

Tarea 11 - NLP

April 20, 2021

Tarea 11

NLP

Descripción: Realizar el análisis sintáctico para un texto dado y visualizar el árbol sintáctico.

Alumno: Miguel Angel Soto Hernandez

```
[ ]: # Importacion de Stanza
!pip install deplacy stanza
import stanza
stanza.download("es")
nlp = stanza.Pipeline("es")

[23]: import re
from bs4 import BeautifulSoup
import urllib.request as urllib2
import requests
import deplacy
```

0.2 Obtener texto

```
[3]: def obtener_texto(tema):
    response = requests.get(f'https://es.wikipedia.org/wiki/{tema}')
    soup = BeautifulSoup(response.text)
    parrafos = soup.find_all('p')
    texto = parrafos[0].text
    return texto

[21]: def texto_a_frases(texto):
    documento_stanza = nlp(texto)
    frases = [frase.text for frase in documento_stanza.sentences]
    return frases

[14]: texto = obtener_texto('Lenguaje natural')
print(texto)
```

En la filosofía del lenguaje, el lenguaje natural es la lengua o idioma hablado o escrito por humanos para propósitos generales de comunicación. Son aquellas

lenguas que han sido generadas espontáneamente en un grupo de hablantes con propósito de comunicarse, a diferencia de otras lenguas, como puedan ser una lengua construida, los lenguajes de programación o los lenguajes formales usados en el estudio de la lógica formal, especialmente la lógica matemática.

```
[22]: frases = texto_a_frases(texto)
      print(frases)
```

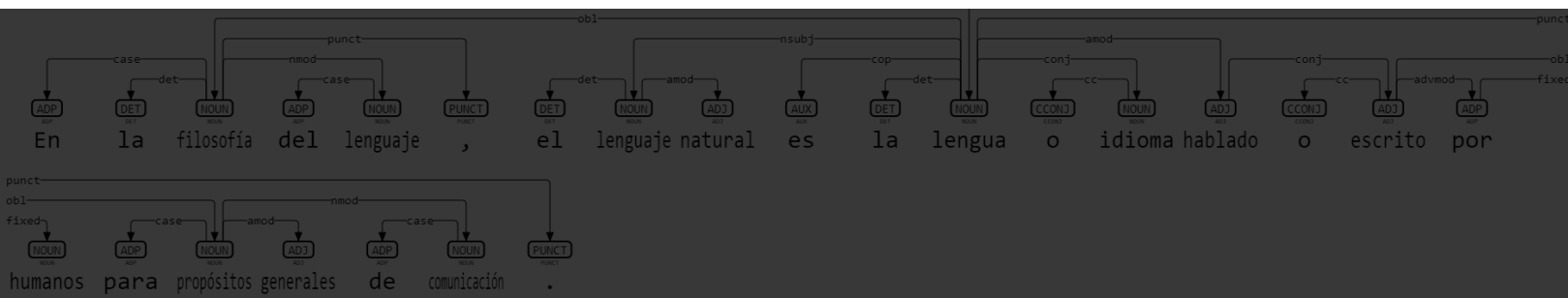
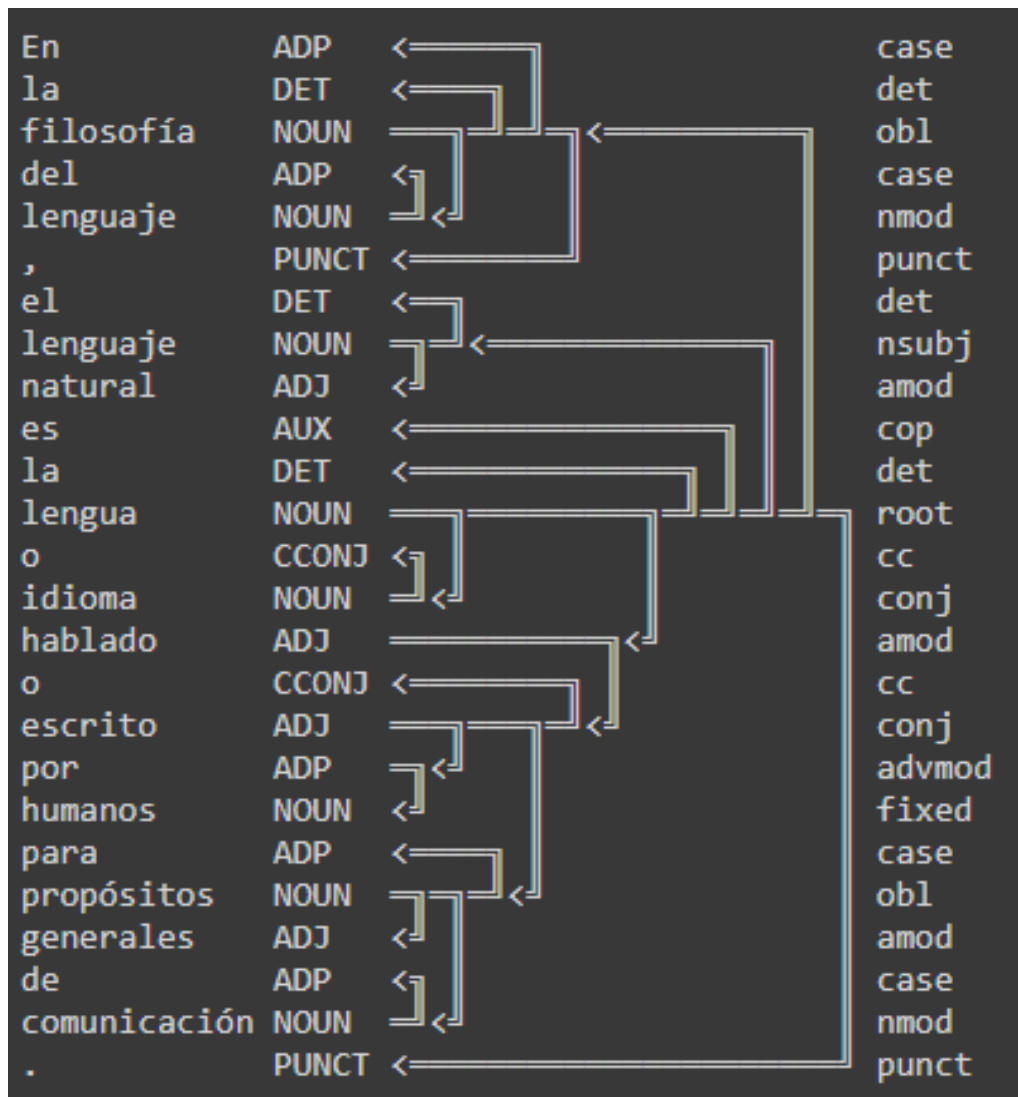
['En la filosofía del lenguaje, el lenguaje natural es la lengua o idioma hablado o escrito por humanos para propósitos generales de comunicación.', 'Son aquellas lenguas que han sido generadas espontáneamente en un grupo de hablantes con propósito de comunicarse, a diferencia de otras lenguas, como puedan ser una lengua construida, los lenguajes de programación o los lenguajes formales usados en el estudio de la lógica formal, especialmente la lógica matemática.']

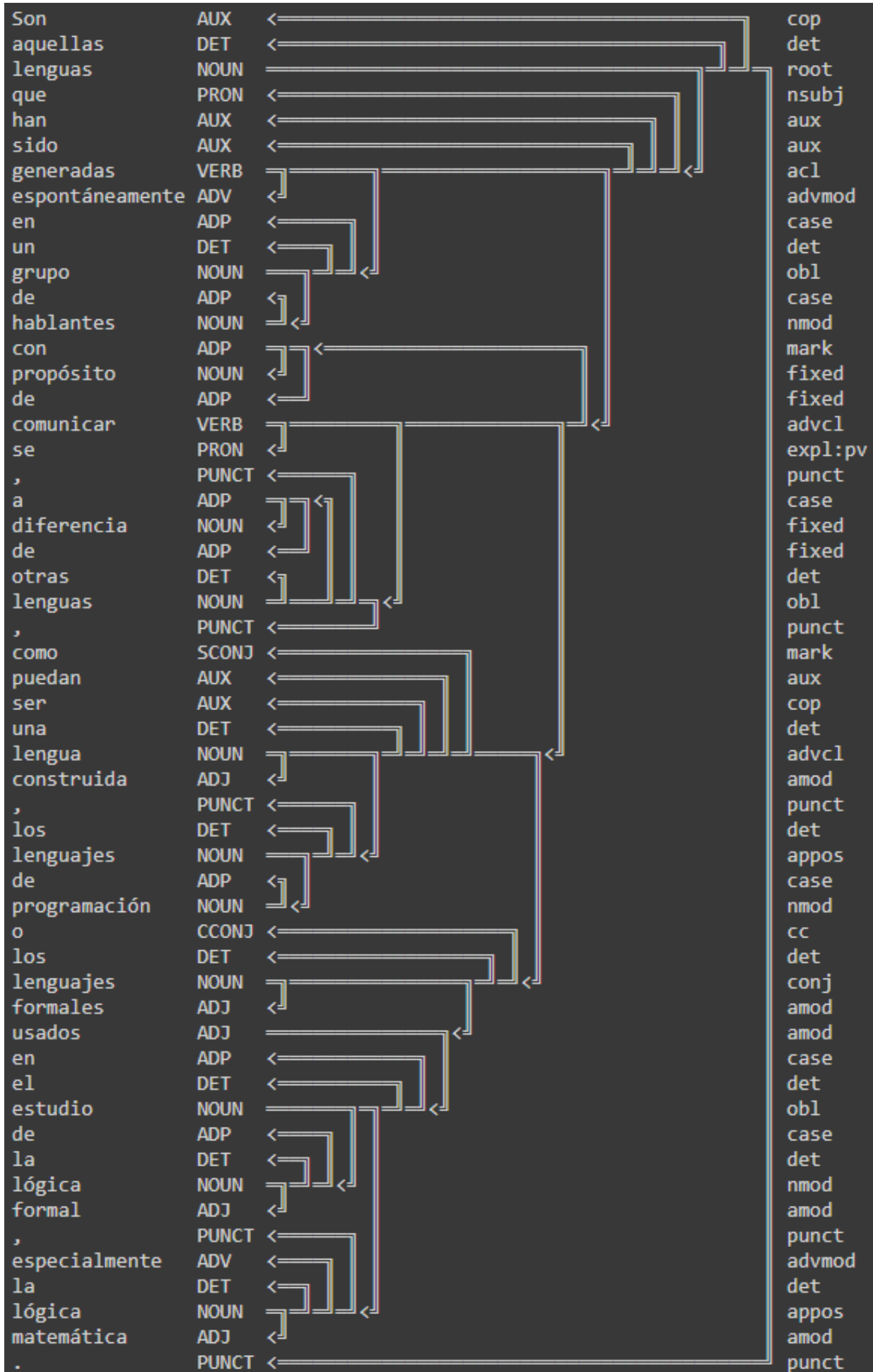
0.3 Análisis sintáctico

```
[26]: def analisis_sintactico(frase):
      documento_stanza = nlp(frase)
      deplacy.render(documento_stanza)
      deplacy.serve(documento_stanza, port=None)
```

```
[27]: for frase in frases:
      analisis_sintactico(frase)
```

en	ADP	ADP	AdpType=Prep	3	case	-	start_char=0 end_char=2
la	DET	DET	Definite=Def Gender=Fem Number=Sing PronType=Art	3	det	-	start_char=3 end_char=5
filosofía		filosofía	NOUN NOUN Gender=Fem Number=Sing	12	obl	-	start_char=6 end_char=15
del	ADP	ADP	AdpType=Preppron	5	case	-	start_char=16 end_char=19
lenguaje		lenguaje	NOUN NOUN Gender=Masc Number=Sing	3	nmod	-	start_char=20 end_char=28
,	PUNCT	PUNCT	PunctType=Comm	3	punct	-	start_char=28 end_char=29
el	DET	DET	Definite=Def Gender=Masc Number=Sing PronType=Art	8	det	-	start_char=30 end_char=32
lenguaje		lenguaje	NOUN NOUN Gender=Masc Number=Sing	12	nsubj	-	start_char=33 end_char=41
natural	ADJ	ADJ	Number=Sing	8	amod	-	start_char=42 end_char=49
es	AUX	AUX	Mood=Ind Number=Sing Person=3 Tense=Pres VerbForm=Fin	12	cop	-	start_char=50 end_char=52
la	DET	DET	Definite=Def Gender=Fem Number=Sing PronType=Art	12	det	-	start_char=53 end_char=55
lengua	NOUN	NOUN	Gender=Fem Number=Sing	0	root	-	start_char=56 end_char=62
o	CCONJ	CCONJ	14 cc	-	-	-	start_char=63 end_char=64
idioma	NOUN	NOUN	Gender=Masc Number=Sing	12	conj	-	start_char=65 end_char=71
hablado	ADJ	ADJ	Gender=Masc Number=Sing VerbForm=Part	12	amod	-	start_char=72 end_char=79
o	CCONJ	CCONJ	17 cc	-	-	-	start_char=80 end_char=81
escrito	ADJ	ADJ	Gender=Masc Number=Sing VerbForm=Part	15	conj	-	start_char=82 end_char=89
por	ADP	ADP	AdpType=Prep	17	advmod	-	start_char=90 end_char=93
humanos	NOUN	NOUN	18 fixed	-	-	-	start_char=94 end_char=101
para	ADP	ADP	AdpType=Prep	21	case	-	start_char=102 end_char=106
propósitos	NOUN	propósito	NOUN NOUN Gender=Masc Number=Plur	17	obl	-	start_char=107 end_char=117
generales	ADJ	general	ADJ Number=Plur	21	amod	-	start_char=118 end_char=127
de	ADP	ADP	AdpType=Prep	24	case	-	start_char=128 end_char=130
comunicación		comunicación	NOUN NOUN Gender=Fem Number=Sing	21	nmod	-	start_char=131 end_char=143
.	PUNCT	PUNCT	PunctType=Peri	12	punct	-	start_char=143 end char=144





1	Son	ser	AUX	AUX	Mood=Ind Number=Plur Person=3 Tense=Pres VerbForm=Fin	3	cop	-
<u>start_char=0 end_char=3</u>								
2	aquellas	aquel	DET	DET	Gender=Fem Number=Plur PronType=Dem	3	det	<u>start_char=4 end_char=12</u>
3	lenguas	lengua	NOUN	NOUN	Gender=Fem Number=Plur	0	root	<u>start_char=13 end_char=20</u>
4	que	que	PRON	PRON	PronType=Int,Rel	7	nsubj	<u>start_char=21 end_char=24</u>
5	han	haber	AUX	AUX	Mood=Ind Number=Plur Person=3 Tense=Pres VerbForm=Fin	7	aux	-
<u>start_char=25 end_char=28</u>								
6	sido	ser	AUX	AUX	Gender=Masc Number=Sing Tense=Past VerbForm=Part	7	aux	-
<u>start_char=29 end_char=33</u>								
7	generadas	generado	VERB	VERB	Gender=Fem Number=Plur Tense=Past VerbForm=Part	3	acl	-
<u>start_char=34 end_char=43</u>								
8	espontáneamente	espontáneamente	ADV	ADV		7	advmod	<u>start_char=44 end_char=59</u>
9	en	en	ADP	ADP	AdpType=Prep	11	case	<u>start_char=60 end_char=62</u>
10	un	uno	DET	DET	Definite=Ind Gender=Masc Number=Sing PronType=Art	11	det	-
<u>start_char=63 end_char=65</u>								
11	grupo	grupo	NOUN	NOUN	Gender=Masc Number=Sing	7	obl	<u>start_char=66 end_char=71</u>
12	de	de	ADP	ADP	AdpType=Prep	13	case	<u>start_char=72 end_char=74</u>
13	hablantes	hablante	NOUN	NOUN	Number=Plur	11	nmod	<u>start_char=75 end_char=84</u>
14	con	con	ADP	ADP	AdpType=Prep	17	mark	<u>start_char=85 end_char=88</u>
15	propósito	propósito	NOUN	NOUN		14	fixed	<u>start_char=89 end_char=98</u>
16	de	de	ADP	ADP	AdpType=Prep	14	fixed	<u>start_char=99 end_char=101</u>
17-18	comunicarse							<u>start_char=102 end_char=113</u>
17	comunicar	comunicar	VERB	VERB	VerbForm=Inf	7	advcl	-
18	se	el	PRON	PRON	Case=Acc Person=3 PrepCase=Npr PronType=Prs Reflex=Yes	17	expl:pv	-
19	,	,	PUNCT	PUNCT	PunctType=Comm	24	punct	<u>start_char=113 end_char=114</u>
20	diferencia	diferencia	NOUN	NOUN	AdpType=Prep	24	case	<u>start_char=115 end_char=116</u>
21	de	de	ADP	ADP		20	fixed	<u>start_char=117 end_char=127</u>
22	de	de	ADP	ADP	AdpType=Prep	20	fixed	<u>start_char=128 end_char=130</u>
23	otras	otro	DET	DET	Gender=Fem Number=Plur PronType=Ind	24	det	<u>start_char=131 end_char=136</u>
24	lenguas	lengua	NOUN	NOUN	Gender=Fem Number=Plur	17	obl	<u>start_char=137 end_char=144</u>
25	,	,	PUNCT	PUNCT	PunctType=Comm	24	punct	<u>start_char=144 end_char=145</u>
26	como	como	SCONJ	SCONJ		30	mark	<u>start_char=146 end_char=150</u>
27	puedan	poder	AUX	AUX	Mood=Sub Number=Plur Person=3 Tense=Pres VerbForm=Fin	30	aux	-
<u>start_char=151 end_char=157</u>								
28	ser	ser	AUX	AUX	VerbForm=Inf	30	cop	<u>start_char=158 end_char=161</u>
29	una	uno	DET	DET	Definite=Ind Gender=Fem Number=Sing PronType=Art	30	det	-
<u>start_char=162 end_char=165</u>								
30	lengua	lengua	NOUN	NOUN	Gender=Fem Number=Sing	17	advcl	<u>start_char=166 end_char=172</u>
31	construida	construido	ADJ	ADJ	Gender=Fem Number=Sing VerbForm=Part	30	amod	-
<u>start_char=173 end_char=183</u>								
32	,	,	PUNCT	PUNCT	PunctType=Comm	34	punct	<u>start_char=183 end_char=184</u>
33	los	el	DET	DET	Definite=Def Gender=Masc Number=Plur PronType=Art	34	det	-
<u>start_char=185 end_char=188</u>								
34	lenguajes	lenguaje	NOUN	NOUN	Gender=Masc Number=Plur	30	appos	<u>start_char=189 end_char=198</u>
35	de	de	ADP	ADP	AdpType=Prep	36	case	<u>start_char=199 end_char=201</u>
36	programación	programación	NOUN	NOUN	Gender=Fem Number=Sing	34	nmod	<u>start_char=202 end_char=214</u>
37	o	o	CCONJ	CCONJ		39	cc	<u>start_char=215 end_char=216</u>
38	los	el	DET	DET	Definite=Def Gender=Masc Number=Plur PronType=Art	39	det	-
<u>start_char=217 end_char=220</u>								
39	lenguajes	lenguaje	NOUN	NOUN	Gender=Masc Number=Plur	30	conj	<u>start_char=221 end_char=230</u>
40	formales	formal	ADJ	ADJ	Number=Plur	39	amod	<u>start_char=231 end_char=239</u>
41	usados	usado	ADJ	ADJ	Gender=Masc Number=Plur VerbForm=Part	39	amod	<u>start_char=240 end_char=246</u>
42	en	en	ADP	ADP	AdpType=Prep	44	case	<u>start_char=247 end_char=249</u>
43	el	el	DET	DET	Definite=Def Gender=Masc Number=Sing PronType=Art	44	det	-
<u>start_char=250 end_char=252</u>								
44	estudio	estudio	NOUN	NOUN	Gender=Masc Number=Sing	41	obl	<u>start_char=253 end_char=260</u>
45	de	de	ADP	ADP	AdpType=Prep	47	case	<u>start_char=261 end_char=263</u>
46	la	el	DET	DET	Definite=Def Gender=Fem Number=Sing PronType=Art	47	det	-
<u>start_char=264 end_char=266</u>								
47	lógica	lógica	NOUN	NOUN	Gender=Fem Number=Sing	44	nmod	<u>start_char=267 end_char=273</u>
48	formal	formal	ADJ	ADJ	Number=Sing	47	amod	<u>start_char=274 end_char=280</u>
49	,	,	PUNCT	PUNCT	PunctType=Comm	52	punct	<u>start_char=280 end_char=281</u>
50	especialmente	especialmente	ADV	ADV		52	advmod	<u>start_char=282 end_char=295</u>
51	la	el	DET	DET	Definite=Def Gender=Fem Number=Sing PronType=Art	52	det	-
<u>start_char=296 end_char=298</u>								
52	lógica	lógica	NOUN	NOUN	Gender=Fem Number=Sing	44	appos	<u>start_char=299 end_char=305</u>
53	matemática	matemático	ADJ	ADJ	Gender=Fem Number=Sing	52	amod	<u>start_char=306 end_char=316</u>
54	.	.	PUNCT	PUNCT	PunctType=Peri	3	punct	<u>start_char=316 end_char=317</u>

