

Adding Semantics to UD markers

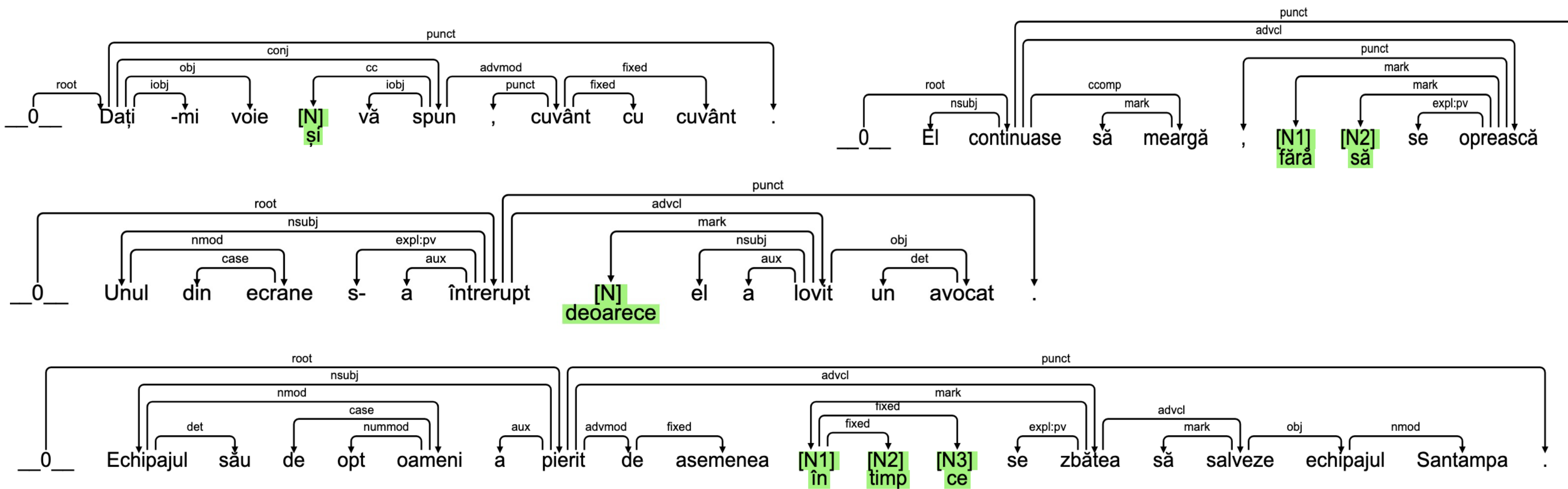
Tudor Voicu¹, Verginica Barbu Mititelu²

¹Tudor Vianu National High School of Computer Science

²Romanian Academy Research Institute for Artificial Intelligence

Conjunctive locutions in UD

Word combinations functioning as a (coordinating or subordinating) conjunction.



The corpus: RRT

Results

- 9,523 sentences; 218511 tokens
- genres: legal, news, fiction, medical, non-fiction, academic, FrameNet translations and Wikipedia (unbalanced distribution)
- UD morpho-syntactic annotation
- CONLL-U format

Conjunction	Frequency	Annotation
<i>în timp ce</i>	29	TEMPORAL:SYNCHRONOUS
'while'	18	COMPARISON:CONTRAST
	2	COMPARISON:CONCESSION:ARG2-AS-DENIER
	1	COMPARISON:SIMILARITY
	1	insufficient context
<i>pentru că</i>	35	CONTINGENCY:CAUSE:REASON
'because'	10	CONTINGENCY:CAUSE+BELIEF:REASON+BELIEF
	4	CONTINGENCY:CAUSE+SPEECHACT:REASON+SPEECHACT
	3	CONTINGENCY:CAUSE+SPEECHACT:RESULT+SPEECHACT
	2	CONTINGENCY:CAUSE:RESULT
<i>înainte de</i>	35	TEMPORAL:ASYNCHRONOUS:PRECEDENCE
'before'		
<i>după ce</i>	48	TEMPORAL:ASYNCHRONOUS:SUCCESSION
'after'	38	TEMPORAL:ASYNCHRONOUS:SUCCESSION
	1	CONTINGENCY:CAUSE:REASON
		TEMPORAL:ASYNCHRONOUS:PRECEDENCE
		TEMPORAL:ASYNCHRONOUS:SUCCESSION
<i>pentru ca să</i>	30	CONTINGENCY:PURPOSE:ARG2-AS-GOAL
'so as to'	1	CONTINGENCY:CAUSE.NEGRESULT
	1	TEMPORAL:ASYNCHRONOUS:SUCCESSION
	1	CONTINGENCY:CONDITION:ARG1-AS-COND?
<i>astfel încât</i>	20	CONTINGENCY:PURPOSE:ARG2-AS-GOAL
'so that'	7	CONTINGENCY:CAUSE:RESULT

Relations

Level-1	Level-2	Level-3
TEMPORAL	SYNCHRONOUS	-
	ASYNCHRONOUS	PRECEDENCE SUCCESSION
CONTINGENCY	CAUSE	REASON
		RESULT
		NEGRESULT
	CAUSE+BELIEF	REASON+BELIEF
		RESULT+BELIEF
	CAUSE+SPEECHACT	REASON+SPEECHACT
		RESULT+SPEECHACT
	CONDITION	ARG1-AS-COND
		ARG2-AS-COND
CONDITION+SPEECHACT	-	
NEGATIVE-CONDITION	ARG1-AS-NEGCOND	
	ARG2-AS-NEGCOND	
NEGATIVE-CONDITION+SPEECHACT	-	
PURPOSE	ARG1-AS-GOAL	
	ARG2-AS-GOAL	
COMPARISON	CONCESSION	ARG1-AS-DENIER
		ARG2-AS-DENIER
	CONCESSION+SPEECHACT	ARG2-AS-DENIER+SPEECHACT
	CONTRAST	-
SIMILARITY	-	
EXPANSION	CONJUNCTION	-
	DISJUNCTION	-
	EQUIVALENCE	-
	EXCEPTION	ARG1-AS-EXCPT
		ARG2-AS-EXCPT
	INSTANTIATION	ARG1-AS-INSTANCE
		ARG2-AS-INSTANCE
	LEVEL-OF-DETAIL	ARG1-AS-DETAIL
		ARG2-AS-DETAIL
	MANNER	ARG1-AS-MANNER
ARG2-AS-MANNER		
SUBSTITUTION	ARG1-AS-SUBST	
	ARG2-AS-SUBST	

Representation

9	,	,	PUNCT	COMMA	-	10	punct	-	-
10	de	de	ADP	Spsa	AdpType=Prep(Case=Acc	7	advmod	-	-
11	asemenea	asemenea	ADJ	Afp	Degree=Pos	10	fixed	-	-
12	-	-	PUNCT	DASH	-	22	punct	-	-
13	de	de	ADP	Spsa	AdpType=Prep(Case=Acc	22	advmod	-	-
14	obicei	obicei	NOUN	Ncms-n	Definite=Ind(Gender=Masc Number=Sing	13	fixed	-	-
15	pentru	pentru	ADP	Spsa	AdpType=Prep(Case=Acc	22	mark	1	CONTINGENCY:CAUSE:REASON
16	că	că	SCONJ	Ccssp	Polarity=Pos	15	fixed	-	-
17	gunoiul	gunoi	NOUN	Ncmsry	Case=Acc,Nom Definite=Def(Gender=Masc Number=Sing	22	nsubj	-	-
18	sau	sau	CCONJ	Ccssp	Polarity=Pos	19	cc	-	-
19	piețișul	piețiș	NOUN	Ncmsry	Case=Acc,Nom Definite=Def(Gender=Masc Number=Sing	17	conj	-	-
20	s-	sine	PRON	Px3-a-y-----w	Case=Acc Person=3 PronType=Prs Reflex=Yes Strength=Weak Variant=Short	22	expl:pv	-	SpaceAfter=No
21	au	avea	AUX	Va-3p	Number=Plur Person=3	22	aux	-	-
22	adunat	aduna	VERB	Vmp-sm	Gender=Masc Number=Sing VerbForm=Part	7	advcl	-	-
23	între	între	ADP	Spsa	AdpType=Prep(Case=Acc	24	case	-	-
24	șanț	șanț	NOUN	Ncms-n	Definite=Ind(Gender=Masc Number=Sing	22	obl	-	-
25	și	și	CCONJ	Ccssp	Polarity=Pos	26	cc	-	-
26	sigiliu	sigiliu	NOUN	Ncms-n	Definite=Ind(Gender=Masc Number=Sing	24	conj	-	-
27	sau	sau	CCONJ	Ccssp	Polarity=Pos	34	cc	-	-
28	pentru	pentru	ADP	Spsa	AdpType=Prep(Case=Acc	34	mark	2	CONTINGENCY:CAUSE:REASON
29	că	că	SCONJ	Ccssp	Polarity=Pos	28	fixed	-	-
30	însuși	însuși	DET	Dh3ms	Gender=Masc Number=Sing Person=3 PronType=Emp	31	det	-	-
31	sigiliul	sigiliu	NOUN	Ncmsry	Case=Acc,Nom Definite=Def(Gender=Masc Number=Sing	34	nsubj	-	-
32	s-	sine	PRON	Px3-a-y-----w	Case=Acc Person=3 PronType=Prs Reflex=Yes Strength=Weak Variant=Short	34	expl:pv	-	SpaceAfter=No