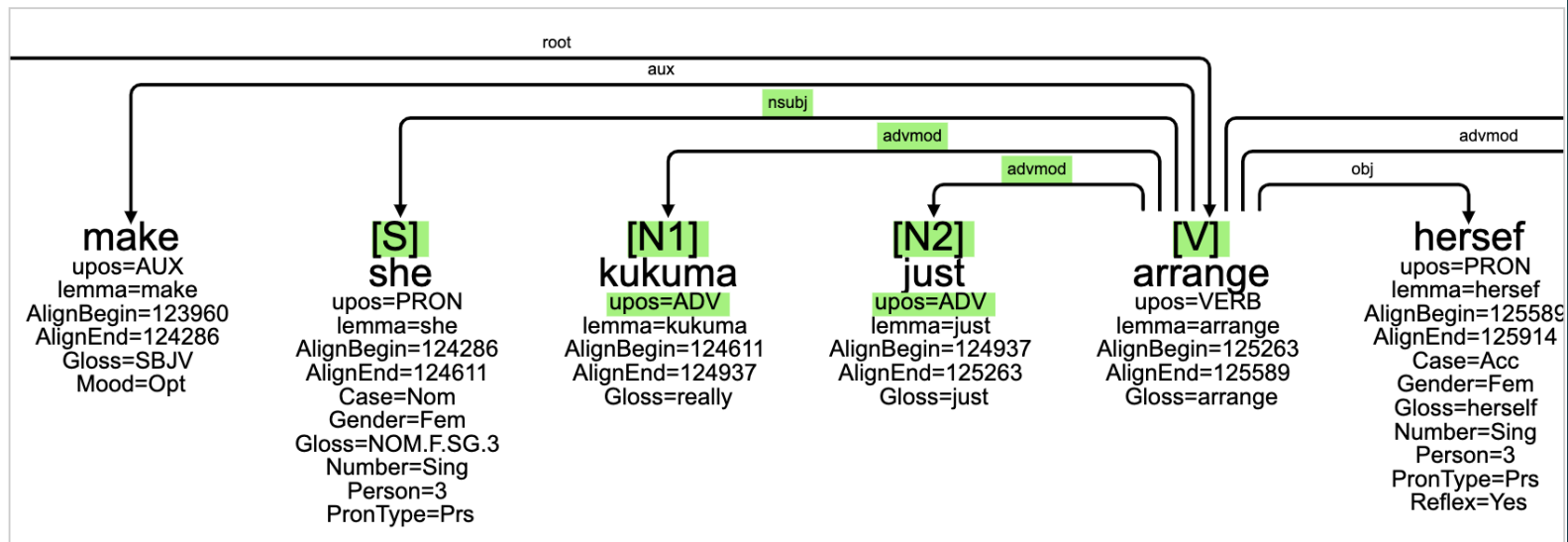
people **dey** suffer //



ADV

- ▶ 438 ADV between AUX and V
- ▶ 293 ADV between S and V
- ▶ 123 ADV between S and AUX

make she kukuma just arrange hersef o //



AUX piling vs ADV piling

- ▶ **Only 8 cases with two ADV**
- ▶ In contrast, it is common to have two AUX
 - ▶ 1300 cases with **two AUX**
+ 103 cases with three AUX
- ▶ and common to have an AUX and an ADV
 - ▶ 438 configurations **AUX ADV V**
 - ▶ 123 configurations **ADV AUX V**

Co-occurrence and order between AUX

► no < go < don < fit < con < dey

<	no	go	don	fit	con	dey
no	–	129	0	17	52	131
go	0	–	18	91	44	127
don	0	0	–	4	1	76
fit	6	0	0	–	4	3
con	0	1	0	1	–	98
dey	0	0	1	1	1	–

Focus on the Negation in Naija : *NO*

3 types of negation have been annotated in the corpus, operating 3 different levels, with 3 different POS

- ▶ Determinative: DET (NOUN Phrase)
- ▶ Negative Predicative Particle: PART (Non-verbal Sentence)
- ▶ Negative Auxiliary (VERB Phrase)

The word *no* is also used as an interjection: INTJ

Hide corpora list

SUD_Naija-NSC@latest [9242 trees, 140836 tokens] last update: 2021/06/25

```

1 % Search for a given lemma (lemmatization is not available for all languages)
2
3 pattern { N [lemma="no"] }
4

```

Relation tables

Clustering 1: No Key Whether

N.upos

 lemma upos xpos features textform/wordform context

sentences order: shuffle

Search

Save

Export

More than 1000 results found in 54.64% of the corpus

863 AUX

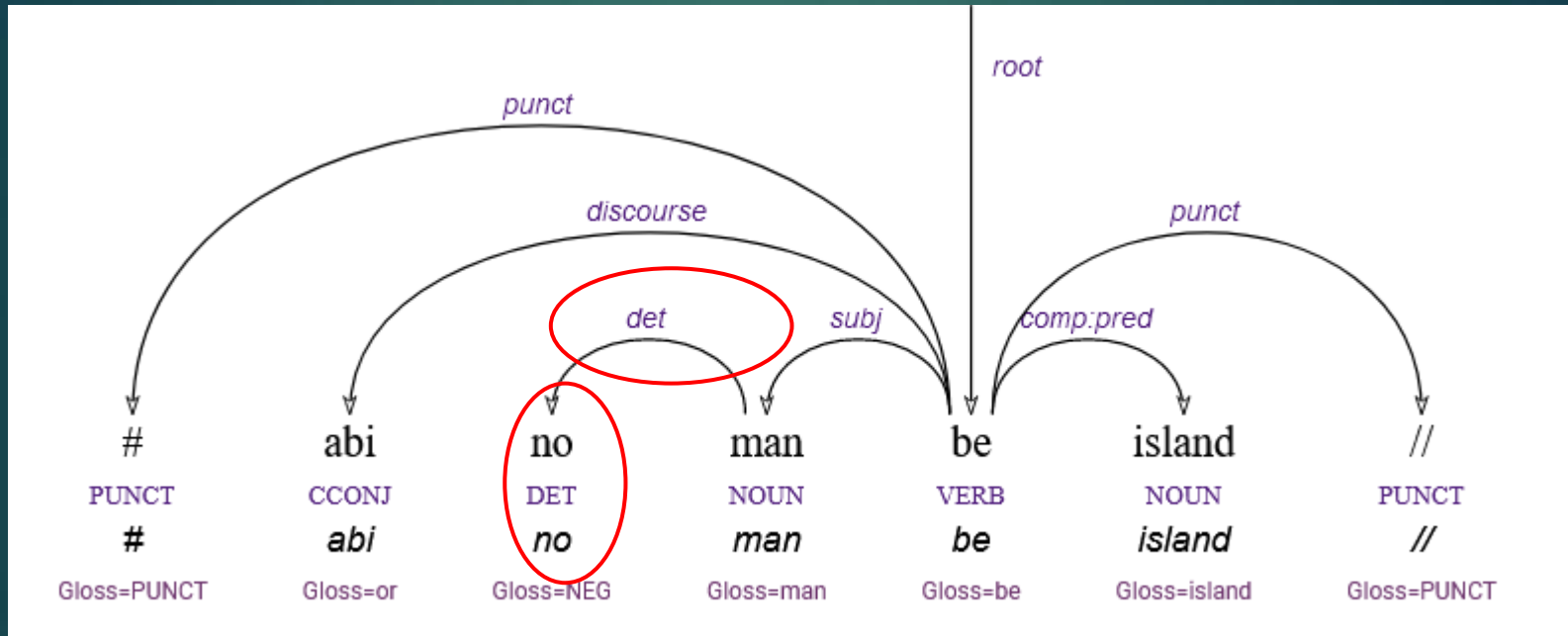
101 PART

25 INTJ

11 DET

AUX	1569	86,3%
PART	183	10,1%
INTJ	45	2,5%
DET	20	1,1%
TOTAL	1813	100%

No = DET



text_en I mean, no man is an island.

text # abi no man be island //

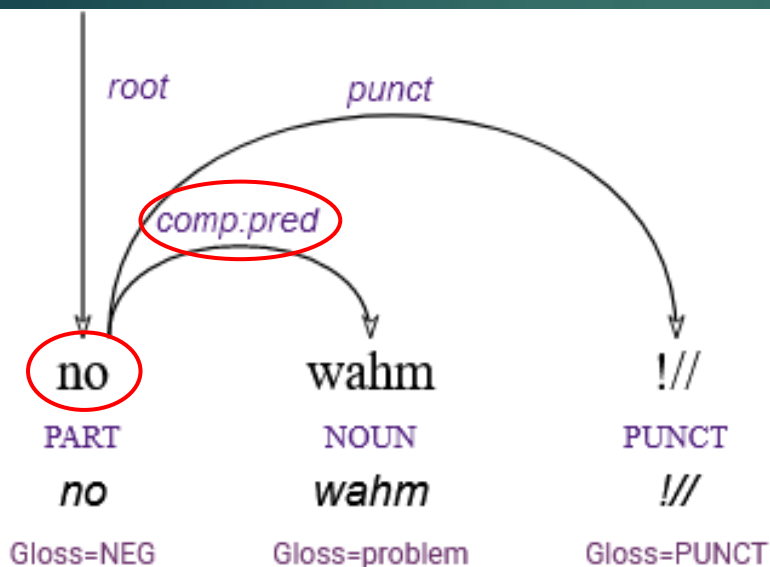
sent_id WAZA_05_Big-Mo_MG_63

```

1
2 pattern { N [lemma="no"] ; N [upos=DET] }

```

No = PART



text_en No problem!

text no wahn !!!

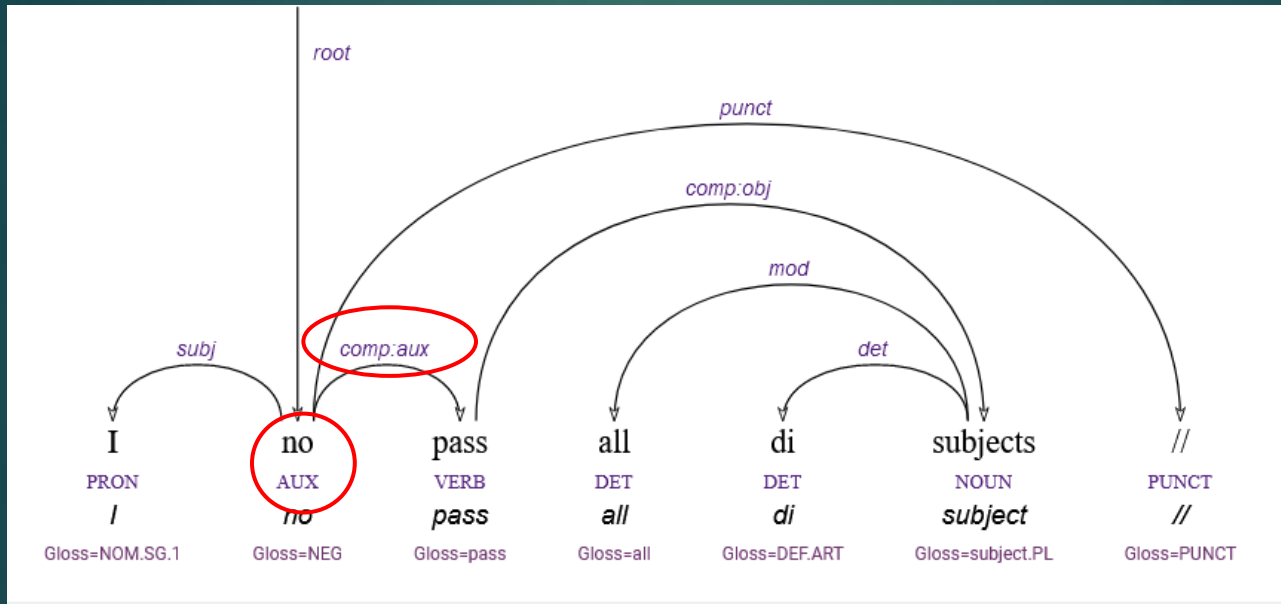
sent_id PRT_05_Ghetto-Life_MG_44

```

1
2 pattern { N [lemma="no"] ; N [upos=PART] }|

```

No= AUX:



`text_en` I didn't pass all the subjects.

`text` I no pass all di subjects //

`sent_id` WAZP_07_Imonirhua-Lifestory_MG_14

```
1  
2 pattern { N [lemma="no"] ; N [upos=AUX] }
```

Why is *NO* an *AUX* or a *PART* and not an *ADV*?

NO + NOUN

The case of noun predication

- ▶ If NO is treated as a DET in noun predication, two problems arise:

text_en No problem!

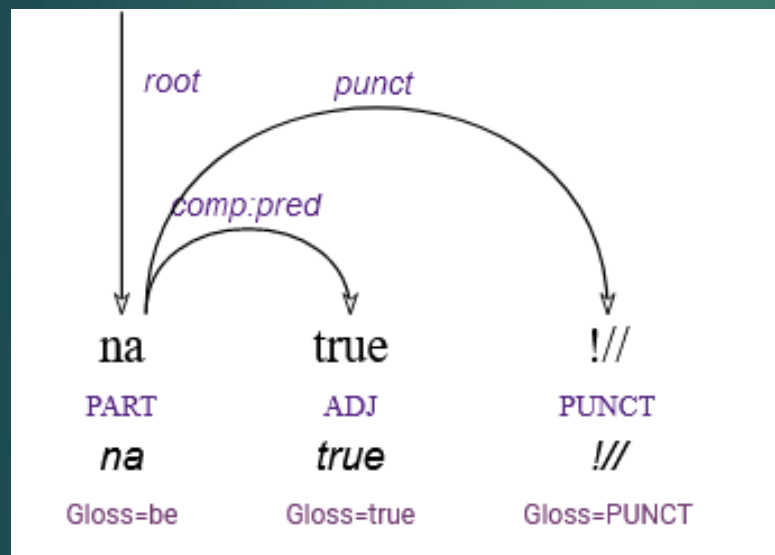
text no wahn !//

sent_id PRT_05_Ghetto-Life_MG_44

1. The noun *wahn* becomes the head of the sentence, and this is not possible outside short answers, insults, etc.
2. Negative determinatives of NOUNS functioning as arguments not common (23) compared to those in nominal predication (183)

Conclusion on non-verbal negative predication

- ▶ NO operates just like the particle *na* in *na true*, 'it's true!'



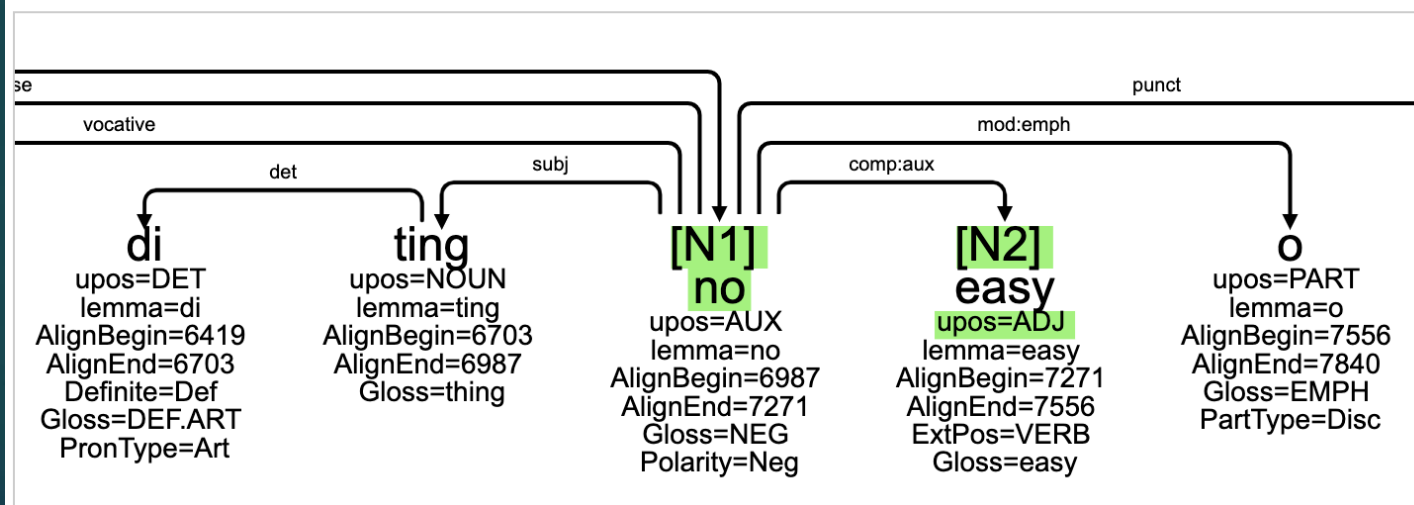
'It is'	na (PART) na true	no be (no = AUX; be = VERB) no be true
'there is'	dey (VERB) wahala dey	no (PART) no wahala

NO + ADJ

Adjectival negation

- ▶ NO co-occurs with ADJ only when ADJ is the predicate
 - ▶ 63 cases with ADJ predicate, 0 case with ADJ functioning as a modifier

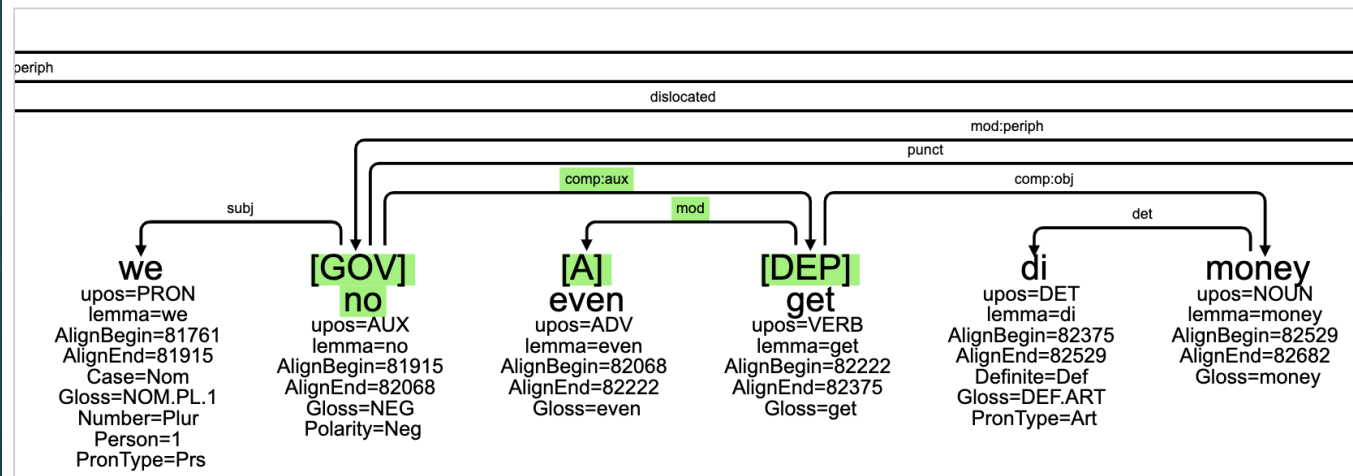
ehn auntie < di ting no easy o //



NO + ADV

- ▶ NO never modifies an ADV
- ▶ But NO can cooccur with an ADV
 - ▶ 78 configurations NO ADV V
 - ▶ to be compared with the only 8 cases of two ADV cooccurring in the preverbal position
 - ▶ This makes it very unlikely that NO is an ADV

as e be now sef < recession < we no even get di money < na im >> be di ting //



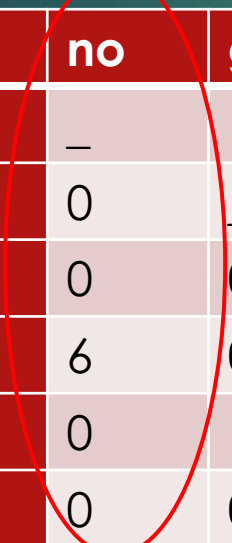
NO + VERB

Verbal negation

- ▶ NO appears in the same position as the TAM markers (=AUX), on the left of the verb
- ▶ Whereas adverbs move quite freely within the group of TAM markers, NO has a set place, just like other TAM markers, as shown by Sylvain

Distribution of NO in the AUX cluster

- ▶ NO only has 2 positions:
 - ▶ At the initial of the AUX group
 - ▶ In second position with a small list of AUX



<	no	go	don	fit	con	dey
no	–	129	0	17	52	131
go	0	–	18	91	44	127
don	0	0	–	4	1	76
fit	6	0	0	–	4	3
con	0	1	0	1	–	98
dey	0	0	1	1	1	–

Position of NO	3 AUX	2 AUX	1 AUX
1	2	56	494
2	1	17	80
3	0	1	–
4	0	–	–

- ▶ **Conclusion: The preferred position is at the initial**
- ▶ If not, the negation is **in second position** after a limited number of AUX: *make, fit, bin, for*. See the count for 1 AUX when filtered on the lemma of the AUX:

70 occurrences [0.11s]

59 make

6 fit

3 bin

2 for

[More results](#)
[Metadata](#)
[CoNLL](#)
[SVG](#)

man fit no forgive at all |c but God

WAZP_07_Imonirhua-Lifestory_MG_70

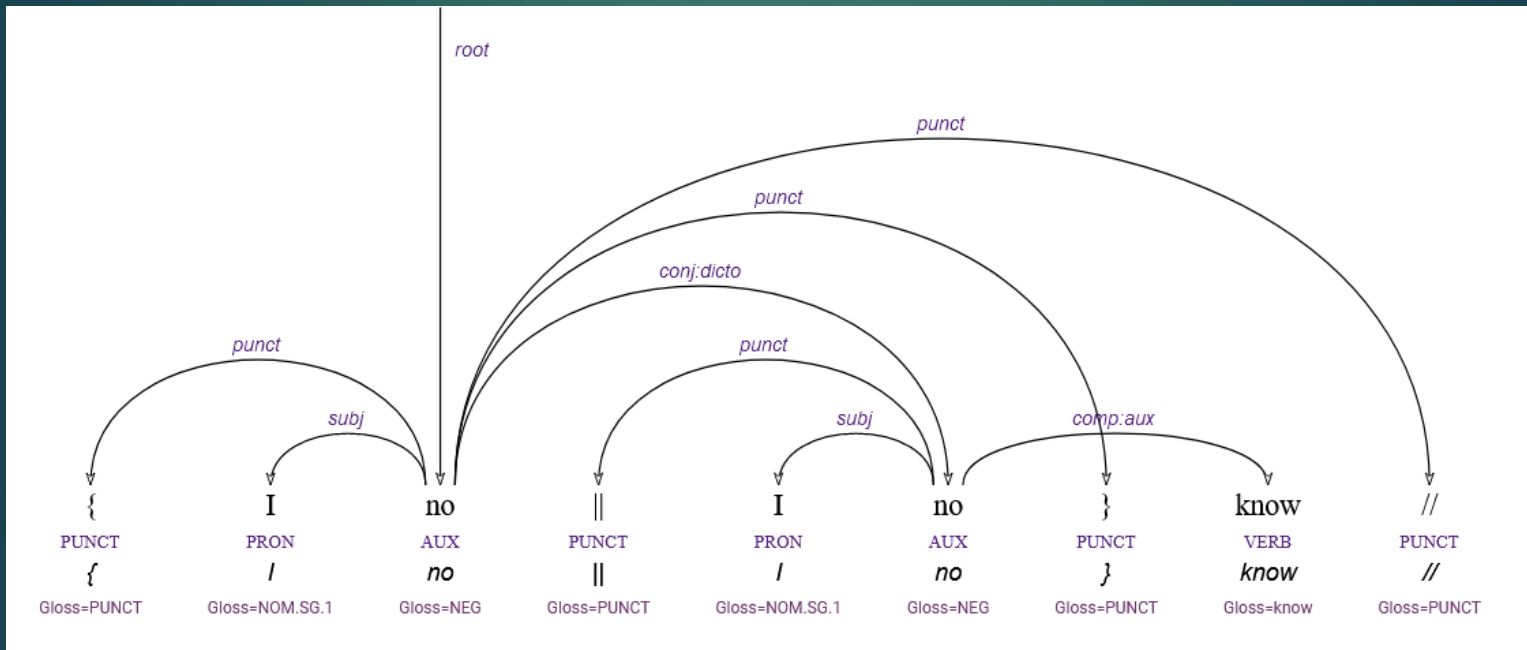
- ▶ One exception stands out: 1 NO that appears in 3rd position after 2 AUX:

Position of NO	3 AUX	2 AUX	1 AUX
1	2	56	494
2	1	17	80
3	0	1	–
4	0	–	–

- ▶ # *my people < na here >+ I go hold brake dis wonderful morning make e for no loss //*
- ▶ ‘My people, this is where I will call it a day this wonderful morning, so that it will not get spoilt.’
- ▶ This actually underlines the possibility of a grammaticalization of the AUX *make* to a purposive SCONJ.

NO and dysfluencies

- ▶ Treating NO as an AUX that can head a sentence explains clearly how certain dysfluencies work in Naija.

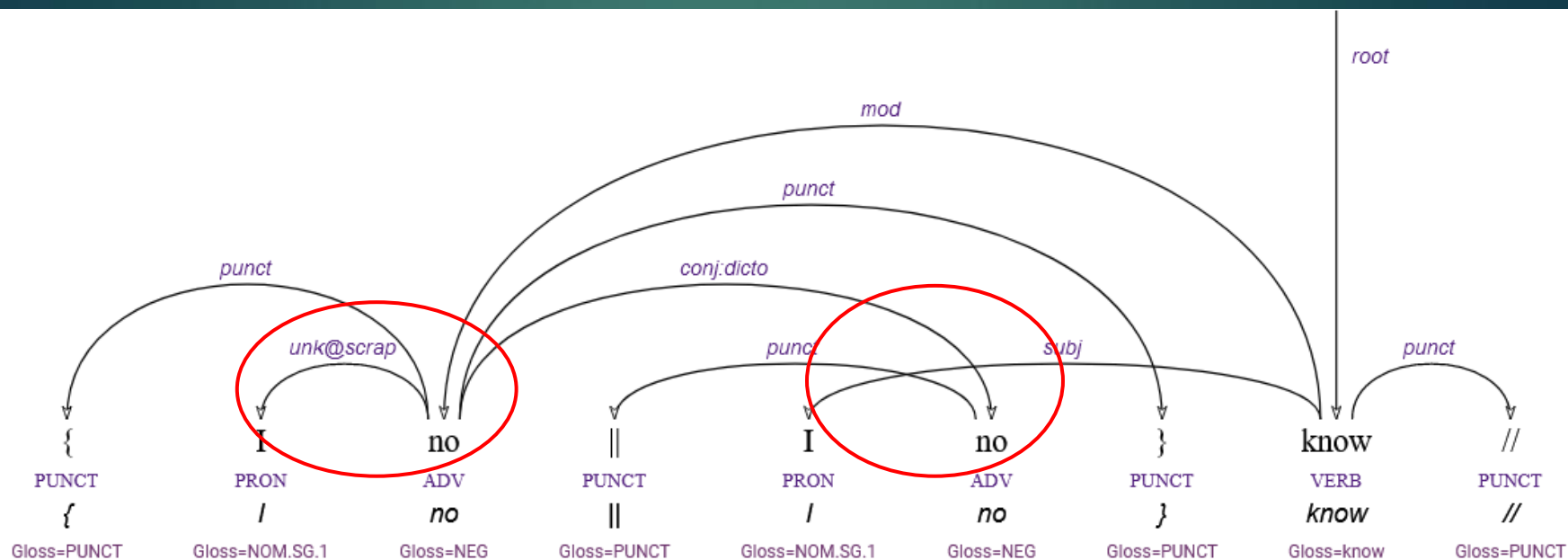


text_en I don't... I don't know.

text { I no || I no } know //

sent_id KAD_22_Chatting-At-The-Restaurant_DG__8

(If you don't, it gets very messy! Believe me!)



Conclusion

- ▶ (1) **Negation** tends to operate high in the syntactic tree
(ADJ and ADV are not negated as modifiers)
- ▶ (2) With **SUD treebanks + GREW**, it is possible to get quickly and rather easily a detailed account of the distribution of words, filter them on their syntactic function and thus pave the way for an extraction of **quantitative grammars**.