

PENN ARABIC TREEBANK GUIDELINES

VERSION 4.95
JUNE 30, 2011

Mohamed Maamouri, Ann Bies, Sondos Krouna,
Dalila Tabessi, Fatma Gaddeche

With much appreciated contributions from
Seth Kulick, Wigdan Mekki, Tim Buckwalter

Linguistic Data Consortium
University of Pennsylvania
3600 Market Street, Suite 810
Philadelphia, PA 19104
bies@ldc.upenn.edu, maamouri@ldc.upenn.edu

Table of Contents

1	Introduction.....	10
1.1	A note about glosses	10
1.2	A Note on Case Marking	10
1.2.1	Diacritization: Vocalized and Unvocalized text /Case and Mood endings.....	10
1.3	A partial glossary of Arabic grammar terminology	11
1.4	A note about punctuation in Arabic texts	11
2	Basic Arabic clause structure.....	11
2.1	Basic sentence structure.....	12
2.1.1	The form of Arabic sentences	12
2.1.2	Annotating basic elements in Arabic sentences.....	15
2.2	Making corrections to morphological/POS tags	16
2.3	Node labels and functional "dashtags"	17
2.3.1	Node labels.....	17
2.3.1.1	Types of clauses	18
2.3.1.1.1	S.....	18
2.3.1.1.2	SBAR	21
2.3.1.1.3	SQ.....	23
2.3.1.1.4	SBARQ.....	25
2.3.1.1.5	FRAG	27
2.3.1.1.5.1	Not FRAG	31
2.3.1.1.6	Node labels not used in ATB	33
2.3.1.2	Types of phrases smaller than clauses	33
2.3.1.2.1	NP.....	33
2.3.1.2.2	VP.....	33
2.3.1.2.3	PP	34
2.3.1.2.3.1	List of true prepositions in the Arabic Treebank.....	36
2.3.1.2.3.2	Specific preposition constructions	41
2.3.1.2.3.3	Some lexical items can be either preposition or particle depending on context	42
2.3.1.2.3.4	List of nouns formerly treated as prepositions in the Arabic Treebank	43
2.3.1.2.4	ADVP	43
2.3.1.2.4.1	List of true adverbs in the Arabic Treebank.....	44
2.3.1.2.5	ADJP	48
2.3.1.2.6	PRT.....	49
2.3.1.2.6.1	Particles annotated under constituents smaller than a sentence	49
2.3.1.2.6.1.1	Vocative particles (VOC_PART).....	49
2.3.1.2.6.1.2	Emphatic particle (EMPHATIC_PART)	50
2.3.1.2.6.1.3	Negative particles (NEG_PART).....	50
2.3.1.2.6.1.4	Future Particle (FUT_PART) and Verb Particles (VERB_PART)	51
2.3.1.2.6.1.5	Other Particles	52
2.3.1.2.6.2	Particles annotated at the sentence level	53
2.3.1.2.6.2.1	Interrogative Particles (INTERROG_PART)	53
2.3.1.2.6.2.2	Focus Particle (FOCUS_PART)	53
2.3.1.2.6.2.3	Other Particles	54

2.3.1.2.7	WHNP.....	55
2.3.1.2.8	WHADVP.....	56
2.3.1.2.9	WHPP.....	59
2.3.1.2.10	CONJP.....	59
2.3.1.3	Other types.....	60
2.3.1.3.1	PRN.....	60
2.3.1.3.1.1	Parentheticals where PRN should be used.....	61
2.3.1.3.1.1.1	Sentential interpolations:.....	62
2.3.1.3.1.1.2	PRN because there is no other solution, as a last resort (train wrecks) 71	
2.3.1.3.1.2	Not PRN anymore.....	72
2.3.1.3.2	NAC (Not A Constituent).....	75
2.3.1.3.2.1	NAC emphasizing modification.....	76
2.3.1.3.2.2	NAC vs. UCP.....	82
2.3.1.3.2.3	NAC vs. Gapping.....	83
2.3.1.3.2.4	NAC vs. prepositional “wa” (waw AlmaEiy~ap).....	84
2.3.1.3.2.5	NAC vs. particle “wa”.....	85
2.3.1.3.3	UCP.....	85
2.3.1.3.4	LST.....	85
2.3.1.3.5	X.....	86
2.3.2	Dashtags! How to use them. How to decide... ..	88
2.3.2.1	–SBJ.....	88
2.3.2.2	–OBJ.....	89
2.3.2.3	–DTV.....	91
2.3.2.4	–CLR.....	93
2.3.2.4.1	NP-CLR.....	95
2.3.2.5	–NOM.....	96
2.3.2.6	–PRD.....	97
2.3.2.7	–LOC.....	99
2.3.2.8	–DIR.....	101
2.3.2.9	–MNR.....	102
2.3.2.10	–TMP.....	103
2.3.2.11	–PRP.....	105
2.3.2.12	–ADV.....	105
2.3.2.13	–HLN.....	107
2.3.2.14	–TTL.....	107
2.3.2.15	–VOC.....	109
2.3.2.16	–TPC.....	110
2.3.2.17	Dashtags not used in ATB.....	113
2.3.3	Dashtag decisions:.....	113
2.4	VP arguments and adjuncts.....	115
2.5	NP arguments and adjuncts.....	116
2.6	Clause combinations.....	117
2.6.1	Coordination.....	117
2.6.1.1	Coordination with a conjunction.....	117
2.6.1.1.1	Coordination with only “>amo 1A” (<i>or not</i>).....	122

2.6.1.2	Coordination without a conjunction.....	122
2.6.1.2.1	When sentences are NOT coordinated.....	123
2.6.1.3	Initial conjunction.....	126
2.6.2	Subordination: The use of SBAR.....	127
2.6.3	Gapping (VP Template Gapping).....	129
2.6.4	Separate sentences.....	130
2.6.5	S-ADV versus SBAR zero relative clause.....	131
2.6.6	SBAR-ADV versus separate S's.....	132
2.7	Empty categories.....	133
2.7.1	Discontinuous Constituents/Rightward Movement.....	134
2.7.2	Ellipsis.....	139
2.8	Clitics and Tokenization.....	139
2.8.1	Clitics.....	139
2.8.2	Tokenization.....	143
2.9.	Attachment decisions:.....	144
2.9	Punctuation.....	145
3	Noun Phrase and Adjective Phrase Structure.....	147
3.1	Heads.....	148
3.1.1	Numbers as heads.....	148
3.1.2	Quantifiers as heads:.....	149
3.1.3	Adjectives as heads.....	150
3.1.3.1	NP's headed by adjectives.....	151
3.1.3.1.1	Single-word NP's headed by adjectives.....	151
3.1.3.1.2	Multi-word NP's headed by adjectives.....	152
3.1.3.1.3	Modification of NP's headed by adjectives.....	153
3.2	Complements.....	154
3.2.1	Complements of nouns.....	154
3.2.1.1	Genitive complement :the construct state (“true” idafa).....	154
3.2.1.1.1	Noun not obligatory in construct state.....	155
3.2.1.1.2	Nouns with obligatory complements.....	157
3.2.1.1.2.1	Numbers (1-10 and hundreds, thousands, millions ...)......	157
3.2.1.1.2.2	Quantifiers.....	158
3.2.1.1.2.3	Nouns formerly treated as prepositions or particle.....	159
3.2.1.1.2.3.1	Nouns expressing possession or accompaniment:.....	160
3.2.1.1.2.3.2	Nouns of similitude.....	161
3.2.1.1.2.3.3	Nouns of exception.....	161
3.2.1.1.2.3.4	Nouns expressing locative or temporal relationships.....	164
3.2.1.1.2.4	Idiomatic annexation/idafa phrases.....	167
3.2.1.1.2.5	Prayer with gerund (masdar) completed by noun or bound pronoun.....	168
3.2.1.2	Accusative complements: tamiyz complements of numbers (11-99).....	169
3.2.2	Complements of adjectives.....	169
3.2.2.1	Semantic specificity of the non-true construct state.....	170
3.2.2.2	Adjectival head of non-true construct state (false idafa).....	172
3.3	Adjuncts.....	175
3.3.1	Adjectival modification.....	176
3.3.1.1	Single token noun, modified by a single token adjective.....	176

3.3.1.2	Multi-token NP, modified by a single token adjective	176
3.3.1.3	Multi-token adjective modifying a single token NP	177
3.3.2	Nominal modification	179
3.3.2.1	"Verbal corroboration"	179
3.3.2.2	Corroboration in meaning	180
3.3.2.3	Restrictive apposition (badal <i>permutation</i>)	182
3.3.2.3.1	Names in apposition	183
3.3.2.3.2	Demonstratives	184
3.3.2.3.3	NP modified by quantificational NP (non-universal)	185
3.3.2.3.4	NP modified by numerical NP	185
3.3.2.4	SBAR-NOM or S-NOM modified by an adjunct	185
3.4	Flat	186
3.4.1	Titles	186
3.4.2	Proper nouns	187
3.4.2.1	Case ending/part-of-speech policy for proper nouns	187
3.4.2.2	Treebanking proper nouns	188
3.4.3	Adjective single word	189
3.4.4	Compound numbers	189
3.4.5	Coordination between single words is flat	189
3.5	Pronoun constructions	190
3.5.1	Forms and cases of pronouns in Arabic	190
3.5.2	Clitic pronouns (bound pronouns)	193
3.5.3	Resumptive pronouns	194
3.5.4	Separating pronouns/ Damiyr Al-faSol	199
3.5.5	Anticipatory pronouns/ Damiyr Al\$a>on / ضمير الشأن	201
3.5.5.1	Not to be confused with referential pronouns following in >an~a or >in~a ..	202
3.6	Relative Clauses	204
3.6.1	Free relatives	205
3.7	Numbers, Quantifiers and Comparatives	207
3.7.1	Numbers: Cardinal vs. ordinal, noun vs. adjective	207
3.7.2	Numbers	212
5 thousands or 20 thousand men		212
3.7.3	Akvr "most"	224
3.7.3.1	Adjective	224
3.7.3.2	Adjective in nominal position	226
3.7.3.3	Noun quantifier	227
3.8	Differentiating when an adjective is an ADJP and when it is an NP-ADV	228
3.9	Adverbial pre-modifiers	229
3.10	NP and idafa rules, the quick guide	231
4	Verb Phrase Structure	232
4.1	Classification of verbs	233
4.1.1	Inflectional verbs	233
4.1.1.1	Event verbs	233
4.1.1.1.1	Intransitive verbs	233
4.1.1.1.2	Transitive with preposition	234
4.1.1.1.3	Transitive verbs	234

4.1.1.1.4	Ditransitive verbs	235
4.1.1.1.5	Event verbs with clausal complements	237
4.1.1.1.5.1	Verbs of speech	237
4.1.1.1.5.2	Verbs of judgment.....	238
4.1.1.1.5.3	حسب Hasiba think of something as.....	239
4.1.1.1.5.4	Verbs of transformation and labeling verbs	240
4.1.1.1.5.5	Test for distinguishing S vs. PP-CLR with verbs of transformation..	242
4.1.2	Non-event verbs	243
4.1.2.1	kAna and layosa.....	243
4.1.2.1.1	kAna	244
4.1.2.1.2	layosa.....	245
4.1.2.2	Sisters of kAna.....	247
4.1.2.2.1	Verbs of becoming أفعال الصيرورة.....	247
4.1.2.2.2	Verbs of duration or continuity أفعال الاستمرار	248
4.1.2.2.3	Verb of simultaneity فعل التوقيت.....	249
4.1.2.2.4	Additional event (not-predicative) uses of related verbs	250
4.1.2.3	Serial verbs: Non-event verbs with sentential complements (S and SBAR) ..	250
4.1.2.3.1	Verbs of beginning and continuation >afEAl a\$~uruE:	250
4.1.2.3.2	Verbs meaning ‘be on the point of..’ >afoEAl Almuqaraba	251
4.1.3	Non-inflectional verbs	252
4.1.4	Pseudo-verbs: <in~a and its sisters	256
4.1.4.1	IA as a pseudo-verb (IA bud~a, etc.), -SBJ only no -PRD	257
4.1.4.2	lakin~a usually heads an S-ADV, with exceptions.....	259
4.1.4.3	>an~a as complementizer, not pseudo-verb.....	259
4.1.5	New S complement construction in MSA	260
4.1.6	Interaction of active, passive, and topicalization with S complements	260
4.2	Subjects	261
4.2.1	Pre-verbal/Topicalized subjects	262
4.2.2	Clausal subjects.....	263
4.2.3	Empty subjects	263
4.3	Direct and indirect objects	265
4.3.1	Clitic object pronouns	267
4.4	Closely related prepositional phrases (PP-CLR)	267
4.4.1	With NP-OBJ and PP-CLR.....	268
4.4.2	With PP-CLR and PP-CLR.....	268
4.5	Adverbial modification	269
4.5.1	Adverbial modifiers vs. arguments	269
4.5.2	S-ADV vs. S-MNR (HAI).....	270
4.5.3	-DIR vs. -LOC	271
4.5.4	NP-TMP vs. NP-ADV	272
4.5.5	Nouns formerly prepositions.....	273
4.6	Passive Verbs	273
4.6.1	Passive of “monotransitive” verbs	274
4.6.2	Passive of ditransitive verbs.....	274
4.6.2.1	Passivized direct object.....	274
4.6.2.2	Passivized indirect object.....	275

4.6.3	Passive of verbs with a clausal complement.....	275
4.6.3.1	SBAR complement	275
4.6.3.2	S complement.....	275
4.6.3.2.1	Verbs of judgement	275
4.6.3.2.2	Verbs of transformation and labelling.....	276
4.6.4	Passive of verbs that are transitive with a PP-CLR:	276
4.6.5	Passive of verbs that are ditransitive with a PP-CLR:	277
4.6.6	Passives in the passive participle form	278
4.7	Active verbs that are semantically similar to passives.....	279
4.8	Floating Quantifiers	279
5	Gerunds (maSdar/gerund) and participles (>ism AlfaEil and >ism AlmafoEuwl).....	280
5.1	Non-verbal reading of gerunds and participles	280
5.1.1	Plural form	280
5.1.2	Single word, not in HAl position	280
5.1.3	Definite maSdar/gerund.....	282
5.1.4	Single word modified by an adjective	283
5.1.5	Single word followed by an adverbial of degree	284
5.1.6	Single word followed by a complement in genitive	284
5.1.7	maSdar/gerund followed by an SBAR complement.....	286
5.1.8	Single word followed by special 'li' PP	287
5.1.9	Adverbial modification	292
5.2	Verbal reading of gerunds and participles	293
5.2.1	Related argument in accusative	294
5.2.1.1	For maSdar/gerund.....	294
5.2.1.2	For participles	295
5.2.2	Argument PPs (PP-CLR and PP-DTV and PP-DIR).....	296
5.2.3	In HAl position	299
5.2.4	Argument in nominative	299
5.2.5	maSdar/gerund or Participle followed by an S or SBAR complement.....	302
5.2.6	Auxiliary kAwon as a maSdar	303
5.2.7	Annotation of verbal readings.....	303
5.2.7.1	maSdar/gerund is S-NOM.....	304
5.2.7.2	Participles are SBAR or adverbial S	304
5.2.7.2.1	Participial relative clauses.....	304
5.2.7.2.2	S-ADV vs. S-MNR (HAl and adverbial participial clauses)	305
5.2.7.2.3	Passive participles with passive trace.....	306
5.3	Distribution of S, S-NOM, S-ADV, NP, ADJP	308
6	Coordination	312
6.1	Flat	313
6.2	Coordination as adjunction	313
6.3	UCP.....	313
6.4	Right node raising (*RNR*)	313
6.5	Gapping/ VP template gapping.....	314
6.6	Multi-word conjunctions/CONJP	316
6.7	Arabic coordination without conjunction	316
6.8	NAC	316

6.9	Initial wa, other initial conjunction.....	316
7	Tricky Constructions.....	316
7.1	Tokens with more than one possible part-of-speech value.....	316
7.1.1	>ay/>ay~ (أَي/أَي).....	316
7.1.2	xalA, EadA, Ha\$A.....	318
7.1.3	Fa.....	319
7.1.4	mA(ما): WHNP, NP, SUB_CONJ, PRT, Verb, etc.....	322
7.1.4.1	mA WHNP.....	322
7.1.4.1.1	Relative Pronoun mA (with trace)	323
7.1.4.1.2	mA interrogative pronoun	324
7.1.4.1.3	mA exclamative pronoun	324
7.1.4.2	Quantifier/Indefinite mA "some" = NP	324
7.1.4.3	mA SUB_CONJ.....	325
7.1.4.3.1	Subordinating Complementizer mA (ما المصدرية mA >al-maSdariy~ah) ..	325
7.1.4.3.2	the conditional mA (ما الشرطية).....	326
7.1.4.4	Particle mA (PRT)	327
7.1.4.4.1	mA negative particle	327
7.1.4.4.2	mA AlzAidap(ما الزائدة)	328
7.1.4.5	mA verb ما المشبهة بليس	329
7.1.4.6	When to split “mA” and when not to split it.....	330
7.1.4.6.1	not split : mA is a part of SUB_CONJ, CONJ or PART	330
7.1.4.6.2	Split it: mA is a SUB_CONJ, REL_PRON, INTEROG_PRON	332
7.1.4.6.3	Split or not split both are possible with different meaning	335
7.1.4.6.3.1	The case of fiymA	335
7.1.4.6.3.1.1	The case of TAlamA	336
7.1.4.6.3.2	The case of bayonamA.....	337
7.1.4.6.3.3	The case of bimA	338
7.1.4.6.3.4	The case of kamA.....	339
7.1.4.6.3.5	The case of kul~amA	339
7.1.4.6.3.6	The case of <in~amA	340
7.1.4.6.4	When words related to “mA” need a trace and when not	341
7.1.4.6.4.1	There should be a trace, when it is a free relative and mA is a REL_PRON	341
7.1.4.6.4.2	There should not be a trace	341
7.1.4.6.4.2.1	When it is a SUB_CONJ.....	341
7.1.4.6.4.2.2	when it is part of SUB_CONJ like kulamA, TAlamA, etc.	342
7.1.5	wa- (وَ).....	342
7.1.5.1	Tests distinguishing the wa of coordination, NAC clauses introduced by wa, and the circumstantial wa.....	346
7.1.5.1.1	wa: SUB_CONJ or CONJ?	346
7.1.5.1.2	wa: CONJ (under NAC) or SUB_CONJ?.....	347
7.1.6	<i*/لِذْ ‘because’, ‘as’, ‘suddenly’	349
7.1.6.1	The subordinating conjunction <i*/ ‘because,’ ‘since,’ ‘as’ (SUB_CONJ) ...	349
7.1.6.2	The particle <i*/ ‘suddenly’ (PART).....	350
7.1.7	Hat~Y (حَتَّى).....	351
7.1.8	lA (لَا).....	353

7.1.9	The particle and the preposition $\text{إِلا} <il\sim A$	355
7.1.9.1	$<il\sim A$ as a Restrictive Particle	355
7.1.9.2	$<il\sim A$ as a Preposition.....	357
7.1.10	mun^*u and mu^*	358
7.1.11	$EalY$: PREP or NOUN.....	359
7.1.12	The particle and preposition $la-$	359
7.1.13	$layosa/ليس$ 'not be', 'not' : $layosa$ has 2 POS values and 2 matching syntactic annotations.	360
7.1.13.1	The perfective verb $layosa/$ 'not be' (PV).....	360
7.1.13.2	The negative particle $layosa$ 'not' (PART)	360
7.2	li heads PP only not SBAR	361
7.3	Dates	361
7.4	Compass directions	361
7.5	Sports scores	362
7.6	Internal combined with external modification, both are annotated (both times count!) 363	
8	Arabic terminology	364
8.1	$fAEil$, $nA\&ib\ fAEil$ and $mubtada'$	364
8.1.1	The $fAEil$ الفاعل (doer).....	364
8.1.2	The $nA\}ib\ fAEil$ نائب الفاعل (the substitute of the doer).....	364
8.1.3	The $mubtada'$ المبتدأ (what is started with)	365
8.2	the $mafEuwl\ bihi$ به المفعول (the recipient of the action)	366
8.3	The $mafEuwl\ muTlaq$ المفعول المطلق (the absolute object)	367
8.4	The $mafEuwl\ li>ajlih$ (the purposive object)	368
8.5	The HAi (state)	368
8.6	The $tamyiyz$	370
8.6.1	Noun complement when the noun that completed is... ..	370
8.6.1.1	NOUN_NUM (from 11 to 99) \rightarrow treebanked as Complementation.....	370
8.6.1.2	Nouns of measurement (meter, liter, kilo...) \rightarrow treebanked as NP-ADV	371
8.6.1.3	Kam in the interrogative sentence.....	371
8.6.2	The $tamyiyz$ is a verb modifier	371
9	Arabic Treebank Notation.....	371
9.1	Node labels and functional "dashtags"	371
9.2	Empty categories.....	372
9.3	Coordination	372
9.4	VP template gapping.....	372
9.5	Coreference	373
10	References	374
11	Index	375

1 Introduction

To be written...

1.1 A note about glosses

To be written...

1.2 A Note on Case Marking

To be written...

1.2.1 Diacritization: Vocalized and Unvocalized text /Case and Mood endings

The description of Arabic language idiosyncrasies and their impact on the annotation process and methodology in Maamouri and Bies (2004) can be summarized in the following points:

- (1) Leaving out the short vowel markers, consonantal length (šaddah), inflection and word-final case and mood markings is typical in most written Arabic. Vocalized MSA text is scarce and limited to a small number of literary, religious or school-related titles.
- (2) Most Arabic NLP applications seem to do away with all diacritics by working from a graphic representation which is stripped of many significant linguistic features, mostly relating to grammatical marking.
- (3) The reader reads the text and interprets its meaning by mentally providing the missing grammatical information (vocalization process) that leads to its/an acceptable interpretation. This amounts to an additional manual/human annotation with decisions that may have a non-trivial impact on the overall annotation routine in terms of both accuracy and speed.
- (4) The graphemic representation of vocalization diacritics is not absolutely necessary for Treebank annotation. However, its presence completes the text and enhances the quality of the linguistic analysis of the targeted corpus. Annotators doing the morphological analysis provide a first reading and an interpretation of the bare text based on an internalized knowledge of the required vowels and case/mood endings ('mental vocalization'). They produce a vocalized output, shown as a text with full vocalic diacritics in the text box of the TreeEditor tool, for syntactic analysis. Syntactic annotators can either accept or challenge the interpretation shown.
- (5) Providing a vocalized text for annotation will decrease the amount of additional ambiguity produced by the lack of grammatical and lexical markings. However, it should be made clear that, like any other language, Arabic will continue to have the usual amount of linguistic ambiguity.
- (6) Since readers have to provide the missing MSA grammar towards understanding and annotating newswire or other text corpora and since the level of internalized MSA grammar

differs drastically sometimes from annotator to annotator, there is an added degree of grammatical inconsistency which will negatively impact inter-annotator agreement rates.

1.3 A partial glossary of Arabic grammar terminology

To be written...

1.4 A note about punctuation in Arabic texts

To be written....

Final punctutation and the

Frequency (?) of independent sentences without intervening final punctuation (resulting in multiple S-expressions on a single line in the .tree files).

Annotators cannot change the punctuation tokens in any way (punctuation cannot be added or deleted).

2 Basic Arabic clause structure

For the most part, our syntactic/predicate-argument annotation of newswire Arabic follows similar annotation principles to the Penn English Treebank. The Penn English Treebank guidelines are available from the University of Pennsylvania Department of Computer and Information Science as the *Bracketing Guidelines for Treebank II Style Penn Treebank Project*, MS-CIS-95-06, www.cis.upenn.edu/~treebank. Our updated Arabic Treebank Guidelines will be available from LDC on-line.

Some points where the Penn Arabic Treebank differs from the Penn English Treebank:

- Arabic subjects are analyzed as VP internal, following the verb.
- Matrix clause (S) coordination is frequent.
- The function of NP objects of transitive verbs is directly shown as NP-OBJ.
- Co-reference is shown always on the node label, never on the empty category token itself.
- Gapping co-reference is always shown as '=' indexing, for *both* the template and the subsequent gap filling items.

An example of a sample annotated sentence is below. This tree includes both the syntactic node labels and the pre-terminal part-of-speech/morphological tag node level.

(TOP
 (S (CONJ wa-::و-::and)
 (VP (PRT (NEG_PART -lam::-لَمْ::did_not))
 (IV3MS+IV+IVSUFF_MOOD:J ya+tasana~a+[null]::يَتَسَنُّ[نُلل]::he/it+
 be_possible+[jus.])
 (PP-TMP (PREP Ealay::عَلَى::on/above)
 (NP (DET+NOUN+CASE_DEF_GEN Al+fawor+i::الْفَوْر::
 the+immediately+[def.gen.])))
 (NP-SBJ (NOUN+NSUFF_FEM_SG+CASE_DEF_NOM maEorif+ap+u::مَعْرِفَةٌ::
 knowledge /acquaintance+[fem.sg.]+[def.nom.])
 (NP (NOUN+CASE_DEF_GEN sabab+i::سَبَب::
 reason/cause+[def.gen.])
 (NP (DET+NOUN+CASE_DEF_GEN Al+HADiv+i::الْحَادِثُ::
 the+accident/mishap+[def.gen.]))))))
 ولم يتسنَّ على الفور معرفة سبب الحادث
 wa+lam + yatasan~a+o + Ealay + Al+fawor+i + maEorifap+u + sabab+i + Al+HADiv+i
 and+did not + be possible + on + the+immediately + knowledge + reason + the+accident
It was not immediately possible to recognize the reason behind the accident.

For ease of reading, however, most of the example trees in these guidelines will be shown without the part-of-speech/morphological tag nodes and without the TOP node, as follows:

(S wa-::و-::and
 (VP (PRT -lam::-لَمْ::did_not)
 ya+tasana~a+[null]::يَتَسَنُّ[نُلل]::he/it+be_feasible/be_possible+[jus.]
 (PP-TMP Ealay::عَلَى::on/above
 (NP Al+fawor+i::الْفَوْر::the+immediately/at_once+[def.gen.]))
 (NP-SBJ maEorif+ap+u::مَعْرِفَةٌ::
 knowledge /acquaintance+[fem.sg.]+[def.nom.]
 (NP sabab+i::سَبَب::reason/cause+[def.gen.]
 (NP Al+HADiv+i::الْحَادِثُ::the+accident/mishap+[def.gen.]))))))
 ولم يتسنَّ على الفور معرفة سبب الحادث
 wa+lam + yatasan~a+o + Ealay + Al+fawor+i + maEorifap+u + sabab+i + Al+HADiv+i
 and+did not + be possible + on + the+immediately + knowledge + reason + the+accident
It was not immediately possible to recognize the reason behind the accident.

2.1 Basic sentence structure

2.1.1 The form of Arabic sentences

The view most common among Arabic grammarians is that there are two types of sentences in Modern Standard Arabic (MSA):

- verbal sentence (جملة فعلية / jumolap fiEoliy~ap): a sentence with a main clause beginning with a verb
- nominal sentence (جملة اسمية / jumolap Aisomiy~ap) : a sentence with a main clause beginning with a noun (in the nominative)

More recently, some grammarians and linguists (such as Fischer, 2002,185) find that there are three sentence patterns in Arabic :

- ‘verbal sentences’ in which the predicate is a verb and the word order is a predicate-subject
- ‘nominal sentences’ in which the predicate is a noun or a pronoun and the word order is a subject-predicate or a predicate-subject
- ‘compound sentences’ in which the predicate is a verbal or nominal clause connected to the subject by a pronoun

Most of the existing research investigating Arabic sentence categorization has led to a great controversy around which word order (verb-subject-object (VSO) or subject-verb-object (SVO)) is the underlying Arabic word-order. Many MSA scholars such as Benmamoun (1992), Mohammad (1989), Bahloul et al. (1992), etc. consider that VSO sentences are based on an underlying SVO structure and that all preverbal NPs are genuine subjects. However, others, such as Fassi Fehri (1993), Ouhalla (1997), etc. follow traditional grammarians and much more recently the ‘Internal Subject Hypothesis’ (which assumes that verbs, in any language, assign theta-roles inside the VP, which means that the argument structure of any verbal predicate is realized within the VP (cf. Koopman & Sportiche, 1991)). These scholars consider the VSO order as the basic word order for MSA with any preverbal NP analyzed not as subject but as a left-dislocated element co-referring with the real subject, a trace which occurs postverbally. Many arguments have been provided by both groups to back up their respective claims but there is no conclusive corpus based evidence. MSA seems to allow both orders freely and the situation varies from text to text, style to style, and context to context. In our analysis of MSA sentences below, we opted for VSO as the underlying word order because we wanted to adhere as much as possible to the most commonly used Arabic grammar approach in order to ease the cognitive load on our annotators.

In the ATB, we limit MSA sentences to the two following types:

- **equational sentences** – similar to ‘nominal sentences’ – which we define in the strictest sense of the term as consisting of a subject and a predicate without an overt verb.
- **verbal sentences** are analyzed as sentences which consist of a main clause with a verb. Following the traditional grammar approach, we analyze the underlying basic sentence structure as verb-initial. Since the verb is actually not always the first element in many sentences in the data, we adopt a TOPICALIZATION STRUCTURE for arguments that are fronted before the verb and allow adverbials, prepositional phrases, and conjunctions to appear freely before the verb, which leads to a main clause beginning with a noun in the nominative traced after the verb (as SBJ, OBJ, complement of PREP or NOUN)

We have the following examples:

- equational sentence with subject and adjectival or nominal predicate:

(S (NP-SBJ ha*A هذا)
 (NP-PRD (NP (ADJP >akobaru أكبر
 (NP HAdivi حادث
 (NP EunfK عنيف))))
 (SBAR (WHNP-1 0)
 (S (VP ta\$ohadu تشهد
 (NP-OBJ (NP hu هـ)
 (NP-1 *T*))
 (NP-SBJ AlminoTaqapu المنطقة)
 (PP-TMP mun*u منذ
 (NP >ay~AmK أيام))))))
 هذا أكبر حادث عنيف تشهده المنطقة منذ أيام

ha*A + >kobaru + HAdivi + EunofK + ta\$ohadu+hu + AlminoTaq + mun*u + >aEowAmK
 this + biggest + incident + violence + witness+him+ the region + since + days
This is the biggest episode of violence ever witnessed for days

- verbal sentence with VP in the beginning of the main clause:

(S (VP hold/convene/conclude+it/they/she |Eaqadat | عَقَدَتْ
 (NP-SBJ the+federation |AlfiydiyAliy~apu | الفيدرالية)
 (NP-OBJ conference/convention |mu&otamarAF | مؤتمر
 journalistic/press |SaHAfiy~AF | صحافياً)
 (PP-LOC in |fiy | في
 (NP hotel |funoduqi | فندق
 (NP Martinez |mArrotiyniyz | مارتنيز
 -
 Beirut |bayoruwta | (بَيْرُوتُ)))
 (NP-TMP yesterday |>amosi | أمس))))
 عقدت الفدرالية مؤتمراً صحافياً في فندق مارتنيز بيروت امس

Eaqadat + Al+fidrAliy~ap+u + mu&otamar+AF + SaHAfiy~AF + fiy + funduq+i +
 mArrotiyniyz + bayruwt + >ams+a
 held + the+federation + conference + press + in + hotel + Martinez + Beirut + yesterday
Yesterday, the federation held a press conference in Martines Beirut hotel.

- verbal sentence with NP-TPC before the VP:

(S (NP-TPC-1 ministers |wuzarA'u | وُزَرَاءُ
 (NP (NP the+foreign_ministry/foreign_office |AlxArijjiy~api | الخارجية)
 and |wa- | و-
 (NP affairs/matters |-\$u&uwni | شُؤُون-
 (NP the+francophone |Alfranokuwfuwniy~api | (الفرنكُوفُونِيَّةُ))))))
 (VP they_+arrive/come_into |yafiduwna | يَفِدُونُ
 (NP-SBJ-1 *T*)
 (NP-MNR in_succession/consecutively/one_after_the_other |tibAEAF | تباعاً)
 (PP-DIR to/towards |<ilaY | إلى
 (NP Beirut |bayoruwta | (بَيْرُوتُ)))))
 وزراء الخارجية وشؤون الفرنفونية يفدون تباعاً الى بيروت

wuzarA'+u + Al+xArijjiy~ap+i + wa + \$u&uen+i + Al+franokuwfuwniy~ap+i + yafiduwna +
 tibAE+AF + <ilaY + bayruwt+a
 ministers + the+foreign office + and + affairs + the+francophone + arrive + consecutively +
 to +Beirut
The ministers of foreign offices and francophone affairs arrive consecutively to Beirut

2.1.2 Annotating basic elements in Arabic sentences

The sentence (S) is at the top level of structure (each "paragraph" also has a Paragraph label above any other brackets). The subject (labeled NP-SBJ) is inside VP after the verb. If the subject precedes the verb, it is labeled NP-TPC and traced to (NP-SBJ *T*) following the verb. All sentences have a subject (-SBJ) and a predicate (VP or -PRD). (NB: The VP is often the same as the S, if nothing precedes the verb.) The following examples are examples only, and many other types of these sentences occur (for example, equational sentences can occur with many different types of constituents as the predicates: NPs, ADJPs, PPs, etc., etc.).

A simple sentence with NP subject following the verb:

```
(S (VP استعاد:::{isotaEAd+a::recover/regain/reclaim+he/it_[verb]
  (NP-SBJ (NP الرئيس::Al+ra}iys+u::the+president/head/chairman+[def.nom.]
    الاميركي::Al+>amiyrokiy~+u::the+American+[def.nom.]
    (NP جورج::juwroj::George
      بوش::buw$::Bush)
    (NP-OBJ عافيت::EAfiy+at+a-::good_health+[fem.sg.]+[def.acc.]
      (NP ه::hu::his/its)
    (NP-TMP امس::>amosi::yesterday)))
  استعاد الرئيس الاميركي جورج بوش عافيته امس
  {isotaEAd+a + Al+ra}iys+u + Al+>amriykiy~+u + juwroj + buw$ + EAfiy+at+a+hu +
  >ams+i
  recover + the+president + the+American + George + Bush + good health+his + yesterday
  The American President George Bush recovered his good health yesterday
```

A simple sentence with pro-drop:

```
(S (VP أخذ::->axolad+a::eternalize/perpetuate/remain+he/it_[verb]
  (NP-SBJ (-NONE- *))
  (PP-CLR الى::<ilaY::to/towards
    (NP النوم::Al+nawom+i::the+sleep+[def.gen.])))
  .....nogloss)
أخذ الى النوم
>axolad+a + <ilaY + Al+nawom+i
perpetuate + to + the+sleep
He retired to bed
```

A simple sentence with a topicalized subject:

```
(S (fa-::fa-::and/so
  (NP-TPC-1 الجليد::Al+jaliyd+u::the+ice+[def.nom.])
  (VP سيطر::sayoTar+a::dominate/seize/command+he/it_[verb]
    (NP-SBJ-1 (-NONE- *T*))
    (PP-CLR على::EalaY::on/above
      (NP الطرق::Al+Turuq+i::the+roads/ways/methods+[def.gen.])))
  (NP-TPC-1 الجليد سيطر على الطرق
    fa + Al+jaliyd+u + sayoTar+a + EalaY + Al+Turuq+i
    so + the+ice + seize + the+roads
    So the ice seized the roads
```

An "equational" sentence with an adjectival predicate:

```
(S (NP-SBJ الملكة::Al+malikap+u::the queen)
  (ADJP-PRD جالسة::jAlisap+N::sitting)
  (PP على::EalaY::on/above
    (NP عروش::Euruw$+i-::thrones+[def.gen.]
      (NP ها::-hA::its/their/her))))
  الملكة جالسة على عروشها
  Al+malikap+u + jAlisap+N + Ealay + Euruw$+i+hA
  the+queen + sitting + on + thrones +her
  The queen is sitting on her thrones
```

An "equational" sentence with a nominal predicate:

```
(S (NP-SBJ السبعة::Al+saboEap+u::the+seven)
  (NP-PRD رقم::raqom+N::number/numeral+[indef.nom.]
    مهم::muhim~+N::important/serious+[indef.nom.]
    السبعة رقم مهم
    Al+saboEap+u + raqom+N + muhim~+N
    the seven + number + important
    Seven is an important number
```

2.2 Making corrections to morphological/POS tags

The morphological/POS tag is the lowest level node in the .tree files. This tag is represented in a separate column of the Treebank annotation tool rather than in the tree itself in order to make the structure of the trees legible for the annotators.

If the morphological/POS tag is incorrect, the Treebank annotator should mark the token with an X node when the incorrect POS tag affects the tree structure.

For example, in the following tree, the verb was incorrectly POS-tagged as a noun. The correct tree annotation is the VP, and the POS needs to be corrected.

Incorrect version:


```
(S-HLN (CONJ wa-||and)
  (NP-TPC-2 (DET+NOUN+CASE_DEF_NOM -Al+maEo$ar+u||the+assembly/
    community+[def.nom.])))
  (VP (NOUN+CASE_DEF_ACC xalof+a||behind+[def.acc.])
    (NP-SBJ-2 *T*)
    (NP-OBJ (NP (DET+NOUN+CASE_DEF_GEN Al+xaTiyb+i||the+orator/
      lecturer/preacher+[def.gen.])))
    (PP (PREP fiy||in)
      (NP (DET+NOUN+NSUFF_FEM_SG+CASE_DEF_GEN Al+xArijiy~+ap+i||
        the+foreign_ministry/foreign_office+
        [fem.sg.]+[def.gen.]))))))
```

Corrected version:

```
(S-HLN (CONJ wa-||and)
  (NP-TPC-2 (DET+NOUN+CASE_DEF_NOM -Al+maEo$ar+u||the+assembly/
    community+[def.nom.])))
  (VP (PV+PVSUFF_SUBJ:3MS xalaf+a|| succeed/come after + he/it)
    (NP-SBJ-2 *T*)
    (NP-OBJ (NP (DET+NOUN+CASE_DEF_GEN Al+xaTiyb+i||the+orator/
      lecturer/preacher+[def.gen.])))
    (PP (PREP fiy||in)
      (NP (DET+NOUN+NSUFF_FEM_SG+CASE_DEF_GEN Al+xArijiy~+ap+i||
        the+foreign_ministry/foreign_office+
        [fem.sg.]+[def.gen.]))))))
```

These corrections will be made in a separate pass using a tool specifically designed to make this type of correction.

2.3 Node labels and functional "dashtags"

In Penn Treebank annotation, when we refer to “node labels” we mean the labels given to the syntactic nodes in the trees. In the released version of the corpora, the part-of-speech/morphological tags form the pre-terminal nodes. However, in the Treebank annotation process, and therefore in these guidelines, the pre-terminal POS/morphological tags are not shown in the trees themselves, since they make the trees very difficult for annotators to read. The POS/morphological tags can be seen by the Treebank annotators in another part the annotation tool.

“Dashtags” mark more or less semantic function (-SBJ subject, -OBJ object, -ADV adverbial, -TMP temporal, -PRD predicate, etc.). They are appended to the syntactic node labels, separated by a dash (NP-SBJ, PP-TMP, ADJP-PRD, for example). Functional dashtags are used only if they are relevant, not on every node label (see VP arguments and adjuncts below).

2.3.1 Node labels

This section includes a description of all of the syntactic node labels used by the Penn Arabic Treebank.

2.3.1.1 Types of clauses

We make a distinction among a number of different clause types: S, SBAR, SQ, SBARQ, and FRAG.

2.3.1.1.1 S

Simple declarative sentence with a verb and with the subject following the verb:

```
(S (VP استعاد ::{isotaEAd+a::recover/regain/reclaim+he/it_[verb]
  (NP-SBJ (NP الرئيس ::Al+ra}iys+u::the+president/head/chairman+[def.nom.]
    الاميركي ::Al+>amiyrokuy~+u::the+American+[def.nom.]
  (NP-OBJ عافيت ::EAfiy+at+a-::good_health+[fem.sg.]+[def.acc.]
    (NP ه ::-hu::his/its)
  (NP-TMP امس ::>amosi::yesterday)))
```

استعاد الرئيس الاميركي عافيته امس

{isotaEAd+a + Al+ra}iys+u + Al+>amriykiy~+u + EAfiy+at+a+hu + >ams+a
 recover + the+president + the+American + good health+his + yesterday
the American President recovered his good health yesterday

Simple declarative sentence with a verb and with the subject preceding the verb:

```
(S (NP-TPC-1 الجليد ::-Al+jaliyd+u::the+ice+[def.nom.]
  (VP سيطر ::sayoTar+a::dominate/seize/command+he/it_[verb]
    (NP-SBJ-1 *T*)
    (PP-CLR على ::EalaY::on/above
      (NP الطرق ::Al+Turuq+i::the+roads/ways/means/methods+[def.gen.])))
```

الجليد سيطر على الطرق

Al+jaliyd+u + sayoTar+a + EalaY + Al+Turuq+i
 the+ice + seize + on + the+roads
The ice seized the roads

Simple declarative sentence without a verb:

```
(S (NP-SBJ السبعة ::Al+saboEap+u::the+seven)
  (NP-PRD رقم ::raqom+N::number/numeral+[indef.nom.]
    مهم ::muhim~+N::important/serious+[indef.nom.])))
```

السبعة رقم مهم

Al+saboEap+u + raqom+N + muhim~+N
 the seven + number + important
Seven is an important number

In passive sentences, the surface subject is tagged –SBJ, the passive trace is indicated with (NP-OBJ *) and coindexed to the surface subject. Arabic does not have logical subjects in passive sentences (the *by* phrase in English passives), so there is no use of the –LGS tag in the Arabic Treebank.

```
(S (VP قتل::qutil+a::be_killed+he/it_[verb]
  (NP-SBJ-1 ثلاثة::valAv+ap+u::three+[fem.sg.]+[def.nom.]
    (NP جنود::junuwd+K::soldiers+[indef.gen.])))
  (NP-OBJ-1 *)))
```

قتل ثلاثة جنود

qutil+a + valAv+ap+u + junuwd+K
 be killed + three + soldiers
Three soldiers were killed

Imperative sentences are labeled S and given a null subject (NP-SBJ *). They can be distinguished from declarative sentences by the POS/morphological tag of the verb, which will include “CV” as the core tag of imperative verbs.

```
(S (VP (CV+CVSUFF_SUBJ:2MS خذ::xu*+::take+you_[verb])
  (NP-SBJ (-NONE- *))
  (NP-OBJ (NP عصا::EaSA-::stick/baton/rod/staff
    (NP ك::-ka::your_[masc.sg.])))
    و::wa-::and
    (NP ستائر::-satA}ir+a-::veil/curtain/cover+[def.acc.]
      (NP ك::-ka::your_[masc.sg.] )))))
```

خذ عصاك وستائرک

xu* + EaSa+ka + wa + satA}ir+a+ka
 take + stick+your + and + curtains+your
Take your stick and curtains

If the name of the addressee appears with the imperative, it is given the dashtag –VOC (vocative). The vocative is NOT coindexed to the null surface subject.

```
(S (VP (CV+CVSUFF_SUBJ:2MS اَعْلَم |AiEolamo |know/be_aware+you )
  (NP-SBJ *)
  (NP-VOC (PRT oh/you |yA | يا )
    Aba |>abA | أبا
    Ammar |Eam~Ar | عَمَّار )
  (SBAR that |>an~a | أَنَّ
    (S (NP-TPC-4 honor/distinction |$arafa | شَرَف
      (NP the+Jerusalem |Alqudosi | الْقُدْس ))
      (VP he/it+deny/refuse |ya>obaY | يَأْبَى
        (NP-SBJ-4 *T*)
        (SBAR to |>an | أَنْ
          (S (VP he/it+be_liberated | yataHar~ara | يَتَحَرَّرُ
            (NP-SBJ *)
            (PP-MNR (PRT however/except/other | <il~A | إِلَّا
              on/above |EalaY | عَلَى
              (NP hands |>ayodiy | أَيْدِي
                (NP the+believer |Almu&ominiyna |
                  الْمُؤْمِنِينَ ))))))))
            اعلم يا ابا عَمَّار ان شرف القدس يابى ان يتحرر الا على ايدي المؤمنين
            <iElam+o + yA + >abA + Eam~ar + >an~a + $araf+a + Al+qudos+i + ya>baY + >an +
              yataHar~a+a + <il~A +EalaY + >ayodiy + Al+mu&miniyn
            be aware + oh + Aba + Ammar + that + honor + the+Jerusalem + refuse + to + be liberated +
              except + on + hands + the+believers
```

Aba Ammar, be aware that the honor of Jerusamel refuses to be liberated except by the believers

Participial, gerund and masdar clauses are annotated as sentential in contexts where the participle/gerund/masdar has a verbal reading. If they are also in a nominal position, they need to have a –NOM dashtag. If they are in an adverbial position, they need to have the appropriate adverbial dashtag. See sections ????? and ????? below for a full discussion of participials, gerunds and masdars.

```
(S (VP >axa*nA::أَخَذْنَا::take/begin+we_[verb]
  (NP-SBJ *)
  (PP-CLR fiy::فِي::in
    (NP Al+{iEotibAr+i::عَتَبَارِ}ال::the+consideration/regard+[def.gen.] ))
  (PP fiy::فِي::in
    (NP taHoDiyAr+At+i-::تَحْضِيرَات::preparation/production+[fem.pl.]+[def.gen.]
      (NP -nA::-نَا::our )))
  (NP-OBJ <imokAn+a::إِمْكَان::possibility/possible+[def.acc.]
    (S-NOM (VP takab~ud+i-::تَكَبُّد-
      ::suffering/sustaining/undergoing+[def.gen.]
      (NP-SBJ -him::-هِمْ::their )
      (NP-OBJ xasA}ir+a::خَسَائِر::
        ::losses/ damages+[indef.acc.] ))))))
  أَخَذْنَا فِي الْإِعْتِبَارِ فِي تَحْضِيرَاتِنَا إِمْكَانَ تَكَبُّدِهِمْ خَسَائِرَ
```

>axa*nA + fiy + Al+<iEtibAr+i +fiy + taHdiyrAt+i+nA + <imkAn+a + takab~ud+i+him +
 xasA}ir+a
 we took + in + consideration + in + preparation+our + possibility + undergoing+their +
 losses
In our preparations, we took in consideration the possibility causing them losses.

S-NOM is used for the complement of a preposition:

```
(S (NP-TPC-1 EanAn::عَنَّان::Annan )
  (VP yu+qal~il+u::يُقَلِّل::he/it+lessen/reduce/diminish+[ind.]
    (NP-SBJ-1 *T*)
    (PP min::مِنْ::from
      (S-NOM (VP rafoD+i::رَفُوض::rejection/refusal+[def.gen.]
        (NP-SBJ bagodAd+a::بَغْدَاد::Baghdad+[indef.gen.] )
        (NP-OBJ Eawod+ap+a::عَوْدَةٌ::return+[fem.sg.]+[def.acc.]
          (NP Al+mufat~i$+iyna::الْمُفَتِّشِينَ
            ::the+inspector/supervisor+[masc.pl.gen.]
            ))))))
      عَنَّانُ يُقَلِّلُ مِنْ رَفْضِ بَغْدَادِ عَوْدَةَ الْمُفْتَشِينَ
```

EanAn + yuqal~il+u + min + rafoD+i + bagodAd+a + Eawodap+a + Al+mufat~i\$+iyna
 Annan + reduce + from + refusal + Baghdad + return + the inspector+[masc.pl.gen]
Annan reduces Baghdad's refusal to allow the inspectors' return.

S-ADV is used for adverbial clauses:

(S (VP -حَلَقَتْ | -Hal~aqat | encircle/whirl+it/they/she
 (NP-SBJ-1 طائرة |TA}irapN | plane)
 (NP-LOC فَوْق |fawoqa | above/over
 (NP مناطق |manATiqi | areas/zones/territories
 (NP الجنوب |Aljanuwbi | the+south/South)))
 (S-ADV (VP مُتَّجِهَةً |mut~ajihapF | directed/aimed
 (NP-SBJ-1 *)
 (NP-DIR-CLR north/North |\$amAlAF | شَمَالاً))))))
 حلقت طائرة فوق مناطق الجنوب متجهة شمالاً
 Hal~aqat + TA}irap+N + fawoqa+a + manATiq+i + Al+januwbi+i + mut~ajihap+F +
 \$amAl+AF
 whirl + plane + above + areas + the+south + directed + north
A plane whirled above the southern areas directed to the north

2.3.1.1.2 SBAR

SBAR is used for relative clauses and subordinate clauses that have either an overt complementizer or an empty one.

Subordinate clause:

(S و::wa-::and
 (VP أوضح ::>awoDaH+a::clarify/explain/indicate+he/it_[verb]
 (NP-SBJ الكولونيل ::Al+kuwluwniyl::the+colonel)
 (SBAR ان ::>an~a::that
 (S (NP-TPC-1 الفحص ::Al+faHoS+a::the+examination/checkup+[def.acc.]
 الطبي ::Al+Tib~iy~+a::the+medical+[def.acc.]
 (VP أظهر ::>aZohar+a::show/manifest/demonstrate+he/it_[verb]
 (NP-SBJ-1 *T*)
 (SBAR ان ::>an~a::that
 (S (NP-SBJ بوش ::buw\$::Bush)
 (PP-PRD ب ::bi-::by/with
 (NP خير ::-xayor+K::good+[indef.gen.]))))))))))
 وأوضح الكولونيل ان الفحص الطبي أظهر ان بوش بخير
 wa+>awoDaH+a + Al+kuwluwniyl + >an~a + Al+faHoS+a + Al+Tib~iy~+a >aZohar+a +
 >an~a + buw\$ + bi+xayor+K
 and + clarified + the+colonel + that + the+examination + the+medical + show + that + Bush
 + with+good
The colonel clarified that the medical examination showed that Bush was in a good condition

Subordinate clauses with null complementizers:

(S (VP *akara ذَكَرَ
 (NP-SBJ *)
 (SBAR *0*
 (S (VP layosa لَيْسَ)
 (PP-PRD fiy فِي
 (NP maqduwri مَقْدُورِ
 (NP hi هِ))
 (NP-SBJ AltaHam~ulu اَلْتَحْمُلُ))))))

ذكر ليس في مقدوره التحمل

*akar+a + layosa + fiy + maqoduwri+hi + Al+taHam~ul+u
 mention + not + in + ability+his + the+patience
He mentioned that he can not stand it

(S (VP عسى::EasaY::perhaps)
 (SBAR-SBJ *0*
 (S (VP يتكاثر::ya+takAvar+u::he/it+become_numerous/
 increase+[ind.]
 (PP-LOC في::fiy::in
 (NP لبنان::lubonAn+a::Lebanon+[indef.gen.]
 (NP-SBJ المخلصون :: Al+muxoliSu+wna [def+nom]))))))
 عسى يتكاثر في لبنان المخلصون

EasaY + ya+takAvar+u + fiy + lubonAn+a + Al+muxoliSu+wna
 may + increase + in + Lebanon + the faithful
May the faithful people increase in Lebanon

(S (VP -ya+jib+u::يَجِبُ::he/it+be_necessary/be_incumbent+[ind.]
 (NP-SBJ Al+{inotizAr+u::اَلْاِنْتِظَارُ::the+waiting/anticipating+[def.nom.]
 (PP-PRP li-::لِ-::for
 (SBAR *0*
 (S (VP -ta+tagay~ar+a::تَتَغَيَّرُ::it/them/her_to+
 be_changed +[sub.]
 (NP-SBJ *)))))))

يجب الانتظار للتغير

yajib+u + Al+<inotiZAr+u + li+tatagay~ar+a
 necessary + the+waiting + to+change
Waiting is necessary for it to change

Relative clause – relative clauses are always adjoined to the NP they modify (see section ??? for more on relative clauses):

(NP-OBJ (NP (NP زوجت::zawoj+at+a-::wife+[fem.sg.]+[def.acc.]
 (NP -::hu::his/its))
 (NP لورا::luwrA::Lora))
 (SBAR (WHNP-2 التي::Al~atiy::which/who/whom_[fem.sg.]
 (S (VP كانت::kAn+at::be/was/were+it/they/she_[verb]
 (VP تتحدث::ta+taHad~av+u::it/they/she+speak/discuss+[ind.]
 (NP-SBJ-2 *T*)
 (PP على::EalaY::on/above
 (NP الهاتف::Al+hAtif+i::
 the+telephone+[def.gen.]))))))))
 زوجته لورا التي كانت تتحدث على الهاتف

zawoj+at+a+hu + luwrA + Al~atiy + kAn+at + ta+taHad~av+u + Ealay + Al+hAtif+i

(SQ (PRT does/do?/did?/is/are? |hal | هل)
 (NP-SBJ (NP measures/steps |<ijorA'Atu | إجراءات
 (NP the+government/administration |AlHukuwmap | الحُكُومَة |
 (ADJP the+taken/adopted |Almut~axa*api | المُنْتَخَذَة |
 (ADJP-PRD guaranteeing |kafiylapN | كَفَيْلَة
 (PP by/with |bi- | بـ
 (NP treatment/therapy/processing | -muEAlajapi | مُعَالَجَة -
 (NP the+conditions/situation/status |
 Al>awoDAEi | الأَوْضَاع |
 the+economic/economical |
 AlAiqtisAdiy~api | الإِقْتِصَادِيَّة |
 هل اجراءات الحكومة المتخذة كفيلة بمعالجة الاوضاع الاقتصادية
 hal + |<ijorA'At+u + Al+Hukuwmap+i + Al+mut~axa*ap+i + kafiylap+N bi+muEAlajap+i +
 Al+>awDAE+i + the+taken + Al+<iqtiSAdiy~ap+i
 does/do?/did?/is/are? + measures + the government + guarantee + with + treatment +
 the+situation + the+ economic
Do the measures taken by the government guarantee the treatment for the economic situations?

(SQ (PRT ا >a did)
 (VP سمعت samiEo+ta hear
 (NP-SBJ *)
 (NP-OBJ قولي qawol+iy saying+my)
 (NP-ADV جيداً jay~id+F well)))
 أ سمعت قولي جيداً؟
 >a samiEo+ta qawol+iy jay~id+F
 did hear saying+my well
Did you hear what I said well?

NB. Yes/no questions using the two interrogative particles “>a” and “hal” are annotated as SQ. However, questions formed using the interrogative pronoun are annotated as SBARQ (see the following section ???). Echo questions are also SQ.

(SQ (VP ذَهَبْتَ *ahabo+ta:: went+you
 (NP-SBJ *)
 (WHADVP أين >ayona :: where)))
 ?)
 ذَهَبْتَ أين؟
 *ahabot+a + >ayona ?
 went + where?
You went where?


```
(SQ (SQ (PRT هل)
        (NP-SBJ هم)
        (PP-PRD من
         (NP القاعدة )))
  أم
  (SQ (NP-SBJ هم)
       (NP-PRD قطاع
        (NP (طرق))))
```

هل هم من القاعدة أم هم قطاع طرق

Do they from AlQaida or they **cutter** the roads

Are they from AlQaida or are they bandits?

NB: need transliteration for the full sentence and in the tree, also needs to have the gloss corrected

The *halo* >*an~a* in the question example below is considered to be an error, since it should be *halo* <*in~a*. The correct annotation is as follows, with the SQ for the sentence and the X node for the >*an~a*.

```
(SQ (PRT هل halo)
    (VP (XP أَنْ >an~a)
        (NP-SBJ ه hu)
        (ADJP-PRD قادم qAdim+N)))
```

هل إنه قادم

Do indeed he coming

Is he coming?

NB: need transliteration for the full sentence and in the tree, also needs to have the gloss corrected

Indirect question complements of verbs of speech are annotated as S or SBAR, not as SQ or SBAR.

```
(S (VP سال
    (NP-SBJ *)
    (S (PRT هل)
        (VP تحسنت
         (NP-SBJ الأوضاع )))))
```

سال هل تحسنت الأوضاع

Asked did get better the situations

He asked whether the situation improved.

NB: need transliteration for the full sentence and in the tree, also needs to have the gloss corrected

2.3.1.1.4 SBARQ

SBARQ is used for questions formed using WH- words.

With nominal Wh- words, WHNP:

```
(SBARQ (WHNP-1 ماذا::mA*A::what)
  (S (NP-PRD لدي::laday::with/by
    (NP لنا::nA::us))
    (NP-SBJ-1 *T*)) )
```

ماذا لدينا

mA*A + laday+nA
 what + possession+our(s)
what do we have?

With adverbial Wh- words, WHADVP:

```
(SBARQ (WHADVP-1 why |limA*A | لماذا)
  (S (VP he/it+pulverize/scatter/blow_up |yanosifu | يَنْسِفُ
    (NP-SBJ Arafat |EarafAt | عَرَفَات)
    (NP-OBJ bridges/_beams_or_bars |jusuwra- | جُسُورَ-
    (NPits/his |-hu | -هُ))
    (ADVP-PRP-1 *T*)) )
  ?)
```

لماذا ينسف عرفات جسوره؟

limA*A + yanosifu + EarafAt + jusuwra+hu
 why + blow up + Arafat + bridges+his
Why does Arafat blow up his bridges?

With a modified WHNP:

```
(SBARQ ف (WHNP-1 (WHNP كَمْ)
  (PP من
    (NP لنا)) )
  (S (PP في
    (NP صَلَات
      (NP و)) )
    (VP يرْتَل
      (NP-SBJ-1 *T*)
      (NP-OBJ الفَاتِحَة))))
```

فَكَمْ مِنَّا فِي صَلَاتِهِ يَرْتَلُ الْفَاتِحَةَ

So how many of us in prayer his chants Surat the fateha

So how many of us chant in prayer Alfateha? [NB: is this translation correct? I'm not quite sure what it means]

NB: need transliteration for the full sentence and in the tree, also needs to have the gloss corrected

```
(SBARQ (WHNP-1 (WHNP اي)
                (PP من
                  (NP هم)))
  (S (NP-SBJ (NP هو)
              (NP-1 *T*))
     (NP-PRD الوزير)))
```

اي منهم هو الوزير؟

Which of them he the minister

Which of them is the Minister?

NB: need transliteration for the full sentence and in the tree, also needs to have the gloss corrected

Indirect question complements of verbs of speech are annotated as S or SBAR, not as SQ or SBARQ.

```
(S (VP (PRT لم)
        نعرف
      (NP-SBJ *)
      (SBAR (WHADVP -1 كيف)
        (S (VP يفكر
              (NP-SBJ غير
                (NP نا)))
            (ADVP-MNR-1 *T*)))))
.)
```

لم نعرف كيف يفكر غيرنا.

We not know how think the others

We have never known how other people think.

NB: need transliteration for the full sentence and in the tree, also needs to have the gloss corrected

NB: why is the WHADVP word in red? Red = problem in this edit, but not sure what the problem is.

2.3.1.1.5 FRAG

FRAG is used for sentence fragments – that is to say, two or more constituents that need to be held together as a grammatical statement, but that are not in a subject-predicate relationship. Note, however, that a single top-level constituent that is not a full sentence is labeled with its constituent type (for example, NP or SBAR). See section ????? for examples. FRAG is used only if at least two or more constituents need to be held together.

Typical examples of FRAG include headlines, news story bylines, sports scores, non-sentential lists, etc. (See section ????? below for other examples that should not be FRAG, and what they should be instead.)

Headlines:

Many headlines are properly treated as FRAG fragments rather than as S sentences. It is important to distinguish between equational sentences (where there is a true subject and

predicate relationship) and fragments (where there may be something that looks like a subject, but often not a real predicate for it).

See also section ????? on the use of the dashtag –HLN.

```
(FRAG (NP تونس)
      :
      (ADJP أهم
        (NP المطالب
          الشعبية )
        (PP في
          (NP مظاهرات
            الخبز
              و
                الحرية))))
```

تونس: أهم المطالب الشعبية في مظاهرات الخبز والحرية

Tunisia: the most important the claims the popular in demonstrations the bread and the freedom

Tunisia: the most important popular claims in the bread and freedom demonstrations.

NB: need transliteration for the full sentence and in the tree, also needs to have the gloss corrected

Bylines:

```
(FRAG (NP-LOC القاهرة::Al+qAhir+ap+u::the+Cairo+[fem.sg.]+[def.nom.])
      -: -:nogloss
      (PP من::min::from
        (NP جمال::jamAl::Jamal/Gamal فهمي::fahomiy::Fahmi))
      : : : : :nogloss)
```

القاهرة من جمال فهمي

Al+qAhirap+u + min + jamAl + fahmiy

Cairo + from + Jamal Fahmi

Cairo from Jamal Fahmi

Scores:

```
(FRAG (NP (NP (NP بورتلاند::buwrotlAnod::Portland ترايل::
              trAyl::NOT_IN_LEXICON بلايزرز::blAyzrz::NOT_IN_LEXICON)
            (NP 96::96::nogloss))
      -: -:nogloss
      (NP (NP (NP كليفلاند::kliyfolAnd::Cleveland كافاليرز::
              kAfAloyiyrz::Cavaliers)
            (NP 64::64::nogloss)))
```

بورتلاند ترايل بلايزرز كليفلاند كافاليرز

buwrotlAnod + trAyl + blAyzrz + 96 + kliyfolAnd + kAfAloyiyrz + 64

Portland + trail + blazers + 96 + Cleveland + Cavaliers + 64

Portland trail blazers 96, Cleveland Cavaliers 64

Lists:

```
(FRAG (FRAG (NP-TMP عام::EAm+a::year+[def.acc.]
              (NP 1981::1981::nogloss))
      (NP (NP 9.51::9.51::nogloss)
          (PP في::fiy::in
              (NP المئة::Al+mi}+ap+i::the+[per]_cent))))
      و::wa::and
      (FRAG (NP-TMP عام::EAm+a::year+[def.acc.]
              (NP 1979::1979::nogloss))
      (NP (NP 3.63::3.63::nogloss)
          (PP في::fiy::in
              (NP المئة::Al+mi}+ap+i::the+[per]_cent+[fem.sg.]+[def.gen.])))
      و::wa::and
      (FRAG (NP-TMP عام::EAm+a::year+[def.acc.]
              (NP 1982::1982::nogloss))
      (NP (NP 2.12::2.12::nogloss)
          (PP في::fiy::in
              (NP المئة::Al+mi}+ap+i::the+[per]_cent+[fem.sg.])))
      وعام 1981 9.51 في المئة و عام 1979 3.63 في المئة و عام 1982 2.12 في المئة
      wa+EAm + 1981 + 9.51 + fiy + Al+mi}ap+i + muqAbil + 3.63 + fiy + Al+mi}ap+i +
      EAm+a + 1979 + wa + 2.12 + fiy + Al+mi}ap + EAm+a + 1982
      and+year + 1981 + 9.51 + per + cent + vs. + 3.63 + per + cent + year + 1979 + and + 2.12 +
      per + cent + year + 1982
      and in 1981 9.51 percent vs. 3.63 percent in 1979 and 2.12 percent in 1982
```

```
(FRAG (NP-ADV ثانيا::vAniy+AF::second/next+[acc.indef.]
      :::::nogloss
      (NP بديل::badiyl+N::substitute/replacement/alternative+[indef.nom.]
          عسكري::Easokariy~+N::military/army+[indef.nom.]
      , :::::nogloss)
```

ثانيا بديل عسكري

vAniy+AF + badiyl+N + Easokariy~+N
 second + alternative + military
Second, a military alternative

Sentence with segmentation errors

We also annotate as FRAG any sentential clause with missing constituents, not because of a particular stylistic choice, but because of an obvious segmentation error where a single sentence was split over two TOP or SU or “Paragraph” segments during an earlier stage of annotation.

```
(TOP (FRAG (FP -wa-and
            (VP -جاء-jA'+a.arrive/come/occur+he/it_[verb]
                (PP في.fiy.in
                    (NP البيان.Al+bayAn+i.the+communiqu/statement))))
```

وجاء في البيان

wa + jA'a + fiy + AlbayAni
 and + came + in + the+statement
And it was reported in the statement

(TOP (S و-·wa-·and
 (VP -لاحظ-·lAHaZ+a·notice/observe+he/it_[verb]
 (NP-SBJ (NP الناطق·Al+nATiq+u·the+spokesman/speaker)
 (PP ب-·bi-·by/with
 (NP -{سم-·{isom+i·name+[def.gen.]
 (NP (NP وزارة·wizAr+ap+i·ministry
 (NP الخارجية·Al+xArijiy~+ap+i·
 the+foreign_ministry))
 (ADJP الباكستانية·Al+bAkistAniy~+ap+i·
 the+Pakistani))))
 (FRAG أن·>an~a·that))))

ولاحظ الناطق باسم وزارة الخارجية الباكستانية أن

wa + lAHaZa + alnATiq + b+{ismi + wizArapi + AlxArijiy~api + AlbAkistAny~api +
 >an~a
 and + noted + the spokesman + with+name + ministry + the+foreign affairs + the+Pakistani
 + that
And the spokesman of the Pakistani ministry of foreign affairs noted that

(TOP (FRAG و-·wa-·and
 (PP-PRD -من·-min·from
 (NP (ADJP أبرز·>aboraz+i·more/most_prominent+[def.gen.]
 (NP التوصيات·Al+tawoSiy+At+i·
 the+recommendations))))
 :·))

ومن أبرز التوصيات

Wa + min + <aborazi + AltawoSiyAti
 And + from + most prominent + the+recommendations
And among the most important recommendations are

(TOP (FRAG و-·wa-·and
 (NP (NP رئيس·-ra}iys+u·president/head/chairman+[def.nom.]
 (NP الهند·Al+hinod+i·the+India+[def.gen.]
 (NP نارايانان·nArAyAnAn·Narayanan))))

ورئيس الهند نارايانان

wa + ra}iysu + Alhinod + nArAyAnAn
 and + president + the+India + Narayanan
And India's president Narayanan

```

(TOP (FRAG وَ-·wa-·and
      (SBAR -أَنْ·->an·to
        (S (VP (PRT لَا·lA·no/not/non-)
              تَمْضِينَ·tu+moDiy+na·you_[fem.pl.]+sign+[fem.pl.]
            (NP-SBJ *)
            (PP-CLR عَلَى·EalaY·on/above
              (NP (NP صُكُوكْ·Sukuwk+i·deeds/documents+[def.gen.]
                  (NP (NP بَيْعْ-·bayoE+i-·sale/selling
                      (NP -كُنْ·-kun~a·your_[fem.pl.])))
                  (NP-ADV بَيْعْ·bayoE+a·sale/selling
                      (NP الرِّقِيقِ·Al+raqiyq+i·
                        the+slaves))))))))))
              وَأَنْ لَا تَمْضِينَ عَلَى صُكُوكْ بِيَعِ الرِّقِيقِ
wa + >an + lA + tumoDiy + EalaY + Sukuwki + bayoEi+kun~a + bayoEa + Alraqiyqi
and + that + not + sign + on + deeds + selling+your + selling + the+slaves
And that you should not sign papers selling you as if you were slaves

```

2.3.1.1.5.1 Not FRAG

Several constructions are easily mistaken as FRAG. These are usually constructions that should be annotated as single constituents (and therefore, not fragments), but that were incorrectly analyzed as consisting of multiple independent constituents. This section contains examples of incorrect FRAG annotation, along with the correct annotation of these examples.

In the below example, the NP *Al>uroduniwuna* and the PP following it are not held together in an obvious grammatical statement, and so should not be adjoined under the same node FRAG.

Incorrect use of FRAG:

```
(TOP (FRAG (NP الازدنيون·Al+>urodun~iy~+uwna·the+Jordanian+[masc.pl.])
  (PP-PRP لـ·li-·for/to
    (SBAR -أَنْ-·->an~a-·that
      (S (NP-TPC-1 -هُم·-hum·they_[masc.pl.])
        (VP يَخْشَوْنَ·ya+xo$+awona·fear/be_afraid+[masc.pl.]
          (NP-SBJ-1 *T*)
            (SBAR أَنْ·>an·to
              (S (VP تَكُونُ·ta+kuwn+a·it/they/she+be
                (NP-SBJ (NP هذه ha*ihi this)
                  (NP المعادلة AlmuEAdal+ap+u
                    equation))
                (NP-PRD (NP مُدْخَلًا·mudoxal+AF·
                  introduction)
                (PP إِلَى·<ilaY·to/towards
                  (NP تَهْمِيْشِ-
                    ·tahomiy$+i-
                    ·excluding
                    (NP -هُمْ ·their
                      ))))))))))))
                الأردنيون لأنهم يخشون أن تكون هذه المعادلة مدخلا إلى تهميشهم
Al>urduniwun + li>an~a+hum + yaxo$awona + >an + takuwna + ha*ihi + AlmuEAdalapu +
  madoxalAF + <ilaY + tahomiy$i+him
the +Jordanians + for+that + fear (they ) + that + be + this + equation + entry + to +
  alienation+their
The Jordanians because they fear that this equation will mean the beginning of their
  alienation
```

Correct annotation:

The NP and the PP should be 2 separate nodes at the top level because the NP and PP do not belong together as a single constituent.

```
(TOP (NP الازدنيون·Al+>urodun~iy~+uwna·the+Jordanian+[masc.pl.])
  (PP-PRP لـ·li-·for/to
    (SBAR -أَنْ-·->an~a-·that
      (S (NP-TPC-1 -هُم·-hum·they_[masc.pl.])
        (VP يَخْشَوْنَ·ya+xo$+awona·fear/be_afraid+[masc.pl.]
          (NP-SBJ *)
            (SBAR أَنْ·>an·to
              (S (VP تَكُونُ·ta+kuwn+a·it/they/she+be+[sub.]
                (NP-SBJ (NP هذه ha*ihi this)
                  (NP المعادلة AlmuEAdal+ap+u
                    equation))
                (NP-PRD (NP مُدْخَلًا·mudoxal+AF·
                  introduction)
                (PP إِلَى·<ilaY·to/towards
                  (NP تَهْمِيْشِ-
                    ·tahomiy$+i-
                    ·excluding
                    (NP -هُمْ ·him·their
                      ))))))))))))
                .
```


الأردنيون لأنهم يخشون أن تكون هذه المعادلة مدخلا إلى تهمة

Al>urduniwun + li>an~a+hum + yaxo\$awona + >an + takuwna + ha*ihi + AlmuEAdalapu +
madoxalAF + <ilaY + tahomiy\$i+him

the +Jordanians + for+that + fear (they) + that + be + this + equation + entry + to +
alienation+their

*The Jordanians because they fear that this equation will mean the beginning of their
alienation*

Note also that single constituents should not have an additional top node FRAG.

2.3.1.1.6 Node labels not used in ATB

The following node labels are not used in the ATB. These node labels are used in the English Treebank, but do not apply for Arabic. Due to technical issues with the annotation tool, a few labels of this kind exist in earlier releases of the Arabic Treebank; they will not be part of any future release.

RRC (non-verbal reduced relative clause)

SINV (inverted sentence)

NX (also no longer used in the English Treebank)

WHADJP (Wh- adjectival phrase)

2.3.1.2 Types of phrases smaller than clauses

2.3.1.2.1 NP

NP is used for noun phrases.

(S (NP-SBJ zayodN زَيْدُ)
(NP-PRD (NP say~idN سَيِّدُ)
(PP fiy فِي
(NP qawomi قَوْمُ
(NP hi هِ)))))

زيد سيّد في قومه

zayod+N + say~id+N + fiy + qawom+i+hi

Zaiyd + master + in + community + his

Zaid rules in his community

2.3.1.2.2 VP

VP is used for verb phrases – that is to say, the constituent that is headed by the verb and contains all of the verbal arguments.

```
(S (VP reiterate/repeat+they | -Eawaduwa | عاؤدوا - |
  (NP-SBJ *)
  (NP-OBJ occupation | AiHotilAla | إحتلال
    (NP the+earth/land | Al+>aroDi | الأرض ))
  .)
```

عاؤدوا إحتلال الأرض-

Eawaduwa + <iHotilAl+a + Al+>aroD+i
 repeat + occupation + the + land
They occupied the land again

VP is also the head of nominal clauses (S-NOM) and adverbial clauses (S-ADV):

```
(S-NOM (VP -{ixotilAs+i:-{ختلاس:: misappropriation+[def.gen.]
  (NP-SBJ maso&uwliyna::مسؤولين::official/functionary+[masc.pl.gen.])
  (NP-OBJ (NP juzo'+AF::جُزء:: portion/part/fraction+[acc.indef.])
    (PP min::من::from
      (NP (NP h`*A::هَذَا::this_[masc.sg.] )
        (NP (Al+mabolag+i::المبلغ::the+amount/sum+[def.gen.]
          ))))
    ))))
```

إختلاس مسؤولين جزءاً من هذا المبلغ

<ixotilAs+i + maso&uwliyna + juzo'+AF + min + h`*A + Al+mabolag+i
 misappropriating + officials + portion + from + this + the+sum
Officials' misappropriating of a portion of this sum

2.3.1.2.3 PP

PP is used for prepositional phrases.

Prepositional phrases that do not fill one of the more specific adverbial roles that the treebank marks with dashtags (see section ????? about dashtags) should be annotated as PP without a dashtag.

Inside VP:

```
(S (VP -Darab+a-::ضرب::strike/hit+he/it_[verb]
  (NP-SBJ *)
  (NP-OBJ -hA::ها::it/them/her )
  (PP fiy::في::in
    (NP wajoh+i-::وَجْه::face/front+[def.gen.]
      (NP -hA::ها::its/their/her ))))
```

ضربها في وجهها

Darab+a+hA + fiy + wajoh+i+hA
 strike+her + in + face+her
He struck her on the face

Adjoined to NP:

```
(NP (NP فهم
      (NP ه))
  (PP ل
    (NP لقضية)))
```

فهمه للقضية

Understanding his of the issue

His understanding of the issue

NB: need transliteration for the full sentence and in the tree, also needs to have the gloss corrected

Complement of adjective, child of ADJP:

```
(NP (NP quw~+At+K::قُوَاثَ::armed_forces+[fem.pl.]+[indef.gen.] )
  (ADJP tAbiE+ap+K::تَابِعَةٌ::
    subordinate/adherent/belonging_[to]+[fem.sg.]+[indef.gen.]
    (PP li-::ل-::to/for
      (NP (NP -Al+wizAr+ap+i::الْوِزَارَةُ::the+ministry+[fem.sg.]+[def.gen.])
        (NP nafos+i-::نَفْس-::same/self+[def.gen.]
          (NP -hA::-هَا::its/their/her ))))))))
```

قوات تابعة للوزارة نفسها

quw~At+K + tAbiEap+K + li+Al+wizArap+i + nafos+i+Ha

armed forces + belonging + to+the+ministry + same+it

Armed forces belonging to the same ministry

Prepositional phrases that do fill one of the more specific adverbial or complement roles that the treebank marks with dashtags should be annotated with those dashtags (PP with dashtag: -CLR, -PRD, -LOC, -TMP, -PRP, -MNR, -DTV...). See section ????? for a complete description of dashtags.

```
(S wa-::وَ-::and
  (VP ->ad~aY+[null]::أَدَّى-نُلل::direct/guide/lead+he/it_[verb]
    (NP-SBJ *`lika::ذَلِكَ::that )
    (PP-CLR <ilaY::إِلَى::to/towards
      (NP suquwT+i::سُقُوطَ::fall/collapse/crash_[aircraft]+[def.gen.]
        (NP 22::22::nogloss
          (NP qatiyl+AF::قَتِيلًا::casualty/dead_person/
            killed_person+[acc.indef.] ))))))))
    وأدى ذلك الى سقوط 22 قتيلاً
```

wa+>ad~aY+o + *`lika + <ilaY + suquwT+i + 22 + qatiyl+AF

and+led + this + to + fall + 22 + killed

This led to killing 22 people

(VP kAn+a::كان::be/was/were _[verb]
 (NP-SBJ Al+AinotiHArIy~uwna ::الانتحاريون:: the+suicidal/suicide+[masc.pl.nom.])
 (PP-LOC-PRD min::من::from
 (NP Al+mamolak+ap+i::المملكة::the+kingdom+[fem.sg.]+[def.gen.]
 Al+Earabiy~+ap+i::العربية::
 the+Arab/Arabic+[fem.sg.]+[def.gen.]
 Al+saEuwdiy~+ap+i::السعودية::
 the+Saudi+[fem.sg.]+[def.gen.])))
 كان الانتحاريون من المملكة العربية السعودية
 kAn+a + Al+<inotiHArIy~uwna + min + Al+ mamolakap+i + Al+Earabiy~ap+i +
 Al+saEuwdiy~ap+i
 were + the+suicidals + from + the+kingdom + the+Arabian + the+Saudi
The suicidals were from the kingdom of Saudi Arabia

NB: Earlier Arabic Treebank corpora allowed “flat PPs” or “multi-word prepositions,” but there are no longer any multi-word prepositions in the annotation. All such constructions show structure, most often as either NP or SBAR complements of the preposition or as an idafa construction (for those nouns that were formerly treated as prepositions).

2.3.1.2.3.1 List of true prepositions in the Arabic Treebank

Following is a list of the true prepositions in Arabic. These are the only words that may head a prepositional phrase (PP) in the Arabic Treebank.

See also section 4.3.1.7 Prepositions: PREP in the morphological/POS guidelines for details and notes on individual prepositions and usage.

Arabic	Buckwalter	Gloss	Syntactic/ Semantic Context	Example	Transliteration	Gloss	Translation
إلى	<ilaY	Until	It indicates the beginning of a time (synonym of Hat~Y)	مُنْذُ الصَّبَاحِ إِلَى الْمَسَاءِ	mun*u AlSabAHi <iLY AlmasA'i	From the morning <u>until</u> the evening	From morning until evening.
		At/To	It indicates the point of arrival	الوصولُ إِلَى إِثْقَاقِيَّةٍ	AlwSwlu <iLY Ait~ifAqiy~p	Reaching <u>to</u> agreement	Reaching to an agreement
		In addition to	It indicates a meaning of addition	وَ لَكِنَّكَ إِلَى ذَلِكَ عَلَى مَوْعِدٍ مَعَ	wa l'kin~aka <iLY *alika EalY mawEidK mEa...	And but you <u>in addition to</u> that on appointment with	But you are, in addition to that, on an appointment with...

عَنْ	Ean	About	It introduces the topic/subject	هَذَا عَنْ عَلاَقَةِ الْشَّرْطَةِ بِالنَّاسِ	ha*A Ean EalAqapi Al\$urTapi biAlnAsi	This <u>about</u> relation the police with the people	This was about the relation between police and citizens
		From		يَجِبُ أَنْ يَمْتَنِعَ عَنِ التَّدخينِ	yajibu >an yamtaniEa Eani Altadxyni	Must to he depart <u>from</u> the smoking	He must refrain from smoking
فِي	Fiy	In	It indicates location	فِي اسْتِقْنَاءِ أَجْرَتِهِ الْجَمْعِيَّةِ الْأَمْرِيكِيَّةِ	fiy <istiftA'K >ajrathu AljamEay~pu Al>amariykiy~a py	<u>In</u> poll carried out the university the American Society	In a poll carried out by the American Society
		By	It is used in multiplication phrases	ثَلَاثَةٌ فِي ثَلَاثَةٍ تُسَاوِي تِسْعَةً	valAvapN fiy valAvapN tusAwiy tisEapN	Three multiplied <u>by</u> three equals nine	Three times three equals nine.
مِنْ	min	From	It indicates the point of origin be it physical or temporal	مِنْ لَنْدُنْ إِلَى دُبَيِّ	min landun <iLY dubay	<u>From</u> London to Dubai	From London to Dubai
مِ-	mi- (as in mim~an)	Of	It selects a proportion or part from a whole.	ثَمَانُونَ مِلْيَارًا مِنْ الدُولَارَاتِ	vamAnwna milyArAF mina AldwlArAt	Eighty billions from the dollars	Eighty billion of Dollars

- ل	li-	Because of	It indicates causality	لِعَيْبٍ فِيهَا	liEaybK fiyhA.	<u>Because of</u> a defect in it	because of some fault in herself.
		In order to	It introduces the purpose of an event	لِيَحْتِ الْمَوْقِفِ	libaHvi Almawqifi.	<u>In order to</u> examine the situation	In order to examine the situation.
		To/At	It describes the point of arrival	الْوَصُولِ لِلْإِثْقَاقِ	AlwSwlu li<it~ifAqiy~p	Reaching <u>to</u> a deal	Arriving at an agreement.
		For	It introduces a temporal expression	لِأَوَّلِ مَرَّةٍ	li>aw~ali mar~apK	<u>For</u> first time	For the first time
		By	It introduces information on ownership	كِتَابٌ لِمَوْلَانَا مُحِي الدِّينِ الْعَرَبِيِّ	kitAbN limawlAnA muHy Aldiyn AlEaraby	Book <u>by</u> master our muHyi AlDiyn AlEarabiy	A book by our master MuHyi AlDiyn AlEarabiy
- ب	bi-	In/At	It introduces information about location	بِمَدْرَسَةِ الصَّبَاحِيَّةِ الثَّانَوِيَّةِ	bimadraspi AlSabAHiy~ap AlvAnawiy~ap	<u>In</u> school al-Sabahiyya the secondary	At the Sabahia secondary school
		With	Instrument	بِيَدٍ مُرْتَعِشَةٍ	biyadK murtaEi\$apK	<u>With</u> hand shaky	With a shaky hand.

على ^[17]	EalaY	Over/On	It introduces a sense of burden or obligation	عَلَيْنَا أَنْ نَعِيَ ذَلِكَ	EalaynA >an naEiya *alika	<u>On</u> us to be aware that	We must be aware of this
		In	It introduces information about the psychological state	أَنَا عَلَى تِلْكَ الْحَالَةِ مِنَ الْكَابَةِ	>anA EalaY tilka AlHALapi mina Alklbapi	I am <u>in</u> that state of depression	I was in that state of gloom
		At	It introduces information about the location of an event	اثناء جلوسه على مفهى	>avnA'a julwsihi EalY maqhaY	Whilst sitting his <u>at</u> a coffee	Whilst he was sitting at a coffee
		On	It gives information about physical direction	دخلت على الطبيب و قلت له...	daxaltu EalaY AlTabybi wa qultu lahu...	Entered (I) <u>on</u> the physician and said to him...	I went into the doctor and said to him...
		Over/on	It gives information about a time span.	هَلْ حَدَثَ ذَلِكَ كُلُّهُ عَلَى مَدَى خَمْسَةِ وَ عَشْرَ عَاماً	hal Hadava *alika kul~uhu EalaY madaY xamspK wa Ei\$rina EamAF	Has/had happen that all it <u>over</u> twenty five year	Did all of that occur over a period of twenty-five years?
حَتَّى	Hat~aY	Until/up to	It indicates the end point of an action	أَكَلْتُ السَّمَكَةَ حَتَّى رَأْسِهَا	>akaltu Alsamakapa Hat~aY ra>sihA	Ate (I) the fish <u>up to</u> head her	I ate the fish up to its head
كَ-	ka-	As Like	It indicates comparison	لَمْ يَكُنْ مُوَفَّقًا كَمُلْحِنٍ	lam yakun muwaf~aqAF kamulaH~inK	Not was (he) successful <u>as</u> composer	He was not successful as a composer
مُنْذُ ^[18] مُنْذُ	mun*u mu*	Since	It indicates the starting time of an event.	مُنْذُ الْحَمَلَاتِ الْإِنْتِخَابِيَّةِ	mun*u AlHamalAti Al<intixAbiy~ap i	<u>Since</u> the campaign of the elections	Since the election campaigns.

و- [واو القسم]]	wa- [the wa- of oath]	By <By God!>	It is used in oaths	وَ اللَّهِ	wa All~ahi	<u>By</u> God	By God!
ت- [تاء القسم]]	ta- [the ta- of oath]	By (By God!)	It is used in oaths	تَاللَّهِ	taAll~Ahi	<u>By</u> God	By God!
ب- [باء القسم]]	bi- [the ba- of oath]	By (By God!)	It is used in oaths	بِاللَّهِ	biAll~Ahi	<u>By</u> God	By God!
عَدَا	EadA	Except	It marks an exception	أَتَى الْجَمِيعُ عَدَا زَيْدٍ	>atY AljamiyEu EadA zaydK	Came the students <u>except</u> Zaid	The students came except Zayd
حَاشَا	HA\$A	Except	It marks an exception	جاء الطلبةُ حَاشَا زَيْدٍ	jA'a AlTalabapu Ha\$A zaydK.	Came the students <u>except</u> Zaid	The students came except Zayd
خَلَا	xalA	Except	It marks an exception	جاء الطلبةُ خَلَا زَيْدٍ	jA'a AlTalabapu xalA zaydK.	Came the students <u>except</u> Zaid	Everyone came except Zayd
و- [وَإِوَا الْمَعِيَّةِ]	wa- [the wa- of accompaniment]	Together with Along with	It expresses the meaning of accompaniment	جاء الطلبةُ وَ زَيْدًا	jA'a AlTalabapu wa zaydAF	Came the students <u>with</u> Zaid	The students came along with Zayd
إِلَّا [حرف استثناء]	<il~A [The exceptive particle]	Except	It marks an exception	جاء الطلبةُ إِلَّا زَيْدًا	jA'a AlTalabapu <il~A zaydAF	Came the students <u>except</u> Zaid	The students came except Zayd

2.3.1.2.3.2 Specific preposition constructions

The preposition *li* is always annotated as heading a PP. It can have either an NP or an SBAR complement. (nb: In a completely different usage, *li* has a particle homophone, see POS guidelines section 4.3.1.9.13 Jussive Particle: JUS_PART.)

NP complement:

```
(PP li-||for/to ل
  (NP -{inotixAb+At+K|انتخابات|elections/selections+[fem.pl.]+[indef.gen.]
    niyAbiy~+ap+K|نيابية|representative/delegated/deputed+
                                     [fem.sg.]+[indef.gen.]})
    لانتخابات نيابية
```

For elections depute

For depute elections

NB: need transliteration for the full sentence and in the tree, also needs to have the gloss corrected

The SBAR complement of *li* can be headed by *>an~a* or it can have a null complementizer.

```
(PP-PRP li-|ل|for
  (SBAR ->an~a |أَنْ|that
    (S (NP-TPC Al+DuguwT+a|الضغط|the+pressures/stresses)
      (VP ta+tazAyad+u|تزايد|it/they/she+outbid+[ind.]
        (NP-SBJ *T*)))))
    لأنَّ الضغط تزايد
```

For that the pressures increase

Because the pressures increase

NB: need transliteration for the full sentence and in the tree, also needs to have the gloss corrected

```
(S فـfa-·and/so
  (VP -تَقْعُدْ-ta+qoEud+u·it/they/she+sit_down/be_seated/remain+[ind.]
    (NP-SBJ هي·hiya·it/they/she)
    (PP-PRP لـli-·for/to
      (SBAR *0*
        (S (VP -نَقِفْ-na+qif+a·us_to+stop/cease/stand+[sub.]
          (NP-SBJ نَحْنُ·naHonu·we)
          (PP-MNR عَلى·EalaY·on/above
            (NP -أَعْصَابْ-aEoSAb+i-·nerves+[def.gen.]
              (NP -لنا·-nA·our)))))))
    .))
```

فتقعد هي لنقف نحن على أعصابنا

fa + taqoEudu + hya + li+naqifa + naHonu + EalaY + >aEoSAbi+nA

and/so + sit + she + for + stand up + us + on + nerves+our

And she sits down so we stand up all anxious

Note that this is a change from previous versions of the Arabic Treebank, in which this *li* could be annotated as a complementizer as well – this *li* is now always a preposition, and always heads a PP when it has a complement.

2.3.1.2.3.3 Some lexical items can be either preposition or particle depending on context

Some prepositions can have a separate function as a particle.

For example, *min*: As a preposition (with the part-of-speech tag PREP):

```
(NP (NP كثير )
  (PP من
    (NP الناس ) ) )
```

كثير من الناس

Many of the people

Many people

NB: need transliteration for the full sentence and in the tree, also needs to have the gloss corrected

As a particle (with the part-of-speech tag ???????[what kind of PART is this ?]):

```
(S (VP ما
  (ADVP-PRD هناك
    (NP-SBJ (PRT من
      (NP أحد ) ) )
```

ما هناك من أحد

Not there of nobody

There is no one.

NB: need transliteration for the full sentence and in the tree, also needs to have the gloss corrected

```
(S (VP جعل
  (NP-SBJ *)
  (NP-OBJ-1 (PRT من
    (NP هـ )
    (S (NP-SBJ-1 *)
      (NP-PRD زعيما ) ) ) ) )
```

جعل منه زعيما

Made he of him leader

He made a leader of him.

NB: need transliteration for the full sentence and in the tree, also needs to have the gloss corrected

2.3.1.2.3.4 List of nouns formerly treated as prepositions in the Arabic Treebank

Please see section 5.1.1 List of Former Prepositions and/or Adverbs for a list of the words that may have been treated as prepositions (or adverbs) in earlier Arabic Treebank corpora. These are all now analyzed as nouns, and as such they can head NPs, but never PPs.

2.3.1.2.4 ADVP

ADVP is used for adverbial phrases, constituents that are headed by a true adverb.

Adverb phrases that do not fill one of the more specific adverbial roles that the treebank marks with dashtags (see section ????? about dashtags) should be annotated as ADVP without a dashtag.

```
(S (VP waSal+a::وَصَلَ::arrive_at/reach/connect+he/it_[verb]
  (PP-DIR <ilaY::إِلَى::to/towards
    (NP Al+minoTaq+ap+i::الْمِنْطَقَةُ::the+area/zone/
      territory+[fem.sg.]+[def.gen.] ) )
    (ADVP >ayoD+AF::أَيْضًا::also+[acc.indef.] )
    (NP-SBJ (NP Al+maboEuwv+u::الْمُبْعُوثُ::the+envoy/representative+[def.nom.]
      Al+ri}Asiy~+u::الرِّئَاسِيُّ::the+presidential/executive+[def.nom.] )
      (PP li-::لِ-::to/for
        (NP -Āl+salAm+i::-الْسَّلَامُ::the+peace+[def.gen.] ) ) ) ) )
    وصل الى المنطقة أيضا المبعوث الرئاسي للسلام
waSal+a + <ilaY + Al+minoTaq+ap+i + >ayoD+AF + Al+maboEuwv+u + Al+ri}Asiy~+u +
li+Al+salAm+i
arrived + to + the+territory + also + the+envoy + the+executive + for+the+peace
The executive envoy for peace also reached the territory
```

Adverb phrases that do fill one of the more specific adverbial or complement roles that the treebank marks with dashtags should be annotated with those dashtags (ADVP with dashtag: -PRD, -LOC, -TMP, -PRP, -MNR...). See section ????? for a complete description of dashtags.

```
(S (VP وَجَدْتُ wajadotu::find+I
  (NP-SBJ * )
  (NP-OBJ هُ hu:: him )
  (ADVP-LOC هُنَاكُ hunAka:: there ) ) )
```

وجدته هناك

wajadotu+hu + hunAka
found+him + there
I found him there

ADVP can also have the –PRD dashtag when it is the predicate of the clause. A functional dashtag can also be used if appropriate (the –LOC in this example), but the –PRD dashtag must come last:

```
(S wa-::و-::and
  (ADVP-LOC-PRD -hunA::هنا::here )
  (NP-SBJ مقر
    (NP الحزب :::::nogloss )))
```

و هنا مقر الحزب

And here the headquarters of the party.

And here is the headquarters of the party.

NB: need transliteration for the full sentence and in the tree, also needs to have the gloss corrected

ADVP can also occur as the complement of a preposition:

After a preposition, hunA and hunAka are still to be annotated as ADVP.

```
(PP-DIR من
  (ADVP hunA هنا ))
```

```
(PP-DIR إلى
  (ADVP hunAka هناك ))
```

N.B.: Sports scores such as "6-4" in "The Phillies won 6-4" should be done as an NP-ADV with a PP modifier, not as ADVP:

```
(NP-ADV (NP 6 )
  (PP -
    (NP 4 )))
```

Please see section **7.9 Sports Scores** for more examples of sports scores.

2.3.1.2.4.1 List of true adverbs in the Arabic Treebank

Following is a list of the true adverbs in Arabic. These are the only words that may head a adverb phrase (ADVP) in the Arabic Treebank.

See also section 4.3.1.6.1 Regular Adverbs: ADV in the morphological/POS guidelines.

Arabic	Buckwalter	Gloss	Syntactic/Semantic Context	Examples
حينذاك	Hiyna*Aka	At that moment Then	It provides the time of an event	كُنْتُ غَائِبًا حِينَذَاكَ kuntu gA }bAF Hiyna*Aka Was (I) absent <u>then</u> I was absent then
عندذاك	Einda*Aka	At that time	It provides the time of an event	كُنْتُ غَائِبًا عِنْدَذَاكَ kuntu gA }bAF Was (I) absent <u>then</u>

		Then		I was absent then
وَقْتًا	waqta*Aka	At that time Then	It provides the time of an event	كُنْتُ غَائِبًا وَقْتًا kuntu gA}bAF Was (I) absent <u>then</u> I was absent then
آنَآك	lna*Aka	At that time Then	It provides the time of an event	كَانَ الْقَلَمُ آنَآكَ كَالْمُسْتَحِيلِ kAna Alqalamu lna*Aka kAlmustaHiyli Was the pen <u>at that time</u> like the impossible In those times (the use of) the pen was nearly impossible
آنَآذٍ	lna}i*K	At that instant Then	It provides the time of an event	
حِينَئِذٍ	Hiyna}i*K	At that moment Then	It provides the time of an event	حِينَئِذٍ سَيَجْلَى الْحَقُّ Hiyna}*K sayatajal~Y AlHaq~u <u>At that time</u> right will be revealed And then the truth will be revealed
بَعْدَئِذٍ	baEda}i*K	After that time	It provides the time of an event	
عِنْدَئِذٍ	Einda}i*K	At that time Then	It provides the time of an event	عِنْدَئِذٍ سَيَجْلَى الْحَقُّ Einda}*K sayatajal~Y AlHaq~u <u>At that time</u> the truth will be revealed And then the truth will be revealed
وَقْتًا	waqta}i*K	At that time Then	It provides the time of an event	كَانَ يَتَرَدَّدُ عَلَى الْمُسْتَشْفَى وَقْتًا kAna yatarad~adu EalY Almusta\$fY waqta}i*K Was (he) frequently step by on the hospital <u>at that time</u> He used to go regularly to that hospital then
عَمَّا	Eam~a}i*K	At that moment	It provides the time of an event	

عامذاك	EAmA*Aka	In that year	It provides the time of an event	
سَنَّاكَ	sanata*Aka	In that year	It provides the time of an event	
يَوْمِي	yawma}i*K	On that day	It provides the time of an event	
يَوْمَكَ	yawma*Aka	On that day	It provides the time of an event	
مُذَّاكَ	mu*~Aka	From then on	It provides the time of an event	
مُدُّكَ	mud*Aka	From then on	It provides the time of an event	
ثَمَّة ثَمَّ	vam~apa vam~a	There	It specifies the location of an event	ثَمَّة وَقْتُ فَاصِلٍ بَيْنَ الْوُصُولِ وَالْإِقْلَاعِ الثَّانِي vam~apa waqtN fASiN bayna AlwSwli wa Al<iqlAEi AlvAniy <u>There</u> is time separating between the arrival and the departure the second And there is a time separating between arrival and the second departure
هُنَا	hunA	Here	It specifies the location of an event	هُنَا الْحُرِّيَّةُ وَ هُنَاكَ الْإِسْتِعْبَادُ hunA AlHur~iyapu wa hunAka Al<istiEbAdu <u>Here</u> freedom and there slavery <u>Here</u> is freedom and there slavery
هَهُنَا	hahunA	Here	It specifies the location of an event	
هُنَاكَ	hunAka	There	It specifies the location of an event	سَاسَافِرُ هُنَاكَ sa>usAfiru hunAka Will travel <u>there</u> I will travel there
هَنَّاكَ	hunAlika	There		
أَيْضًا	>ayDAF	Also Too	It has the meaning of addition	زَارَهُ أَيْضًا zArahu >ayDaA Visited (he) him <u>also</u> He visited him also
إِذَا	<i*AF	Hence	It expresses the consequence of an event	إِذَا وَ بِكُلِّ الْوُضُوحِ وَالصَّرَاحَةِ نَجَحَ الْإِرْقَابُ فِي قَتْلِ مَوْسَمِ سِيَّاحِي <i*AF wa bikul~I AlwuDuwHi wa

إذن	<i*an	So		AlSarAHapi najaHa Al<irhaAbu fiy qatli mawsimK siyAHiy~K Hence and in all the clarity and the frankness succeeded the terrorism in killing its season the tourist Hence and in all clarity and frankness, terrorism succeeded in destroying tourist session.
رُبَّما	rub~amA	Perhaps	It expresses doubt	رُبَّما يَبْدُو المَوْقِفُ صَعْبًا rub~amA yabduw Almawqifu SEbAF Perhaps seems the situation difficult It may be that the situation is difficult
هَكَذَا	h`ka*A	In this way Like that	It gives the manner with which the action took place	كَانَ واقِفًا هَكَذَا kAna wAqifAF haka*aA was standing in this way
فَقَطْ	faqaT	Only That is all	It adds the nmeaning of restriction	فَقَطْ ارَدْتُ أَنْ اناقِشَ faqaT >aradtu >an >unAqi\$a Only wanted I to discuss I only wanted to discuss
فَحَسْبُ	faHasbu	Only	It adds the meaning of restriction	عَلَيْهِ أَنْ يَفِي بَوَعْدِهِ فَحَسْبُ Ealyhi >an yafiy biwaEdihi faHasbu Has he to keep his promise <u>only</u> He has to keep his promise that is all
قَطْ قَطْ	qaT qaT~u	At all Ever	It reinforces the meaning of negation	مَا سَمِعْنَا بِمِثْلِهِ قَطْ mA samiEnA bimivlihi qaT~u Not heard of the like before <u>ever</u> We have not heard the like of it at all
سَيِّمًا ^[14]	siy~amA	Also		
طالما ^[15]	TalamA	So often		

إِذَاكَ	<i*~Aka	Then		
سَاعَتَيْكَ	sAEata}i*K	That hour/time		
عَامَيْكَ	EAm}i*K	That year		
قَبْلَيْكَ	qabola}i*K	Before then		
لَحْظَتَيْكَ	laHoZata}i*K	That moment		
لَيْلَتَيْكَ	layolata}i*K	That night		

2.3.1.2.5 ADJP

ADJP is used for adjectival phrases consisting of two or more tokens (that is, an adjective with a modifier or a complement).

Adjectival phrases can be the predicate of an equational sentence or clause. When they are the predicate, they are marked with the –PRD dashtag: ADJP-PRD.

```
(S (VP -<in~a:::إِنَّ:::that
  (NP-SBJ Al+lfAq+a:::الآفاق:::the+horizons/provinces+[def.acc.] )\ )
  (ADJP-PRD wADiH+ap+N:::واضحة:::clear/obvious/
    visible+[fem.sg.]+[indef.nom.]
    jid~+AF:::جداً:::very/much+[acc.indef.] )))
```

إنَّ الآفاق واضحة جداً

<in~a + Al+lfAq+a + wADiHap+N + jid~+AF
indeed + the+horizons + visible + very
The Horizons are very visible

Adjectival phrases can also modify noun phrases. When they do, they are adjoined to the NP.

```
(S (VP kAn+at:::كانت:::be/was/were+it/they/she_[verb]
  (NP-SBJ (NP |l+At+u-:::آلات:::instrument/apparatus/
    appliance/machine+[fem.pl.]+[def.nom.]
    (NP -hA:::-ها:::its/their/her ))
  (ADJP Al+Easokariy~+ap+u:::العسكرية:::the+military/
    army+[fem.sg.]+[def.nom.] ))
  (ADJP-PRD mutaTaw~ir+ap+F:::مُتَطَوِّرَةٌ:::developed/
    advanced+[fem.sg.]+[indef.acc.] )))
```

كانت آلاتها العسكرية متطورة

kAnat + l|At+u+hA + Al+Easokariy~+ap+u + mutaTawir~+ap+F
were + machines+its + the+military + advanced
Its military machines were advanced

In some constructions, it is possible for an adjective phrase to be the only child of an NP. This occurs when the ADJP (an adjective with a modifier or a complement) is taking the place of a noun phrase, and in a sense there is a null semantic head of the NP. In these constructions, the ADJP node is used to show the complementation or modification of the adjective (just as with all other uses of ADJP), but the NP node is added above it to show the NP role in the sentence. Frequently this construction involves an ADJP in false idafa (see section ???? for a full description of false idafa).

```
(S (NP-SBJ (NP it/he |-huwa | هُوَ)
  (NP same/self |nafosu- | نَفْسُ-
    (NP its/his |-hu | -هُ)))
  (PP-PRD from |min | مِنْ
    (NP (ADJP senior/important |kibAri | كِبَار
      (NP the+lover/fond_of |AlEA$iqiyana | (العاشقين))))))
    هو نفسه من كبار العاشقين
```

huwa + nafos+u+hu + min + kibAr+i + Al+EA\$iqiyana

he + same+his + from + important + the+lovers

He himself is one of the important lovers

2.3.1.2.6 PRT

PRT is used for all particles.

2.3.1.2.6.1 Particles annotated under constituents smaller than a sentence

All of the particles below are annotated as part of various constituents other than the sentence.

2.3.1.2.6.1.1 Vocative particles (VOC_PART)

They are annotated under an NP-VOC node.

(S (VP اَعْلَم |AiEolamo |know/be_aware+you)
 (NP-SBJ *)
 (NP-VOC (PRT oh/you |yA | يا)
 Aba |>abA | أبا
 Ammar |Eam~Ar | عَمَّار)
 (SBAR that |>an~a | أَنْ
 (S (NP-TPC-4 honor/distinction |\$arafa | شَرَفَ
 (NP the+Jerusalem |Alqudosi | الْقُدْسِ))
 (VP he/it+deny/refuse |ya>obaY | يَأْبَى
 (NP-SBJ-4 *T*)
 (SBAR to |>an | أَنْ
 (S (VP he/it+be_liberated | yataHar~ara | يَتَحَرَّرُ
 (NP-SBJ *)
 (PP-MNR (PRT however/except/other | <il~A | إِلَّا
 on/above |EalaY | عَلَى
 (NP hands |>ayodiy | أَيْدِي
 (NP the+believer |Almu&ominiyna |
 الْمُؤْمِنِينَ))))))))))))
 اعلم يا ابا عمار ان شرف القدس يابى ان يتحرر الا على ايدي المؤمنين
 <iElam+o + yA + >abA + Eam~ar + >an~a + \$araf+a + Al+qudos+i + ya>baY + >an +
 yataHar~a+a + <il~A +EalaY + >ayodiy + Al+mu&miniyn
 be aware + oh + Aba + Ammar + that + honor + the+Jerusalem + refuse + to + be liberated +
 except + on + hands + the+believers
 Aba Ammar, be aware that the honor of Jerusamel refuses to be liberated except by the
 believers

2.3.1.2.6.1.2 Emphatic particle (EMPHATIC_PART)

The emphatic particle لـ (la-/definitely) can cliticize onto different constituents such as verbs, the adverb طالما (TAlamA/so often), prepositional phrases, or adjective phrases. It adds intensity to the token it attaches to. It is always part of the phrase it modifies.

NB: do we have an example of this?

2.3.1.2.6.1.3 Negative particles (NEG_PART)

Most negative particles modify verbs; and therefore, are annotated inside the VP. Exceptions to this are laysa and la which, as negative particles, can also modify NPs, PPs, or any other constituent.

(S (VP (PRT لن::lan::[will]_not/never)
 (NP-SBJ *)
 (PP-CLR ب::bi-::with/by
 (NP الامن::Al+>amoni::the+security/safety))
 (SBAR-ADV حتى::Hat~aY::until/up_to
 (S (VP يرفعوا::ya+rofaE+uwA::they+lift/raise/increase
 (NP-SBJ *)
 (NP-OBJ أيدي::>ayodiy+a-::hands+[def.acc.]
 (NP هم::-hum::their)
 (PP عن::Ean::from/about/of
 (NP أمت::>um~ati-::nation/people
 (NP لنا::-nA::our)))
 لن يَحْلُمُوا بِالْأَمْنِ حَتَّى يَرْفَعُوا أَيْدِيَهُمْ عَنْ أُمَّتِنَا
 lan + yaHlumuW + bi+Al>amni + Hat~Y + yarfaEuW + >aydiy+him + Ean +
 >um~ati+nA

They will not dream of security unless they take out their rule from our nation.

NB: need gloss for the full sentence

Negative particles that modify only specific phrases should be placed inside the modified phrase.

(S (NP-SBJ (NP (PRT لا
 (زيد)
 و
 (NP (PRT لا
 (عمرو)
 (ADJP-PRD قادم)
 .)

لا زيد ولا عمرو قادم.

Not Zayd and not Amr coming

Neither Zayd nor Amr is coming.

NB: need transliteration for the full sentence and in the tree, also needs to have the gloss corrected

2.3.1.2.6.1.4 Future Particle (FUT_PART) and Verb Particles (VERB_PART)

The future particle سَوْفَ (sawfa/will) modifies verbs by adding the feature of future tense to them.

Verb particles are verb modifiers in that they reinforce the perfective aspect of the verb they precede or add to it a meaning of uncertainty. They are always annotated inside the VP.

(S (NP-SBJ هذه::h`*ihi::this/these)
 (NP-PRD (NP ظاهرة::ZAhir+ap+N::phenomenon+[fem.sg.]+[indef.nom.])
 (SBAR (WHNP-1 *0*)
 (S (VP (PRT قد::qad::may/might)
 يسبب::yu+sab~ib+u-::he/it+cause/produce/provoke+[ind.]
 (NP-OBJ (NP ها::-hA::it/them/her)
 (NP-1 *T*))
 (NP-SBJ الخوف::Al+xawof+u::the+fear+[def.nom.]))))))
 .)

وهذه ظاهرة قد يسببها الخوف

wa + h`*ihi + ZAhirap+N + qad + yusab~ib+u+hA + Al+xawof+u
 and + this + phenomenon + may + cause+ it + the + fear
And this is a phenomenon that fear might cause

2.3.1.2.6.1.5 Other Particles

The restrictive particle إلا (<il~A/except for) modifies the constituent it introduces, adding a meaning of restriction and exclusiveness.

(S (NP-TPC-4 سراب::sarAb+a::mirage+[def.acc.]
 (NP التفاوض::Al+tafAwuDi::the+negotiation/consultation))
 (VP (PRT لن::lan::[will]_not/never)
 يصبح::yu+SobiH+a::he/it+become/begin/start+[sub.]
 (NP-SBJ-4 *T*))
 (NP-PRD حقيقة::Haqiyq+ap+F::truth/reality+[fem.sg.]+[indef.acc.])
 (NP-TMP (PRT لا::<il~A::however/except/other)
 بعد::baEoda::after
 (S-NOM (VP عودة::Eawodapi::return
 (NP-SBJ الجولان::Al+jawolAn::the+Golan
 السوري::Al+suwriy~::the+Syrian
 المحتل::Al+muHotal~i::the+occupied)
 (PP-CLR الى::<ilaY::to/towards
 (NP اصحاب::>aSoHAb+i-::owners/originators
 (NP ه::-hi::its/his))))))
 سرابُ التفاوض لن يُصبح حقيقة إلا بعد عودة الجولان السوري المحتل
 sarAbu + AltafAwuDi + lan + yuSbiHa + HaqiyqapF + <il~A + baEda + Eawodapi +
 AljuwlAni + Alsuwriy + AlmuHtal~i
The negotiation mirage will not become real except after the occupied Syrian Golan is back to its owners.

NB: need both gloss and translation here

The particle Hat~Y and >ay modifies any phrase other than VP.

The particle min is always annotated under an NP-SBJ in negated sentences.

The restrictive particle <in~amA is annotated under a phrasal constituent node when it is used in negative constructions. When it is used to focus on the proposition as a whole, it is annotated as a sentential particle.

The restrictive particle *siwY* is always annotated under a PP node.

The particle *rub~a* is always annotated under an NP node: **rub~a** is a particle referring to quantity, and is annotated as follows.

```
(S (NP-SBJ (PRT rub~a رُبُّ)
      Eu*orK عُذْرُ)
  (ADJP-PRD >aqobaHu أَقْبَحُ
    (PP min مِنْ
      (NP *anobK ذَنْبُ))))
```

رب عذر أقبح من ذنب

Rub~a + Eu*orK + >aqobaHu + min + *anobK
 Some (of) + excuse + uglier + of + misdeed
Some excuses are worse than misdeeds (Arabic maxim)

Note that the subject carries the genitive case.

2.3.1.2.6.2 Particles annotated at the sentence level

The particles below are annotated at the sentence level. They are not part of any other constituent in the sentence.

2.3.1.2.6.2.1 Interrogative Particles (INTERROG_PART)

Interrogative particles introduce questions and are annotated under a PRT node inside an SQ.

```
(SQ (PRT does/do?/did?/is/are? |hal | هَلْ)
  (NP-SBJ (NP measures/steps |<ijorA'Atu | إجراءاتُ
    (NP the+government/administration |AlHukuwmap_i | (الحُكُومَةُ))
    (ADJP the+taken/adopted |Almut~axa*api | (الْمُتَّخَذَةُ))
    (ADJP-PRD guaranteeing |kafiylapN | كَفِيلَةٌ
      (PP by/with |bi- | بِ-
        (NP treatment/therapy/processing | -muEAlajapi | مُعَالَجَةٍ -
          (NP the+conditions/situation/status |
            Al>awoDAEi | الأَوْضَاعُ
              the+economic/economical |
                AlAiqotiSAdiy~api | (الاِقْتِصَادِيَّةُ))))))
    هل اجراءات الحكومة المتخذة كفيلة بمعالجة الاوضاع الاقتصادية
  hal + |<ijorA'At+u + Al+Hukuwmap+i + Al+mut~axa*ap+i + kafiylap+N bi+muEAlajap+i +
    Al+>awDAE+i + the+taken + Al+<iqtiSAdiy~ap+i
  does/do?/did?/is/are? + measures + the government + guarantee + with + treatment +
    the+situation + the+ economic
  Do the measures taken by the government guarantee the treatment for the economic
    situations?
```

2.3.1.2.6.2.2 Focus Particle (FOCUS_PART)

The focus particle, أَمَا (>am~A/as for), is annotated at the sentence level.

```
(S (PRT اَلْمَا::>am~A::as_for/concerning)
  (NP-TPC-1 اَلْاِسْلَاطَةُ::Al+suloTapu::the+power/authority/rule)
  ::,
  (PRT فَ::fa-::and/so)
  (VP لَيْسَتْ::layos+at::not_be+it/they/she_[verb]
    (NP-SBJ-1 *T*)
    (ADJP-PRD مُسَالِمَةٌ::msAlmp::nogloss)))
```

أَمَا السُّلْطَةُ فَلَيْسَتْ مُسَالِمَةٌ

>am~A + AlsulTapu + fa+layosat + musAlimapF.

As for the government, then it was not pacifist.

NB: need both gloss and translation here

2.3.1.2.6.2.3 Other Particles

The restrictive particle إِثْمًا (<in~amA/indeed) emphasizes the proposition it introduces as compared to another one.

ثُمَّ (la-/then) is a response conditional particle (RC_PART) that is always cliticized onto verbs.

The response conditional particle فَ (fa-/then) (RC_PART) introduces the main clause that expresses the result of a condition expressed in the subordinate clause.

The particle أَلَا >alA emphasizes the meaning of the command verb that it modifies.

The particles إِذْ (<i*/suddenly) and إِذَا (<i*A/suddenly), introduce a proposition that expresses a sudden action occurring in the meantime of another action.

The particle مَا (mA) is annotated at the sentence level following an adjective in the indefinite accusative and in the position of an NP-ADV.

The particle كَأَنَّمَا (ka>an~amA/as if) modifies the proposition it introduces.

```
(S (SBAR-ADV اِنْ::<in::if/whether
  (S (S (VP (PRT لَمْ::lam::did_not)
    يَجْمَعُ::yu+jomaE+o::he/it+be_agreed_unanimously
    (NP-SBJ اَلْعَرَبُ::Al+Earabu::the+Arabs)))
    و::wa-::and
    (S (VP (PRT لَمْ::lam::did_not)
      يُوَحِّدُوْا::yu+waH~id+uwA::they+unite/unify
      (NP-SBJ *)
      (NP-OBJ كَلِمَةٌ::kalim+at+a-::word/remark/speech
        (NP هُمْ::hum::their)))
      ,
      (PRT فَ::fa-::and/so)
      (NP-TPC-2 اَلْقَضِيَّةُ::Al+qaDiy~apu::the+problem/issue
        اَلْفِلَسْطِيْنِيَّةُ::Al+filasoTiyiny~apu::the+Palestinian)
```

(PP-PRD ل::la-::to/for
 (NP (NP ها:::-hA::it/them/her_[it/she_has,_they_have])
 (NP-2 *T*))))
 (NP-SBJ عدة::Eid~apu::several/numerous/many
 (NP اوجه::>awojuh::nogloss)))
 إن لم يُجمع العربُ ولم يُوحَّدوا كَلِمَتُهُمُ فَالْقَضِيَّةُ الْفِلِسْطِينِيَّةُ لَهَا عِدَّةٌ أَوْجُهٌ
 <in + lam + yujomiEo + AlEarabu + wa+lam + yuwaH~iduwA + kalimatahum
 fa+AlqaDiy~apu + AlfilisTyniy~apu + la+hA + Eid~apu + >awojuhK
If Arabs do not agree and unite their voices then the Palestinian problem will have several facets.

NB: need both gloss and translation here

(S (NP-TPC-2 (NP السلطات::Al+suluTAtu::the+authorities
 النقدية::Al+naqodiy~apu::the+monetary/cash)
 (PP في::fiy::in
 (NP (NP عدد::EadadK::number/quantity/issue)
 (PP من::min::from
 (NP الدول::Al+duwali::the+states/countries))))))))
 (NP-ADV غالبا::gAlibAF::generally/commonly/typically
 (PRT ما::mA)
 (VP تعتمد::ta+Eotamid+u::it/they/she+depend/rely_[on]+[ind.]
 (NP-SBJ-2 (-NONE- *T*))
 (NP-OBJ (NP اسعار::>asoEAra::prices/rates
 (NP فائدة::fA{id+ap+K::benefit/use)
 (ADJP مرتفعة::murotafiEapK::elevated/rising))
 (PP-PRP ل::li-::to/for
 (NP (NP لدفاع::-Al+difAE+i::the+defense+[def.gen.])
 (PP عن::Ean::from/about/of
 (NP اسعار::>asoEAri::prices/rates
 (NP الصرف::Al+Sarofi::the+diverting/spending))))))))))
 السلطات النقدية في عدد من الدول غالبا ما تعتمد اسعار فائدة مرتفعة للدفاع عن اسعار الصرف.
 AlsuluTAtu + Alnaqdy~apu + fiy + EadadK + min + Alduali + gAlibAF + mA + taEotamidu
 + AsoEAra + fA{idapK + murotafiEapF + li+AldifAEi + Ean + AsoEAri + AlSarfi
It is often that monetary institutions in a number of countries employ high interest rates to protect stock market prices.

NB: need both gloss and translation here

2.3.1.2.7 WHNP

WHNP is used for nominal question words or nominal relative pronouns (nominal WH- words).

See list of REL_PRON and INTEROG_PRON in section ??? about PRON in the POS guidelines

(PP from |min | من
 (NP (NP the+numbness/indifference/confusion |Al*~uhuwli | الذُّهُولِ
 (SBAR (WHNP-3 which/who/whom |Al~a*iy | الَّذِي
 (S (VP raise/provoke/agitate+it/they/she |>avArat- | أَثَارَتِ
 (NP-SBJ *)
 (NP-OBJ (NP it/him |-hu | -هُ)
 (NP-SBJ-3 *T*))))))
 من الذُّهُولِ الَّذِي أَثَارَتِهِ

min + Al+*uhuwli+i + Al~a*iy + >avArat+hu
 from + the+confusion + which + generate+it
Because of the confusion it generated

The WH- word can also be phonetically null – these are shown as zero (WHNP *0*):

(S (VP (PRT indeed_ |-laqad | لَقَدْ
 hear/listen+we |samiEonA | سَمِعْنَا
 (NP-SBJ *)
 (NP-OBJ (NP speech/statement/remark/talk/saying |kalAmAF | كَلَاماً
 encouraging |mu\$aj~iEAF | مُشَجَّعاً
 (SBAR (WHNP-4 *0*)
 (S (VP give/provide+he/it |>aEoTaY | أَعْطَى
 (NP-SBJ-4 *T*)
 (NP-OBJ hopes/wishes | |mAlAF | آمالاً
 large/great/important/major/senior
 |kabiyrapF | كَبِيرَةً))))))
 لقد سمعنا كلاماً مشجعاً أعطى آمالاً كبيرةً

laqad + samiEonA + kalAm+AF + mu\$aj~iE+AF + >aETaY + |mAl+F + kabiyrap+F
 indeed + heard + speech + encouraging + gave + hopes + great
We have heard some encouraging speech which gave great hopes

2.3.1.2.8 WHADVP

WHADVP is used for adverbial WH- words.

See list of REL_ADV in the POS guidelines

(S fa-::ف-::and/so
 (SBAR-ADV (WHADVP-2 -kul~amA::كُلَّمَا::whenever)
 (S (VP {ibotaEad+nA::{بِتَعَدْنَا::move_away/eschew +we_[verb]
 (NP-SBJ *)
 (PP-CLR Ean::عَنْ::from/about/of
 (NP (NP Al+difAE+i::الدِّفَاع::the+defense)
 (PP Ean::عَنْ::from/about/of
 (NP Huquwq+i-::حُقُوق-::rights/law
 (NP -nA::-نَا::our))))
 (ADVP-TMP-2 *T*)))))
 (VP {iqotarab+nA::{قَتَرَبْنَا::get_close_to/approach+we_[verb]
 (NP-SBJ *)
 (PP-CLR min::مِنْ::from
 (NP Al+HAj+ap+i::الْحَاجَّةُ::the+need+[fem.sg.]+[def.gen.]
 wa-::وَ-::and
 -Al+xawof+i::الْخَوْفِ::the+fear+[def.gen.]))))
 فكلما ابتعدنا عن لِدفاع عَنْ حقوقنا اقتربنا من الحاجة و الخوف
 fa + kul~amA + AibotaEadonA + Ean + Huquwqi+nA + AiqotarabonA + min + AlHajapi +
 wa + Alxawofi
 and + whenever + stay away+we + from + defense + from + rights+our + get close+we +
 from + the+need + and + the+fear
And whenever we abandon fighting for our rights, we are more exposed to need and fear

(S wa-::وَ-::and
 (VP -waSal+a::وَصَلَ::arrive_at/reach/connect+he/it_[verb]
 (NP-SBJ t\$iyiniy::تَشِينِي:Cheney)
 (PP-DIR <ilaY::إِلَى::to/towards
 (NP (NP saloTan+ap+i::سُلْطَنَة::sultanate+[fem.sg.]+[def.gen.]
 (NP EumAn+a::عُمَان::Oman+[indef.gen.]))
 (SBAR (WHADVP-1 Hayovu::حَيْثُ::where/whereby /given_that)
 (S (VP {ilotaqaY+[null]::{لِتَقَى[نُلل]::
 meet/encounter+he/it_[verb]
 (NP-SBJ *)
 (NP-OBJ Al+suloTAn+a::السُّلْطَانُ::the+Sultan+[def.acc.]
 qAbws::قابوس::NOT_IN_LEXICON)
 (ADVP-LOC-1 *T*))))))))
 ووصل تشيني الى سلطنة عمان حيث إلتقى السلطان قابوس
 wa+waSal+a + t\$iyiniy + <ilaY + saloTanap+i + EumAn+a + Hayovu + <ilotaqaY +
 Al+suloTAn+a + qAbuws
 and+arrived + Cheney + to + Sultanate + Oman + where + met + Sultan + Kaboos
Cheney arrived in Oman Sultanate where he met Sultan Kaboos

NB. When the WHADVP متى/ mataY /when is heading an SBAR-TMP, there are two levels of modification : internal (i.e. within the SBAR-ADV) and external (i.e. at the level of the main clause).

Examples:

1. متى جعت تناول غداءك

When be hungry you, get you lunch your
 When you get hungry have lunch.

NB: need transliteration

2. متى أمطرت لا أريد الخروج

When it rained it, don't like I to go outside

When it rains, I don't like going out.

NB: need transliteration

3. سأكتب الدرس متى فهمته

Will I write the lesson when understand I him

I'll write the lesson when I understood it.

NB: need transliteration

Each sentence has two events:

1. a "be hungry" event and a "get lunch" event
2. a "raining" event and a "like to go" event
3. a "write" event and a "understand" event

In all of these examples, BOTH events have temporal modifiers. The temporal modifiers for the two events are related (they're close in time to each other, or they might refer to each other), but they are still two separate time modifiers.

Because the two different modifiers are modifying two different events, they are both marked. In these examples, they are both marked –TMP. And they both have to be there – you can't leave out the trace in the subordinate clause, or else the event in that clause will be lonely because it's missing its modifier. We need to know when the understanding happened just as much as we need to know when the writing happened.

The temporal modifier for the subordinate clause (the "mataY" clauses in these examples) is the trace of "mataY":

```
(SBAR-TMP (WHADVP-1 متى
  (S (VP فهمت
    (NP-SBJ *)
    (NP-OBJ له)
    (ADVP-TMP *T*)) ) ) )
```

← event=understand

← temporal modifier of "understand" event

NB: need transliteration for the full sentence and in the tree, also needs to have the gloss and the translation, and the Arabic sentence

The temporal modifier for the main clause is the subordinate clause itself:

```
(S (VP (PRT سأكتب
  (NP-SBJ *)
  (NP-OBJ الدرس )
  (SBAR-TMP (WHADVP-1 متى
    (S (VP فهمت
      (NP-SBJ *)
      (NP-OBJ له)
      (ADVP-TMP *T*)) ) ) ) ) )
```

← event=write

← temporal modifier of "write" event

(ADVP-TMP *T*)))))

سأكتب الدرس متى فهمته

Will I write the lesson when understand I it

NB: need transliteration for the full sentence and in the tree, also needs the translation

```
(S (VP (PRT سأكتب
      (NP-SBJ *)
      (NP-OBJ الدرس )
      (SBAR-TMP (WHADVP-1 متى
                  (S (VP فهمت
                      (NP-SBJ *)
                      (NP-OBJ له)
                      (ADVP-TMP *T*) ) ) ) ) ) ) ) )
```

NB: need transliteration for the full sentence and in the tree, also needs to have the gloss and the translation, and the Arabic sentence

2.3.1.2.9 WHPP

WHPP is used for prepositional question or relative clauses. These will be PPs that have WHNP complements. WHPP is a PP headed by a PREP which governs a WHNP.

```
(SBARQ (WHPP-1 for/to |li- | ل-
        (WHNP who/whom | -man | -من))
  (S (VP I+withdraw/pull_out |>anosaHibu | أنسحب
      (NP-SBJ *)
      (PP-1 *T*) ) )
  ?)
```

لمن أنسحب

li+man + >anosaHib+u ?
for+who + withdraw ?
for whom do I withdraw ?

```
(SBARQ (WHPP-1 على
        (WHNP (WHNP اي )
              (PP من
                (NP هم ) ) ) ) )
  (S (VP وقع
      (NP-SBJ الاختيار )
      (PP-CLR-1 *T*) ) ) )
```

على ايّ منهم وقع الاختيار؟

On which of them came/ fell the choice
Which one was chosen?

NB: need transliteration for the full sentence and in the tree, also needs to have the gloss checked

2.3.1.2.10 CONJP

CONJP is used for multi-word conjunctions (usually 'wa' with another CONJ).

```
(S (VP they_+want/desire/intend |yuriyduwna- | يُرِيدُونَ-
  (NP-SBJ *)
  (NP-OBJ me |-niy | -ني)
  (SBAR to |>an | أن
    (S (VP I+be |>akuwna | أَكُونُ
      (NP-SBJ *)
      (NP-PRD either |<im~A | إِمَّا
        (NP prisoner/captive |>asiyrAF | أَسِيرًا)
        (CONJP and |wa- | وَ-
          either |-<im~A | -إِمَّا)
        (NP exiled/evicted/pariah |TariydAF | طَرِيدًا))))))
    يريدونني أن أكون إما أسيرًا وإما طريدًا
```

yuriyd+uwn+a+niy + >an + >akuwn+a + <im~A + >asiyr+AF + wa+<im~A + TariydAF +
 wanted+they+me + that + were + either + prisoner + and+either + exiled + and+either +
 killed

They want me to be either a prisoner of war or exiled or killed

```
(S (S (VP (PRT lam لَمْ
  ya>oti يَأْتِ
  (NP-SBJ zayodN زَيْدٌ )
  (PP-TMP <ilaY إِلَى
    (NP Al|Ana الآن ) )
  (CONJP وَ -::and
    -لْكِين::but )
  (S (NP-SBJ huwa هُوَ )
    (PP-PRD fiy فِي
      (NP AlTariqi الطَّرِيقِ ) ) ) ) ) )
```

لم يأت زيد إلى الآن ولكن هو في الطريق

lam + ya>oti+o + zayod+N + <ilaY + Al|Ana + wa+l'akino + huwa + fiy + Al+Tariqi+i
 not + come + Zaid + till + now + and+but + he + in + the+way
Zaid has not come yet but he is in his way

Note that iDafatAF ilAY, and Ewa EAlay should not be annotated as CONJP. These words tagged formerly in ATB as CONJP should now be NP-ADV.

2.3.1.3 Other types

2.3.1.3.1 PRN

PRN is used for true sentential parentheticals. It will no longer be used for instances of syntactically integrated adjunction or modification (even if parentheses punctuation marks are used, as in this sentence).

PRN should be attached as high as possible in the tree, just like commas. Parentheticals are considered as PRN if they occur in the middle of the sentence. If they are at the end of the sentence, they are annotated as a new coordinated S except for discourse fillers (c.f. [2.3.1.3.1.2.1.1](#)).

The domain of instances where the PRN node is used has become smaller. Formerly, any constituent or group of constituents bounded by brackets was enclosed in a PRN node. Similarly, other punctuation (mainly commas and dashes) could set off a PRN node if an annotator decided that a span was an aside or otherwise not part of the rest of the sentence.

Now, PRN is to be used for "sentential interpolations". **This means that every PRN should dominate some flavor of an S-node (with certain exceptions outlined below).**

What this means for annotators is that most PRN nodes have to be removed (or not put in to begin with), and whatever constituents were beneath it must be integrated into the rest of the sentence, which often entails adding structure. For example,

Incorrect use of PRN:

```
(NP زياد·ziyAd·Ziad
  علي·Ealiy~·Ali
  حسن·Hasan·Hassan
  قواسمة·qwAsmp·NOT_IN_LEXICON)
  (PRN -LRB-·
    (NP 33·
      (NP سنّة·san+ap+F·year+[fem.sg.]+[indef.acc.] )
    -LRB-· ) )
```

زياد علي حسن قواسمة (33 سنة)

ziyAd+ Ealiy + Hasan + qawAsomap + (33 + sanapF)

Zyad + Ali + Hasan + Qawasma + (33 + years)

Zyad Ali Hasan Qawasma (33 years)

Correct annotation:

```
(NP (NP زياد·ziyAd·Ziad
  علي·Ealiy~·Ali
  حسن·Hasan·Hassan
  قواسمة·qwAsmp·NOT_IN_LEXICON)
  (NP -LRB-·
    33·
    (NP سنّة·san+ap+F·year+[fem.sg.]+[indef.acc.] )
    -LRB-· ) )
```

زياد علي حسن قواسمة (33 سنة)

ziyAd+ Ealiy + Hasan + qawAsomap + (33 + sanapF)

Zyad + Ali + Hasan + Qawasma + (33 + years)

Zyad Ali Hasan Qawasma (33 years)

with the PRN removed and an NP node inserted. Note that the parentheses brackets should be paired and attached low. That is, they should still be inside the constituent that was originally marked off by the PRN.

2.3.1.3.1.1 Parentheticals where PRN should be used

The instances where the usage of PRN is still sanctioned can be divided into two categories: sentential interpolations and train wrecks.

2.3.1.3.1.1.1 Sentential interpolations:

This is when a complete sentence S that is syntactically independent is inserted into a larger S context.

In Arabic, typical parenthetical clauses would be, for example,

1. Constructions which praise or vilify someone such as expressions like الحمد لله AlHamodu lil'ah (Thank God!), لا سمح الله lA samHa All'ah (God forbid!), etc.
2. Inserted sentential comments or additional information. يعني 'YaEoniy' (which means), in speech data.
3. The sentential modification of nouns starting with هو wa huwa, or هي wa hya (and + personal pronoun), when it comes in the middle of a main sentence.

This use of PRN will dominate a complete clausal (usually S) node.

```

(S (VP -إِنَّ·-<in~a·indeed
  (NP-SBJ (NP (NP أَيُّ·>ay~+a·any+[def.acc.]
    (NP سُوءٌ·suw'+K·offense/bad/ill/miss-+[indef.gen.])))
  (SBAR (WHNP-4 *0*)
    (S (VP (PRT قد·qad·may/might)
      يُطَاوِلُ·yu+TAwil+u·he/it+contend_with+[ind.]
      (NP-SBJ-4 *T*)
      (PRN (S (VP (PRT لَا·lA·no/not/non-)
        سَمَحَ·samaH+a·allow/permit+the
        (NP-SBJ اللَّهُ·All~`h+u·
          Allah/God)))))
      (NP-OBJ أَحَدٌ·>aHad+a·one+[def.acc.]
        (NP الزُّمَلَاءُ·Al+zumalA'+i
          the+colleagues
          (NP الْمُهَنْدِسِينَ·Al+muhanodis+iyna·
            the+engineer))))))
    (NP هُوَ·huwa·it/he))
  (PP-PRD بِ·bi-·by/with
    (NP -مِثَابَةً·-mavAb+ap+i·virtually/tantamount_to
      (NP (NP {عُدَاءٌ·{iEotidA'+K·attack/aggression/assault)
        (PP عَلَى·EalaY·on/above
          (NP (NP الْجِسْمِ·Al+jisom+i·the+body/form+[def.gen.]
            الْمُهَنْدَسِيِّ·Al+hanodasiy~+i·the+engineering
            اللَّبْنَانِيِّ·Al+lubonAniy~+i·the+Lebanese)
          (PP بِ·bi-·by/with
            (NP (ADJP -كَامِلٌ·-kAmil+i-·complete/full
              (NP -هِ·-hi·its/his))))))))))
    .))

```

إِنَّ أَيُّ سَوْءٍ قَدْ يَطَاوِلُ لَا سَمَحَ اللَّهُ أَحَدَ الزُّمَلَاءِ الْمُهَنْدِسِينَ هُوَ بِمِثَابَةِ اعْتِدَاءٍ عَلَى الْجِسْمِ الْهَنْدَسِيِّ الْلَبْنَانِيِّ بِكَامِلِهِ
 <in~a + >ay~a + suw}K + qad + yuTAwilu + lA + samHa + All~`hu + >aHada +
 AlzumalA'i + Almuhandisiyna + huwa + bi+mavAbapi + <iEotidA'K + EalaY +
 Aljisomi + Al+handasy~i Al+lubnAniy~i + bi+kAmili+hi
 Indeed + any + harm + may + reach + not + allow + God + some + the+colleagues +
 the+engineeres + it/he + with+tantamount + aggression + on + body + the+engineering +
 the+Libanese + with+its/his + totality
Any harm that reaches, God forbid, one of our colleague engineers, is equal to a harm to all engineers in Lebanon.

Also treated as parenthetical in ATB, are collocations like رَحِمَهُ اللَّهُ RaHimahu Al`ah (God saves his soul), عَلَيْهِ السَّلَامُ Ealayohi Alsalamu (Peace be upon him), صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ Sal~aY Al`ahu Ealayohi wa sal~ama that modify NPs. These collocations are treated as PRN even if they fall at the end of a sentence.

NB: correct all PRN examples (throughout whole guidelines) – put PRN at highest possible level, like commas

Examples:

```

(S (NP-SBJ الْحَبِّ need transliteration and gloss here)
  (PRN (S (VP أَعَزَّ need transliteration and gloss here
    (NP-OBJ ك need transliteration and gloss here)
    (NP-SBJ اللَّهُ need transliteration and gloss here)

```

(NP-PRD هزل
ثم
جدّ need transliteration and gloss here))

الحبّ أعزّك الله هزل ثمّ جدّ

The love strengthen you God fun then seriousness.

Love (may you be cherished by God) is fun then things becomes more serious.

Need transliteration

(S (VP qAla قال gloss here
(NP-SBJ *)
(PRN (S (VP taEAlAY تعالّ gloss here
(NP-SBJ *)))
(S : "wa
(PP-PRD bi ب
(NP Al+wAlid+ayoni ((الوالدين
(NP-SBJ احسانا))))

قال تعالى: "و بالوالدين احسانا".

Said be exalted :” and with parents integrity”

God (may he be exalted) said: “and you will be dutiful to your parents”

Need transliteration

The collocations below are always annotated as PRN, even when they come at the end of the sentence:

Approve God on him.
May God approve of him.

رضي الله عنه

Venerate God his face.
May He be venerated

كرّم الله وجهه

Have mercy on him God
May God have mercy on him.

رحمه الله

Lengthen Godlife his
May God lengthen his life.

أطال الله عمره

Insulted him God
May God curse him.

لعنه الله

Not allow God
God forbid.

لا سمح الله

Not appraise God.
God forbid.

لا قدر الله

(S (VP ذهب need transliteration and gloss here
 (NP-SBJ *)
 (PP إلى ·<ilaY·to/towards
 (NP مُنْتَدَى ·munotadaY·conference/fair/assembly
 (NP الدُّكْتُور ·Al+dukotuwri·the+doctor/Doctor/Dr.+[def.gen.]
 جَمَال ·jamAl·Jamal/Gamal
 الأتاسي ·AlAtAsy·NOT_IN_LEXICON)))
 (PRN -
 (S (VP رَحِمَ ·raHim+a·have_mercy_with+he/it
 (NP-OBJ -هُ ·-hu·it/him)
 (NP-SBJ اللَّهُ ·All~`h+u·Allah/God))))
 ذهب إلى منتدى الدكتور جمال الأتاسي رحمه الله
 <ilaY + munotadaY + Al+dukotuwri + jamAl + Al>atAsiy + Rahima+hu + All~`hu
 Went he to + conference + the+doctor + Jamel + Alatassi + have mercy with + him + God
 He wento the conference of Dr. Jamal Alatassi, God rests his soul

A PRN node can contain more than one sentence in coordination; these sentences are coordinated under the same PRN node.

.... قال الله جلّ وعلا
 Said God be majestic and be elevated....
 God (May he be above everything) said....

(S (VP قال need Arabic and gloss here
 (NP-SBJ *)
 (PRN (S (S (VP صَلَّى need Arabic and gloss here
 (NP-SBJ Al~Ah+u الله need Arabic and gloss here)
 (PP-CLR علي need Arabic and gloss here
 (NP hi ه need Arabic and gloss here))))
 wa و need Arabic and gloss here
 (S (VP سَلَّمَ need Arabic and gloss here
 (NP-SBJ *)))))
 (S.....)))

قال صلى الله عليه وسلم.....

Said pray God on him and greet
 He (Peace be on him) said....

(NP (NP الإمام ·Al+<imAm+i·the+leader/imam+[def.gen.]
 الْكَبِيرَ ·Al+kabiyri·the+large/great/important/major/senior+[def.gen.]
 (NP عَلِيَّ ·Ealiy~·Ali
 بن ·bin·bin
 أَبِي ·>abiy·Abi طالب ·Talib·Talib/Taleb
 (PRN (S (PP-PRD عَلَيَّ ·Ealay~·on/above
 (NP -هُ ·-hi·it/him)
 (NP-SBJ السَّلَام ·Al+salAm+u·the+salute/salutation+[def.nom.])))
 الإمام الكبير علي بن أبي طالب عليه السلام
 Al<imAmu + Al+kabiyru + Ali + <ibn + Abiy + TALib + Ealayo+hi + Al+salAmu
 The+Imam + the+great + Ali + Ibn + Abi + Taleb + on+him + the+peace
 The Imam Ali Ibn Abi Taleb, peace be upon him.

Note that the collocation **إن شاء الله** <in \$A'a Al'ah (if God wills) is always treated as an SBAR-ADV and **not parenthetical**.

```
(S (S (VP (PRT لَقَدْ·laqad·indeed_[has/have])
  وَقَعْنَا·waq~aE+nA·sign+we_[verb]
  (NP-SBJ *)
  (NP-OBJ (NP {تَفَاقٌ·{it~ifAq+a·agreement/accord/treaty+[def.acc.]
    (NP الشَّرَاكَةِ·Al+$irAk+ap+i·the+partnership))
    (NP-ADV مَعَ·maE+a·with+[def.acc.]
      (NP {تَحَادٍ·Al+{it~iHAd+i·the+union+[def.gen.]
        الأوروپيَّ·Al+>uwruwb~iy~i·the+European}))
      (PP-MNR بِـ·bi-·by/with
        (NP {الأَحْرُفِ-·Al+>aHoruf+i·the+letters+[def.gen.]
          الأولى·Al+>uwlaY·the+first})))
  وَ-·wa-·and
  (S (SBAR-ADV -إن-·-<in·if/whether
    (S (VP شَاءَ·$A'+a·want/desire/wish+he/it_[verb]
      (NP-SBJ هُوَ·All~h+u·Allah/God+[def.nom.])))
    (NP-TMP بَعْدَ·baEod+a·after+[def.acc.]
      (NP ثَلَاثَةَ·valAv+ap+i·three+[fem.sg.]+[def.gen.]
        (NP أَشْهُرَ·>a$ohur+K·months+[indef.gen.])))
    (VP يُوقَعُ·yu+waq~aE+u·he/it+be_signed+[ind.]
      (NP-SBJ-3 *)
      (NP-OBJ-3 *)
      (NP-MNR نِهَائِيًا·nihA}iy~+AF·finally/in_the_end+[acc.indef.])))
    لقد وقعنا اتفاق الشراكة مع الاتحاد الأوروبي بالأحرف الأولى وإن شاء الله بعد ثلاثة أشهر يوقع نهائياً
```

Laqad + waq~aEonA + <it~ifAq + Al+\$arAkapi + maEa + Al+<it~ihAdi +
 Al+>uwruwb~iy~i + bi+Al+>aHorufi + Al+>uwlaY + wa+>ino + \$A'a + All~Ahu +
 baEoda + valAvapi + >a\$ohurK + yuwaq~aEu + nihA}iy~AF
*We have signed the partnership treaty with the European Union with initials and if God wills
 it will be fully signed in three months*

If this SBAR-ADV comes in a position where it interrupts an uninterruptable constituent (an idhafa, for example), it should still be put at the S or VP level, but the remainder of the interrupted constituent needs to be ICH-traced to its original position.

```
(S (NP-SBJ (NP المذاهب
  (الاربعة)
  (NP هي))
  (PP-PRD في
    (NP ظلال
      (NP-1 *ICH*)))
  (SBAR-ADV ان
    (S (VP شاء
      (NP-SBJ الله))))
  (NP-1 القرآن و السنة))
  المذاهب الاربعة هي في ظلال ان شاء الله القرآن و السنة
```

The trends the four she in shadows if allow God the Coran and the Sunnah.

The four trends are following the Coran and the Sunah if God wills it.

NB: need transliteration here and in the tree

PRN also is used for parentheticals (marked in the text with parentheses, or –LRB- and –RRB-) that contain sentential meta-comments and additional information that is external to the matrix sentence.

```
(S (VP ضَرَبَ.Darab+a.strike/hit+he/it_[verb]
  (NP-SBJ *)
  (NP-OBJ الرِّقْمَ.Al+raqom+a.the+number
    القياسي.Al+qiyAsiy~+a.the+record/)
  (PP-MNR (PP بـ-bi-.by/with
    (NP -الحجم-.Al+Hajom+i.the+size))
    (PRN -LRB-
      (S (NP-SBJ (ADJP آخِرُ|xir+u.last
        (NP لَوَحَاتٍ painting
          (NP -هِ-hi.his)))
        (NP-PRD 200.
          (NP مِثْرٍ.mitor+K.meter
            مُرَبَّعٍ.square )))
        -RRB-))
      و-wa-.and
      (PP -بـ-bi-.by/with
        (NP -السرعة-.Al+suroE+ap+i.the+speed))
    (S (VP نَفَّذَ.accomplish+he
      (NP-SBJ *)
      (NP-TMP الصَّيْفَ.the+summer
        الماضي.the+past)
      (NP-OBJ (NP لَوْحَةٌ.painting
        عملاقة.gigantic)
        (PP مِن.min.from
          (NP عَشْرَةٌ.ten
            (NP أمتار.meters
              مُرَبَّعَةٍ.square))))
        (PP-TMP في.fiy.in
          (NP (ADJP أَقَلُّ.>aqal~+a.less
            (PP مِن.min.from
              (NP سَاعَةٌ.hour))))))
        ضرب الرقم القياسي بالحجم (آخر لوحاته 200 متر مربع) و بالسرعة (نفذ الصيف الماضي لوحة
          عملاقة من 10 أمتار مربعة في أقل من ساعة)
      Daraba + Al+raqoma + Al+qiyAsiy~a + bi+Al+Hajomi + lxiru + lawoHAti+hi + 200 +
        mitorK + murab~aEK + wa+bi+Al+suroEapi + naf~a*a + Al+Sayofa + Al+mADiya +
        lawoHapF + min + 10 + >amotArK + murab~aEapK + fiy + >aqal~a + min + sAEapK
        Hit + the+number + the+records + with + the+size + latest + painting+his + 200 + meters +
        square + and+with+the+speed + executed + the+summer + the+previous + painting + of
        + 10 + meters + square + in +less +of + hour
      He set a record in size (his last painting measures 200 square meters) and in time (he
        finished last summer a 10 square meter painting in less than an hour)
```

(S (VP إنْ need transliteration and gloss here
 (NP-SBJ (NP المؤمن need transliteration and gloss here)
 (PRN (S (VP أعني need transliteration and gloss here
 (NP-SBJ *)
 (NP-OBJ المؤمن الحقيقي need transliteration and
 gloss here))))
 (NP (NP هو need transliteration and gloss here))
 (NP-PRD (NP ذلك need transliteration and gloss here)
 (SBAR (WHNP-1 الذي need transliteration and gloss here)
 (S (VP يصدق need transliteration and gloss here
 (NP-SBJ-1 *T*)
 (PP في need transliteration and gloss here
 (NP (NP قول need transliteration and gloss
 here
 (NP ه need transliteration and gloss
 here))
 و need transliteration and gloss here
 (NP فعل need transliteration and gloss
 here
 (NP ه need transliteration and gloss
 here))))))))))

إنّ المؤمن أعني المؤمن الحقيقي هو ذلك الذي يصدق في قوله وفعله

the believer mean+I the believer the real he that who be sincere in words+his and deed+his
The believer (I mean the true believer) is the one who is sincere in his words and his deeds.

NB: need transliteration here and in the tree

Another typical construction that is annotated as PRN in the ATB is the sentential modification of nouns starting with wa/and and a personal pronoun or demonstrative pronoun (هو wa huwa and he, or هي wa hiya and she, or ذلك wa *alika and this تلك wa tiloka and this etc.), when it comes in the middle of a main sentence.

(S (VP إنْ need transliteration and gloss here
 (NP-SBJ-1 (ADJP آخر
 (NP (NP أجل need transliteration and gloss here)
 (PP ل need transliteration and gloss here
 (NP قبول need transliteration and gloss here
 (NP الترشيحات need transliteration and gloss
 here))))
 (PRN و need transliteration and gloss here
 (S (NP-SBJ ذلك need transliteration and gloss here)
 (PP-TMP-PRD في need transliteration and gloss here)
 (NP موفى need transliteration and gloss here
 (NP شهر need transliteration and gloss here
 (NP سبتمبر need transliteration and gloss
 here))))))
 (S (VP (PRT قد need transliteration and gloss here)
 فات need transliteration and gloss here
 (NP-SBJ-1 *)))))
 إنّ آخر أجل لقبول الترشيحات وذلك في موفى شهر سبتمبر قد فات

Need transliteration here

Indeed last term for accepting the candidates and this in end month September probably passed.

Indeed the deadline for accepting candidates (That was at the end of September) has probably passed.

(S و-·wa-·and
 (VP -يواجه-·yu+wAjih+u·he/it+face/be_faced_with/confront+[ind.]
 (NP-SBJ (NP الضباط ·Al+Dub~AT+u·the+officers+[def.nom.]
 (NP الثلاثة ·Al+valAv+ap+u·the+three+[fem.sg.]+[def.nom.]
 ,
 (PRN و-·wa-·and
 (S (NP-SBJ هم-·-hum·they_[masc.pl])
 (NP-PRD (NP الجنرال ·Al+jinirAl+u·the+general
 بورييس ·buwriys·Boris
 فادييف ·fAdyyf·NOT_IN_LEXICON)
 و-·wa-·and
 (NP -القومانداُن ·-Al+quwmanodAn+u·
 the+commandant
 إيغور ·<iyguwr·Igor
 تيخونوف ·tyxwnwf·NOT_IN_LEXICON)
 و-·wa-·and
 (NP -الكولونيل ·-Al+kuwluwniyl+u·
 the+colonel
 ميخائيل ·miyxAl·Mikhail
 ليفتشنكو ·lft\$nkw·NOT_IN_LEXICON))))
 (NP-OBJ تهمّة ·tuhom+ap+a·accusation/charge+[fem.sg.]+[def.acc.]
 (NP الإهمال ·Al+<ihomAl+i·the+neglect/carelessness))))
 ويواجه الضباط الثلاثة وهم الجنرال بورييس فادييف والقومانداُن إيغور تيخونوف والكولونيل ميخائيل ليفتشنكو تهمّة الإهمال
 wa+yuwAjihu + Al+Dub~Atu + Al+valAvapu + wa+hum + Al+generAl + buwris + fAdiyef
 + wa+Al+quwmundwun + <iguwr + tyxuwnuwf + wa+Al+kuwluwniyl + mixa}iyl +
 lefot\$inkuw + tuhomapa + Al+>ihomAl
 and + face + the + officers + the + three + and + they + the + General + Boris + Fadiev + and
 + the + commandant + Igor + Tikhonoff + and + the + colonel + Mikhael + Levchenko +
 the + charge + the + neglect
*And the three officers, who are the general Boris Fadiev, the commandant Igor Tikhonoff
 and the colonel Mikhael Levchenko, face a charge of neglect.*

Note that wa huwa/ wa hiya instances, if they don't come up in the middle of a main sentence, are annotated as complete independent sentences.

In addition, it is important in these instances not to confuse PRN with SBAR-ADV, the wAw Al Hal with personal pronoun. PRN is used when the meaning of the PRN clause is not related to the verb, as in

```

(S (VP زرت
  (NP-SBJ *)
  (NP-OBJ القاهرة)
  (PRN و
    (S (NP-SBJ هي)
      (NP-PRD (NP مدينة)
        (SBAR (WHNP-1 *0*)
          (S (VP بنيت
            (NP-SBJ-1 *T*)
            (NP-OBJ-1 *)
            (PP على
              (NP يد
                (NP الفاطميين ) ) ) ) ) ) ) ) ) ) ) )
  (PP من
    (NP جديد ) ) ) )

```

زرت القاهرة وهي مدينة بنيت على يد الفاطميين من جديد

Need transliteration here and in the tree

Visited Cairo and she city was built on hand Fatimides from new.

I visited Cairo (it is a city built by Fatimedes) again.

SBAR-ADV is used when what comes after the wa is related to the verb, as in

```

(S (VP زرت
  (NP-SBJ *)
  (NP-OBJ القاهرة)
  (SBAR-ADV و
    (S (NP-SBJ هي)
      (PP-PRD في
        (NP أوج
          (NP إشراق
            (NP ها ) ) ) ) ) ) ) )
  (PP-TMP منذ
    (NP سنين
      (NP طويلة ) ) ) )

```

زرت القاهرة وهي في أوج إشراقها منذ سنين طويلة

Need transliteration here and in the tree

Visited Cairo and she in peak sunrise its since years long

I visted Cairo years ago when it was at its peak.

(S (VP -قُوبِلُوا-quwbiI+uwA·be_met_[with]/be_faced_[with]+they_[verb]
 (NP-SBJ-3 *)
 (NP-OBJ-3 *)
 (PRN وَ-wa-and
 (INTJ -يَا-yA·oh/you)
 (PP لـ-li-for/to
 (NP -الأسف-Al+>asaf+i·the+regret/sorry+[def.gen.])))
 (PP-MNR بِ-bi-by/with
 (NP (NP -المماطلة-Al+mumATal+ap+i·the+procrastination)
 وَ-wa-and
 (NP -تزييف-tazoyiyf+i·forgery/falsification+[def.gen.]
 (NP الحقائق-Al+HaqA}iq+i·the+truth/facts))))))
 قُوبِلُوا وَيَا لِلْأَسْفِ بِالمماطلة وتزييف الحقائق

quwbiIwA + wa+yA li+Al>asafi + bi+AlmumATalapi + wa+taziyyfi + AlHaqA}iqi
 were faced + and + Oh + for + sorrow + with + the+ procrastination + and + falsification +
 the + truths

They were faced, unfortunately, with procrastination and falsification of reality.

(S (VP عسى·EasaY·perhaps
 (PRN وَ-wa-and
 (NP -مليون-·miloyuwn+u·million+[def.nom.]
 (NP عسى·EasaY·perhaps)))
 (SBAR-SBJ أن·>an·to
 (S (VP يَكُونُ·ya+kuwn+a·he/it+be+[sub.]
 (PP-PRD لـ-la-to/for
 (NP -كُنْ-kun~a·you_[fem.pl.]_[you_have]))
 (NP-SBJ (NP أَصْوَاتٌ·>aSowAt+N·votes+[indef.nom.]
 (PP في·fiy·in
 (NP {نِتخاب:{inotixAb+i·election/selection
 (NP شُرَكَاء·\$urakA'+i·partners
 (NP -حَيَات·HayA+at+i·life
 (NP -كُنْ-kun~a·your))))))))))
 عسى ومليون عسى أن يكون لكن أصوات في انتخاب شركاء حياتكن

EasaY + wa+miloywn + EasaY + <an + yakuwna + la+kun~a + >aSwAtuN + fiy + lintixAbi
 + \$urakA'i + HayAtikun~a
 Hopefully + and + million + hopefully + that + be + for + you + voices + in + election +
 partner + lives+yours

Hopefully, and a million hopefully, you will have a say in electing the partner of your life.

2.3.1.3.1.2 Not PRN anymore

There are a few common constructions where PRNs were often used, and now should not be used any more. These are exemplified below:

Renamings – should be adjunction with the parentheses inside the adjoined constituent

(NP (NP -حقّ·-Haq~+i·truth/right+[def.nom.]
 (NP النّقْض·Al+naqoD+i·the+contradiction/destruction+[def.gen.]))
 (NP -LRB-·
 الفَيْتُو·Al+fiytuw·the+veto
 -RRB-·))

حقّ النقض (الفيتو)

Haq~u + AlnaqoDi + (Al+fiytuw)

Right + the+overturn + (the+veto)

The right to overturn (veto)

Additional modifiers of NP – should be adjunction with the parentheses inside the adjoined constituent

(S (NP-TPC-3 (NP مُعْظَم·muEoZam+a·most_of/the_majority_of+[def.acc.]
 (NP الأبنية·Al+>aboniy+ap+i·the+buildings))
 (SBAR -LRB-·
 (WHNP-2 الَّتِي·Al~atiy·which/who/whom_[fem.sg.])
 (S (VP تَقَعُ·ta+qaE+u·it/they/she+be_located+[ind.]
 (NP-SBJ-2 *T*)
 (NP-TMP حَالِيًا·Haliy~+AF·currently+[acc.indef.])
 (NP-LOC قُرْب·qurob+a·near+[def.acc.]
 (NP الحَرَم·Al+Haram+i·the+holy site+[def.gen.]
 الشَّرِيف·Al+\$ariyf+i·the+holy)))
 -RRB-·))
 (VP وقع waqaEa happen
 (NP-SBJ إعادة <iEAdap+u redoing
 (NP تهيئة tahoyi<ap+i
 (NP (NP ها hA)
 (NP-3 *T*))))))))

معظم الأبنية (التي تقع حاليا قرب الحرم الشريف) وقع إعادة تهيئتها.

muEoDamu + Al>abobyapi + Al~atiy + taqaEu + Haliy~AF + quroba + AlHarami +
 Al\$ariyfi **Needt he rest of the transliteration here**

most + the+buildings + that + be located + currently + near + the+site + the+holy took place
 redoing restauration its

Most of the buildings (that are now close to the Holy Site) were rehabilitated.

Additional sentential/VP modifiers – should be in the usual modifier position, with the parentheses inside the constituent

```
(NP (NP القِمة .Al+qim~+ap+i.the+summit+[fem.sg.]+[def.gen.]
    العَرَبِيَّة .Al+Earabiy~+ap+i.the+Arab/Arabic+[fem.sg.]+[def.gen.] )
(SBAR (WHNP-1 الَّتِي .Al~atiy.which/who/whom_[fem.sg.] )
    (S (VP (PRT سَ .sa-.will)
        -تُعَقَّدُ .-tu+Eoqad+u.it/they/she+be_held/be_convened+[ind.]
        (NP-SBJ-1 *T*)
        (NP-OBJ-1 *)
        (PP-LOC في .fiy.in
            (NP بَيْرُوت bayoruwt+a.Beirut+[indef.gen.] ) )
        (SBAR-ADV -LRB-.
            إِذَا .<i*A.if/whether
            (S (VP أَرَادَتْ .>arAd+at.want/desire+it/they/she
                (NP-SBJ إِسْرَائِيلُ .<isorA)iyl+u.Israel)
                -RRB-. ) ) ) ) ) ) )
```

القمة العربية التي ستعقد في بيروت (إذا أرادت اسرائيل)
 Alqim~apu + AlEarabiy~apu + Al~atiy + sa + tuEoqadu + fiy + bayoruwt + <i*A + >arAdat
 + <isorA)iyl
 The+summit + the+Arab + that + will + be held + in + Beirut + if + wants + Israel
The Arabic summit that will be held in Beirut (if Israel wants).

Bylines – should be FRAG or NP (not PRN or X) of necessary – most of these are NP in Arabic

In general, PRN should never be an allowable top node.

```
(TOP (NP -LRB-.
    (NP و .w.
        ص .S.
        ف .f. )
    (NP رُوَيْتِرْز .ruwyotirz.Reuters)
    '
    (NP أ ب .>b.A.P. )
    -RRB-. ) )
```

(و, ص, ف, رويترز, أ ب)

w.S.f.+ ruwyotiz + >a.b
 AFP + Reuters + AP
 (AFP, Reuters, AP)

Bylines are generally NPs in Arabic, since they usually consist of a string of noun phrases.

End of the paragraph – parentheticals at the end of a paragraph should be split off to form a new, coordinated sentence

(S (VP ضَرَبَ·Darab+a·strike/hit+he/it_[verb]
 (NP-SBJ *)
 (NP-OBJ الرَقْمُ·Al+raqom+a·the+number
 القياسي·Al+qiyAsiy~+a·the+record/)
 (PP-MNR (PP بِ·bi-·by/with
 (NP - الحِجْمُ -Al+Hajom+i·the+size))
 (PRN -LRB-
 (S (NP-SBJ (ADJP آخِرُ·|xir+u·last
 (NP لَوْحَاتٍ painting
 (NP - هِ·-hi·his)))
 (NP-PRD 200·
 (NP مِثْرٍ·mitor+K·meter
 مُرَبَّعٍ·square)))
 -RRB-·)
 وَ·wa-·and
 (PP -بِ·-bi-·by/with
 (NP - السَّرْعَةُ -Al+suroE+ap+i·the+speed))
 (S (VP نَفَّذَ·accomplish+he
 (NP-SBJ *)
 (NP-TMP الصَّيْفُ·the+summer
 الْمَاضِي·the+past)
 (NP-OBJ (NP لَوْحَةٌ·painting
 عِمْلَاقَةٌ·gigantic)
 (PP مِنْ·min·from
 (NP عَشْرَةٌ·ten
 (NP أَمْتَارٍ·meters
 مُرَبَّعَةٍ·square))))
 (PP-TMP فِي·fiy·in
 (NP (ADJP أَقَلُّ·>aqal~+a·less
 (PP مِنْ·min·from
 (NP سَاعَةٌ·hour)))))))))
 ضرب الرقم القياسي بالحجم (آخر لوحاته 200 متر مربع) و بالسرعة (نفذ الصيف الماضي لوحة
 عملاقة من 10 أمتار مربعة في أقل من ساعة
 Daraba + Al+raqoma + Al+qiyAsiy~a + bi+Al+Hajomi + lxiru + lawoHAti+hi + 200 +
 mitorK + murab~aEK + wa+bi+Al+suroEapi + naf~a*a + Al+Sayofa + Al+mADiya +
 lawoHapF + min + 10 + >amotArK + murab~aEapK + fiy + >aqal~a + min + sAEapK
 Hit + the+number + the+records + with + the+size + latest + painting+his + 200 + meters +
 square + and+with+the+speed + executed + the+summer + the+previous + painting + of
 + 10 + meters + square + in +less +of + hour
*He set a record in size (his last painting measures 200 square meters) and in time (he
 finished last summer a 10 square meter painting in less than an hour)*

2.3.1.3.2 NAC (Not A Constituent)

NAC is used to hold a conjunction and a constituent together when they need to operate as a unit, in spite of the fact that they do not together form a typical constituent (Not A Constituent). The structure for coordination in the Treebank requires a node level for coordination, multiple conjuncts, and the conjunction (see section ???? on coordination). In the context where only the

conjunction and a single conjunct are present, they do not together form a “constituent” that exists in any other context in the Treebank. They do, however, need to be annotated together as a single node, and the node used for this purpose is NAC. (NB: This is parallel to the use of NAC in the English Treebanks for examples such as *He walked to school, and slowly.*)

NAC is used with the conjunction *wa* 9 for two types of construction: (1) emphasizing modification and (2) discontinuous coordination. NAC can occur with any conjunction or CONJP.

2.3.1.3.2.1 NAC emphasizing modification

The NAC construction can be used to introduce modification (of either VP or NP) in order to emphasize it. In this construction, the conjunction is omissible without affecting the rest of the sentence.

```
(S (VP (PRT قد::-qad::[has/have])
  بتنا::bit+nA::become/remain+we_[verb]
  (NP-SBJ-1 *)
  (S (VP ندرک::nu+dorik+u::we+comprehend/realize+[ind.]
    (NP-SBJ-1 *)
    (NAC و::wa-::and
      (PP-MNR ب::-bi-::by/with
        (NP وعي::-waEoy+K::consciousness/awareness/
          attention+[indef.gen.]
          کامل::kAmil+K::complete/full/
          integral+[indef.gen.])))
    (SBAR-NOM-OBJ (WHNP-3 ماذا::mA*A::what)
      (S (VP حصل::HaSal+a::occur/happen/
        take_place+he/it_[verb]
        (NP-SBJ-3 *T*))))))
  قد بتنا ندرک ووعي کامل ماذا حصل
```

qad + bit+nA + nudorik+u + wa+bi+waEoy+K + kAmil+K + mA*A + HaSal+a
 have + become + realize + and+with+consciousness + complete + what + happen
With complete conscious, we have realized what happened

(S (VP (PRT -qad::قد::[has/have])
 >avobat+a::أثبت::ascertain/establish+he/it_[verb]
 (NP-SBJ *)
 (NP-ADV sAbiq+AF::سابقاً::formerly/earlier+[acc.indef.])
 (NP-OBJ (NP jadAr+at+a::جدارته::worthiness/aptitud
 (NP -hu::هو::his/its))
 ,::,::nogloss
 (NAC wa-::وَ::and
 (PP (NP-ADV -xuSuwS+AF::-خُصُوصاً::especially)
 bi-::بِ-::by/with
 (NP -<izA'i::إزاء::towards
 (NP Al+muEAriD+iyna::المُعَارِضِينَ::the+opponent
)))))

قد أثبت سابقاً جدارته وخصوصاً بإزاء المعارضين

-qad >avobat+a sAbiq+AF jadAr+at+a- -hu, wa- -xuSuwS+AF bi- -<izA`+i
 Al+muEAriD+iyna

Has established he formerly aptitude and especially towards the opponents.

He has formerly established his aptitude especially towards the opponents.

When the NAC emphasizes a relative clause modification, the relative clause must begin with a relative pronoun:

(S (VP -{isotamaE+a::-{سَمِعَ::listen+he/it_[verb]
 (NP-SBJ *)
 (PP-CLR <ilaY::إلى::to/towards
 (NP (NP muEAnA+ap+i::مُعَانَاة::hardship/suffering/
 effort+[fem.sg.]+[def.gen.]
 (NP Al+>ahAliy::الْأَهَالِي::the+families/indigenous_people))
 (NAC wa-::وَ::and
 (SBAR (WHNP-1 -Al~atiy::الَّتِي::which/who/whom_[fem.sg.])
 (S (VP tarak~az+at::تَرَكَّزَتْ::concentrate/
 focus+it/they/she_[verb]
 (NP-SBJ-1 *T*)
 (PP-CLR EalaY::عَلَى::on/above
 (NP Al+>awoDAE+i::الْأَوْضَاع::the+
 conditions/
 situation +[def.gen.]
 Al+maEiy\$iy~+ap+i::الْمُعِيشِيَّة::
 the+livelihood+
 [fem.sg.]+[def.gen.]
))))))))

إستمع الى معاناة الاهالي والتي تركزت على الأوضاع المعيشية

{isotamaE+a + <ilaY + muEAnA+ap+i + Al+>ahAliy + Al~atiy + tarak~az+at + EalaY +
 Al+>awoDAE+i + Al+maEiy\$iy~+ap+i

listen + to + hardship + the+indigenous people + which + concentrate + on + the+conditions
 + the+livelihood

*He listened to the hardship of the indigenous people which concentrated on the livelihood
 conditions*

```

(S (VP قاد qAd+a·lead/guide+he/it_[verb]
  (NP-SBJ *)
  (NP-TMP أخيراً >axiyr+AF·last/latest/recent/latter+[acc.indef.])
  (NP-OBJ (NP (NP المفاوضات Al+mufAwaD+At+i·the+negotiations/talks)
    (NP-ADV بينَ bayon+a·between/among+[def.acc.]
      (NP (NP الأردن Al>urdun~i·the+Jordan+[def.gen.]
        وـ wa-·and
        (NP الولايات -Al+wilAy+At+i·the+States
          المتحدة Al+mut~aHid+ap+i·the+United))))))
  (NAC وـ wa-·and
    (SBAR (WHNP-2 التي -Al~atiy·which/who/whom_[fem.sg.])
      (S (VP توجت tuw~ij+at·be_crowned+it/they/she_[verb]
        (NP-SBJ-2 *T*)
        (NP-OBJ-2 *)
        (PP-CLR بـ bi-·by/with
          (NP توقيع -tawoqiY+e+i·signing
            (NP (NP {تفاق: {it~ifAq+i treaty
              (NP التجارة Al+tijAr+ap+i·
                the+commerce
                الحرة Al+Hur~+ap+i·
                  the+free))
            (NP-ADV بينَ bayon+a-
              between
              (NP -هُما -humA·
                them))))))))))

```

قاد أخيراً المفاوضات بين الأردن والولايات المتحدة و التي توجت بتوقيع اتفاق التجارة الحرة بينهما
 qAda + >axiyrAF + AlmufAwaDAti + bayona + Al>urdun + wa+alwilAyAti +
 Almut~aHidapi + wa + Al~atiy + tuw~ij+at + bi+tawoqiYei + Ai~ifAqi +
 Al+tijAr+ap+i + Al+Hur~+ap+i + bayona+humA
 conducted + recently + the+negotiations + between + Jordan + and+the+states +
 the+united + and + which + be crowned + with+signing+treaty + trade + free +
 between+them
*He conducted lately the negotiations between Jordan and the United States, that were
 crowned with the signing of a trade treaty between them*

NAC may be rightward moved, in which case it will need to be traced with *ICH* to the location
 in the tree where it is interpreted:
 (see section ?????????????? on *ICH*)

(S (VP تقوم::ta+quwm+a::it/they/she+undertake/carry_out+[sub.]
 (NP-SBJ *)
 (PP-CLR ب::bi-::by/with
 (NP (NP ضم::-Dam~+i::joining/addition/annexation+[def.gen.]
 (NP (NP الكتل::Al+kutal+i::the+blocs/groups+[def.gen.]
 الاستيطانية ::Al+{isotiyTAniy~+ap+i::
 the+settlement/colonial+[fem.sg.]+[def.gen.]
 الكبرى::Al+kuboraY::the+larger/
 largest/greater/greatest)
 (NAC-1 *ICH*))
 (PP الي::<ilay-::to/towards
 (NP ها::-hA::it/them/her))
 (NAC-1 و::wa-::and
 (SBAR (WHNP-1 التي::-Al~atiy::which/who/whom_[fem.sg.])
 (S (VP تضم::ta+Dum~+u::it/they/she+
 include/comprise/incorporate+[ind.]
 (NP-SBJ-1 *T*)
 (NP-OBJ آلاف :: lAf+a:: thousands
 (NP المستوطنين ::
 Al+musotawoTin+iyna::
 the+settler+[masc.pl.gen.]
)))))))))))

تقوم بضم الكتل الاستيطانية الكبرى اليها و التي تضم آلاف المستوطنين

taquwm+a + bi+Dam~+i + Al+kutal+i + Al+<istiyTAniy~ap+i + Al+kuboraY + <ilay+hA +
 wa+Al~atiy + taDum~+u + lAf+a + Al+musotawoTin+iyna
 undertake + with + joining + the+blocks + the+colonial + the+largest + to+it + and+which +
 include + thousands + the+settlers

It involves the largest colonial groups that have thousands of settlers

N.B. When the NAC emphasizes a relative clause modification, the relative clause should begin with an overt relative pronoun; if not, it is a case of PRN or SBAR-ADV with wAw Al+HAl.

(S (NP-TPC-1 (NP رئيس::ra}iysu::president/head/chairman
 (NP الشرطة::Al+\$uroTapi::the+police+[fem.sg.])
 (PP-LOC في::fiy::in
 (NP إقليم::<iqoliymi::region/district
 (NP السند::Al+sinod::the+Sind))))

(PRN و::wa-::and

(S (NP-SBJ عاصمة::-EASimatu-::capital_city
 (NP ه::-hu::his/its))

(NP-PRD كراتشي::karAto\$iy::Karachi)))

(VP qAla qAla said
 ...)

رئيس الشرطة في إقليم السند وعاصمته كراتشي قال.....

ra}iysu Al+\$uroTapi fiy <iqoliymi Al+sinod wa- -EASimatu- -hu karAto\$iy qAla...
 ra}iysu + Al\$uroTapi + fiy + <iqoliymi + Alsind + wa+EASimatu+hu + KaraAt\$i

The police chief in the district of Sind (Karatshi is its capital) said....

NAC can occur with any conjunction or CONJP.

(S wa- و
 (VP -qAl+a قال
 (NP-SBJ maSodar+N مصدر
 <iEolAmiy~+N اعلامي
 rasomiy~+N رسمي
 qaTariy~+N قطري)
 (PP li-ل
 (NP -wikAl+ap+i وكالة
 (NP frAnos bris برس))
 (SBAR >an~a أن
 (S (NP-TPC-1 ra}iys+a رئيس
 (NP wuzarA'+i وزراء
 (NP qaTar قطر))
 (VP (PRT sa-س)
 -ya+Eoqid+u يعقد
 (NP-SBJ-1 *T*)
 (NP-OBJ mu&otamar+AF مؤتمرا
 SuHufiy~+AF صحفيا)
 (NP-TMP HawAlay حوالي
 (NP (NP Al+sAE+ap+i الساعة
 Al+vAniy+ap+i الثانية)
 (PP min من
 (NP baEod+i بعد
 (NP Zuhor+i ظهر
 (NP (NP Al+yawom+i اليوم
 (NP Al+vulAvA'+i
 (الثلاثاء))))))))))
 (NAC (CONJP wa- و
 -l`kino لكن)
 (PP bi-ب
 (NP -duwn+i دون
 (NP <iEoTA'+i اعطاء
 (NP (NP maziyd+K مزيد
 (PP min من
 (NP Al+tafASiyl+i التفاصيل))))))))))
 .)

و قال مصدر اعلامي رسمي قطري لوكالة فرانس برس أن رئيس وزراء قطر سيعقد مؤتمرا صحفيا حوالي الساعة الثانية من بعد ظهر اليوم الثلاثاء ولكن بدون اعطاء مزيد من التفاصيل.

Needs transliteration here

And said source media official Qatari to agency France Presse that president ministers Qatar will hold conference journalistic around the hour the second from after afternoon the day the Tuesday and but giving more of the details.

And a Qatari official media source said to AFP that the Qatari Prime Minister will hold a press conference around two of the afternoon today, Tuesday, but without giving any further details.

NAC is also used when there is discontinuous coordination (in which case an *ICH* trace is usually necessary as well).

(S و wa
 (NP-TPC-1 هِيَ hiy~a she)
 (VP دَأَبَتْ da>ab+at get used
 (NP-SBJ-1 *T*)
 (PP-CLR عَلَى EalaY to
 (NP (NP استراق } isotirAq+I stealing
 (NP النوم Al+nawom+i the sleep))
 (PP-LOC (PP فِي fiy in
 (NP السَّيَّارَةُ Al+say~Ar+ap+i the car))
 (NAC-2 *ICH))
 (S-ADV (VP مُتَنَقِّلَةً mutanaq~il+ap+F moving
 (NP-SBJ-1 *)
 (PP-DIR مِنْ mino from
 (NP تَجْمُوع tajam~uE+K meeting
 {إِنْتِخَابِيَّ {inotixAbiy~+K electoral))
 (PP-DIR إِلَى <ilaY to
 (NP آخَر |xar+K another))
 (NAC-2 أَوْ >awo or
 (PP-LOC عَلَى EalaY on
 (NP مَتْنِ matoni board
 (NP (NP الطَّائِرَةُ Al+TA<irap+i the plane)
 (SBAR (WHNP-3 الَّتِي Al~Atiy which)
 (S (VP تَتَنَقَّلُ ta+tanaq~al+u move
 (NP-SBJ *)
 (PP-LOC فِي fiy in
 (NP (NP هَا hA it)
 (NP-3 *T*))
 (NP-TMP يَوْمِيًا yawomiy~+F daily)
 (NP-DIR بَيْنَ bayona between
 (NP (NP الْوَلَايَاتِ Al+wilAy+At+i
 the states))))))))))
 وَهِيَ دَأَبَتْ عَلَى اسْتِرَاقِ النَّوْمِ فِي السَّيَّارَةِ مُتَنَقِّلَةً مِنْ تَجْمُوعِ إِنْتِخَابِيٍّ إِلَى آخَرٍ أَوْ عَلَى مَتْنِ الطَّائِرَةِ الَّتِي تَتَنَقَّلُ فِيهَا
 . يَوْمِيًا بَيْنَ الْوَلَايَاتِ

wa hiy~a da>ab+at EalaY {isotirAq+i Al+nawom+i fiy Al+say~Ar+ap+i mutanaq~il+ap+F
 mino tajam~uE+K {inotixAbiy~+K <ilaY |xar+a >awo EalaY matoni Al+TA<irap+i
 Al~Atiy ta+tanaq~al+u fiy hA yawomiy~+F bayona Al+wilAy+At+i
 And she worked on stealing sleep in the car moving from gathering electoral to another or on
 aboard the plane that moves she in it every day between the states
 She got used to having a nap in the car while she was going from one electoral assembly to
 another or in the plane while she was flying her from state to state.

(PP على::EalaY::on/above
 (NP أمل::>amali::hope/wish))
 (SBAR ان::>an::to
 (S (VP يفي::ya+fiy+a::he/it+fulfill/carry_out+[sub.]
 (NP-SBJ المسؤولون::Al+maso&uwluwna::the+functionary)
 (PP-CLR ب::bi-::by/with
 (NP (NP بقية::baqiy~api::remainder/remnant
 (NP وعود::wuEwdi-::promises
 (NP هم::him::their))
 (NAC-1 *ICH*))
 (PP-TMP في::fiy::in
 (NP (ADJP اقرب::aqorabi::soonest/nearer
 (NP فرصة::furoSapK::opportunity/chance
 ممكنة::mumokinapK::possible)))
 (NAC-1 و::wa-::and
 (NP (NP-ADV خاصة::xAS~apF::especially)
 اعادة::iEAdapu::return/repetition
 (NP تأهيل::ta>ohiyli::habilitating
 (NP البنية::Al+bunoyapi::the+structure
 التحتية::Al+taHotiy~api::the+under/
 infra-)))))))))
 على أمل أن يفي المسؤولون ببقية وعودهم في أقرب فرصة ممكنة وخاصة اعادة تأهيل البنية التحتية.
 ElaY + <amali + <an + yafiy + Almaso&wlwna + bi+baqy~api + wuEwdi+him + fiy +
 <aqrabi furSapK + mumkinapK + wa+xASapF + <iEAdapu + ta>ohiyli + AlbunaY +
 AltaHtiy~api.
*With the hope that the officers keep the rest of their promises as soon as possible and
 especially re-habilitating the infra-structures.*

2.3.1.3.2.2 NAC vs. UCP

Note that NAC should never be used for examples where two unlike constituents are coordinated. UCP should be used for this type of coordination.

Some examples of this type were incorrectly annotated as NAC in previous releases – these are all correctly annotated as UCP below.

(S (NP-TPC-1 (NP -h`*ihi::هـَـذِهِ::this/these)
 (NP Al+muEAdal+ap+u::المُعَادَلَةُ::the+equation+[fem.sg.]+[def.nom.]))
 (VP ta+Eoniy+[null]::تَعْنِي[نُلل]::it/they/she+mean/signify/concern+[ind.]
 (NP-SBJ-1 *T*)
 (UCP-ADV (NP-ADV mubA\$ar+ap+F::مُبَاشَرَةً::directly +[fem.sg.]+[indef.acc.]
 wa-::وَ-::and
 (PP -bi-::بِ-::with
 (NP (PRT -la-::لَا-::no_[without])
 tajomiy1+K::تَجْمِيلٌ::embellishment+[indef.gen.])))
 (NP-OBJ <ilogA'+a::إِلْغَاءٌ::cancellation/abrogation/repeal+[def.acc.]
 (NP Al+wikAl+At+i::الوَكَالَاتُ::the+agencies+[fem.pl.]+[def.gen.]
 Al+tijAriy~+ap+i::التِّجَارِيَّةُ::the+commercial/
 business+[fem.sg.]+[def.gen.]
 Al+HaSoriy~+ap+i::الْحَضْرِيَّةُ::
 the+exclusive+[fem.sg.]+[def.gen.]))))
 هذه المعادلة تعني مُباشرةً وبلا تجميل إلغاء الوكالات التجارية الحصرية
 h`*ihi + Al+muEAdal+ap+u + taEoniy + mubA\$ar+ap+F + wa+bi+la + tajomiy1+K +
 <ilogA'+a + Al+wikAl+At+i + Al+tijAriy~+ap+i + Al+HaSoriy~+ap+i
 this + the+equation + mean + directly + and+with+no + embellishment + cancellation +
 the+agencies + the+commercial + the+executive
 With no embellishment, this equation means cancellation of the executive commercial
 agencies

(NP (NP المنطقة·Al+minoTaq+ap+i·the+area/zone/territory+[fem.sg.]+[def.gen.])
 (UCP (ADJP المنزوعة·Al+manozuE+ap+i·the+removed/taken_away
 (NP السلاح·Al+silAH+i·the+[military]_service_branch))
 و-·wa-·and
 (SBAR (WHNP-1 الَّتِي·-Al~atiy·which/who/whom_[fem.sg.])
 (S (VP تُعْتَبَرُ·tu+Eotabar+u·it/they/she+be_considered
 (NP-SBJ-1 *T*)
 (S (NP-SBJ-1 *)
 (NP-PRD (NP ملاذٌ·malA*+AF·shelter/sanctuary)
 (PP لـ·li·for/to "·"·nogloss
 (NP القُوَات·Al+quw~+At+i·the+armed_forces
 الثَّوْرِيَّةُ·Al+vaworiy~+ap+i·
 the+revolutionary
 المُسَلَّحَةِ·Al+musal~aH+ap+i·
 the+armed))))))
 المنطقة المنزوعة السلاح والتي تعتبر ملاذاً للقوات الثورية المسلحة
 AlminoTaqapi + AlmanozuwEapi + alsilAHi + wa+Al~atiy + tuEotabar + malA*AF +
 li+AlquwAti + Alvawory~api + Almusal~aHapi
 The+region + the+removed + the+arms + and+that + be considered + shelter +
 for+the+forces + the+rebel + the+armed
 The disarmed region, that is considered as a good shelter for the armed rebel forces.

2.3.1.3.2.3 NAC vs. Gapping

See section ????? on NAC, and section ????? on gapping.

NAC should not be confused with gapping. In a coordination structure such as the one in this example, the interpretation is of gapping – i.e., xyz participated in it, and abc participated in it

also. This is not an instance of discontinuous coordination needing NAC because the coordination is not discontinuous (the “wa” is adjacent to the first NP conjunct).

(S (PP-TMP -مُنْذُ-muno*u.since/ago
 (NP (NP ذَٰلِكَ`lika.that)
 (NP الوقت .Al+waqot+i.the+time/moment/period_of_time)))
 (VP (VP صَارَتْ .SAr+at.become/begin_to+it/they/she_[verb]
 (NP-SBJ=2 الكَتَائِبُ .Al+kataA}ibu+u.the+brigades/squadrons/corps)
 (PP-PRD=3 فِي .fiy.in
 (NP جانب .jAnib+K.side/aspect+[indef.gen.])))
 وَ .wa.and
 (VP (NP-SBJ=2 "الْقُوَّاتُ .Al+quw~+At+u.the+armed_forces
 ".)
 (PP-PRD=3 فِي .fiy.in
 (NP جانب .jAnib+K.side/aspect+[indef.gen.]
 آخَرَ .|xar+a.other/another/latest+[indef.gen.]))))))
 ومنذ ذلك الوقت صارت الكتائب في جهة والقوات في جانب آخر
 wa+muno*u + *alika + Alwaqoti + SARat + AlkatA}ibu + fiy + jAnibK + wa+Alquw~Atu +
 fiy + jAnibK + lxara
 and+since + that + time + become + the+brigades + in + side + and+the+forces + in + side +
 other
And since then, the brigades are one side and the (armed) forces on another

2.3.1.3.2.4 NAC vs. prepositional “wa” (waw AlmaEiy~ap)

(S وَ .wa.and
 (PP-LOC -مِنْ-min.from
 (ADVP هُنَاكَ .hunAka.there))
 (VP تَوَجَّهَ .tawaj~ah+a.go_towards/turn_towards/face+he/it_[verb]
 (NP-SBJ *)
 (PP وَ .wa.with
 (NP مُرَافِقِي- .murAfiq+iy-.companion/adjutant+[masc.pl.acc.]
 (NP -هِ-hi.his/its)))
 (PP-DIR إِلَى .<ilaY.to/towards
 (NP مَقَرُّ .maqar~+i.center/headquarters/residence+[def.gen.]
 (NP وَزَارَةِ .wizAr+ap+i.ministry+[fem.sg.]+[def.gen.]
 (NP الشَّبَابِ .Al+\$abAb+i.the+youth/youths+[def.gen.]
 وَ .wa.and
 -الرِّيَاةِ .-Al+riyAD+ap+i.the+sport))))))
 ومن هناك توجه ومرافقيه إلى مقر وزارة الشباب والرياضة
 wa+min+hunAka + taw~ajaha + wa+murAfiqiy+hi + <ilaY + maqar~I + wizArapi +
 Al\$abAbi + wa+AlriyADapi
 and+from+there + headed + with+adjutants + to + headquarter + minestry + the+youth +
 and+the+sport
*And from there is went along with with his adjutants to the headquarter of the minestry of
 youth and sports.*

2.3.1.3.2.5 NAC vs. particle “wa”

There are two types of “wa” that are optional in the sentence: the conjunction “wa” that introduces an NAC, and the particle “wa” that precedes an SBAR with >ano. NAC wa introduces optional modifiers of the sentence, while particle wa introduces obligatory clausal elements in the sentence. In both cases the wa can be omitted. NAC wa can be replaced by any other CONJ or CONJP, while wa as particle cannot.

```
(S (VP سبق
    (SBAR-SBJ (PRT و)
        أن
        (S (VP رأيت
            (NP-SBJ *)
            (NP-OBJ ذلك ) ) ) ) ) ) )
```

.سبق و أن رأيت ذلك

NB: need transliteration for the full sentence and in the tree, also needs to have the gloss and the translation, and the Arabic sentence

2.3.1.3.3 UCP

UCP is used for coordinated constituents of different types (Unlike Coordinated Phrase). In this example, and adjective phrase (ADJP) is coordinated with a prepositional phrase (PP).

```
(NP (NP Al+>asoraY::الأسرى::the+prisoners/captives )
    (UCP (ADJP Al+kuwayotiy~+iyna::الكويتيين::the+Kuwaiti+[masc.pl.gen.] )
        wa-::و-::and
        (PP -min::-من::from
            (NP Al+jinosiy~+At+i::الجنسيات::the+nationality/
                citizenship+[fem.pl.]+[def.gen.]
                Al+>uxoraY::الأخرى::the+other/another/additional ) ) ) ) )
```

الأسرى الكويتيين ومن الجنسيات الأخرى
 Al+>asraY + Al+kuwayotiy~+iyna + wa+min + Al+jinosiy~At+i + Al+>uxoraY
 the+prisoners of war + the+Kuwaitis + and+from + the+nationalities + the+other
The prisoners of war of a Kuwaiti or other nationalities

2.3.1.3.4 LST

LST is used for list markers.

(S and |wa- | و-
 (NP-SBJ this |-h`*A | -ه`ذا)
 (NP-PRD arrangement/organization/preparation |tarotiybu | تَرْتِيبُ
 (NP the+groups/teams/troupes |Alfiraqi | الفِرَقِ))
 :)

(FRAG (NP collection/group/bloc |majomuwEapu | مَجْمُوعَةُ
 (NP the+Atlantic/NATO |Al>aTolasiy~ | الأَطْلَسِيَّ))
 :)
 (FRAG (LST 1)
 -
 (NP Boston |buwsoTun | بُوسْطُنْ
 Celtics |salotiks | سَلْتِكْس)
 (ADVP 6-2))
 (FRAG (LST 2)
 -
 (NP New_Jersey |niyuwjiyroziy | نِيُوجِيزِي
 Nets |nits | نِتْس)
 (ADVP 6-2))
 (FRAG (LST 3)
 -
 (NP Orlando |>uwrolAnoduw | أَوْرْلَانْدُو
 Magic |mAjiyk | مَاجِيك)
 (ADVP 5-3))
 (FRAG (LST 4)
 -
 (NP Philadelphia |fiylAdilofiyA | فِيلَادِلْفِيَا
 76)
 (ADVP 4-3)))

وهذا ترتيب الفرق: مجموعة الأطلسي: 1. بوسطن سلتيكس 6-2. 2. نيو جيرزي نتس 6-2. 3. أورلاندو ماجك 5-3. 4. فلادلفيا 4-3

wa+h`*A + tarotiyb+u + Al+furaq+i + majomuwEap+u + Al+>aTolasiy~ + 1 + buwsoTun +
 salotiks + 6-2 + 2 + niyuwjiyroziy + nits + 6-2 + 3 + >uwrolAnduw + mAjiyk + 5-3 + 4 +
 fiylAdilofiyA + 4-3
 and+this + arrangement + the+teams + group + the+Atlantic + 1 + Boston + Celtics + 6-2 + 2
 + New Jersey + nits + 6-2 + 3 + Orlando + Magic + 5-3 + 4 + Philadelphia + 4-3
*Here is the teams arrangement: The Atlantic Group: 1. Boston Celtics 6-2 2.New Jersey Nits
 6-2 3.Orlando Magic 5-3 4.Philadelphia 76ers 4-3*

2.3.1.3.5 X

X is used to mark any unresolvable problems in the text, errors in the source text and to mark technical pieces of the source text that are not part of the linguistic text.

Any error of this type should be marked as X. For example, tokenization errors should be marked X (for later splitting or merging), passive errors that require a change to the verb token should be marked X, etc.

In this example, fiy and mA are two independent words that are merged together. Inserting the trace of mA is made impossible because of this incorrect tokenization. So the two merged words should be annotated as X, and the rest of the sentence should be annotated as normally as

possible. This tokenization will be corrected (along with the X being removed and the trace inserted) in the process of publishing the annotated corpus.

X should be placed as low as possible in the tree.

```
(S (VP يتحدث ya+taHad~av+u talk
  (NP-SBJ *)
  (PP (SBAR-NOM (WHNP -1 (X فيما fiymA))
    (S (VP (PRT لا la not)
      يخص ya+xuS~+u concern
      (NP-SBJ-1 *T*)
      (NP-OBJ • hu him))))))
```

يتحدث فيما لا يخصه

Says he in what not concern him

He is speaking about things that don't concern him

NB: need transliteration here

Below is an example of segmentation problem that should be Xed. This example is in Iraqi Arabic.

```
(S (VP (X حشونا )
  (NP-SBJ *)
  (PP-LOC في
    (NP بيت
      ( ) ) ) ) )
```

حشونا في بيت واحد

Put they us in house one.

They crammed us into one house.

NB: need transliteration, also in the tree

```
(S (NP-TPC-1 (NP (NP تلك
  (NP الأراضي )
  (SBAR (WHNP-2 التي )
    (S (VP (X يسكنوها
      (NP-SBJ * ) ) ) ) )
  (VP كانت
    (NP-SBJ-1 *T*)
    (PP ب
      (NP الأصل ) )
    (PP-PRD ل
      (NP نا ) ) ) ) ) )
```

تلك الأراضي التي يسكنوها كانت بالأصل لنا

Those the lands that inhabit they her was in the origin to us

Those lands that they inhabit were originally ours.

NB: need transliteration for the full sentence and in the tree,

Long stretches of foreign language text should be Xed. Include all consecutive foreign language words under a single X. However, single words in a foreign language that are used in the expected way in the Arabic sentence should be annotated as a regular part of the tree (i.e., code switching in general should not be Xed).

X is used for passive verbs whenever they are marked as active at the POS level or the reverse. Only use X for the verb inside the VP.

```
(S (VP (X أَسْعَدْتُمْ >asoEado+tum make happy+you)
  (NP-SBJ-1*)
  (NP-OBJ-1*)
  (NP-ADV أَوْقَاتًا >awoqAt+F times)))
```

أَسْعَدْتُمْ أَوْقَاتًا

Be made you happy times.

Have a nice time.

NB: need transliteration

X is also used for unresolvable hamza typo problems, along with any other unresolvable technical problem in the text.

2.3.2 Dashtags! How to use them. How to decide...

2.3.2.1 -SBJ

Marks the subject of the sentence.

- Every S should have exactly one constituent marked as -SBJ, either a direct child of S and sister of the predicate (for equational sentences), or a direct child of the VP.
- In the case of topicalized subjects, the trace is marked as the subject (for example, NP-SBJ-1 *T*), not the moved constituent (which is marked -TPC).
- If the sentence has an empty or pro-dropped subject, it is shown as (NP-SBJ *).
- Possible subjects are NP-SBJ, S-NOM-SBJ, SBAR-NOM-SBJ, S-SBJ, SBAR-SBJ.

```
(S (S (NP-SBJ this |h`*A | |هَذَا)
  (ADJP-PRD natural/normal |TabiyEiy~N |طَبِيعِي))
and |wa- | وَ-
(S (VP we+welcome/receive |-nuraH~ibu |نُرَحِّبُ-
  (NP-SBJ *)
  (PP-CLR with/by |bi- |بِ-
    (NP it/him |-hi | -ه)))))
```

هذا طبيعي ونرحب به

h`*A + TabiyEiy~+N + wa+nuraH~ib+u + bi+hi

this + normal + and+we+welcome + with+him

This is normal and we welcome it

Most subjects will be NP-SBJ noun phrases, either overt or empty. However, clausal subjects are also possible.

(S (VP it/they/she+reward/repay/compensate |tukAfi}u | تكافئ
 (NP-SBJ the+state/country |Ald~awolapu | الدَّوْلَةُ)
 (NP-OBJ (NP the+discipline/precision |AlAinoDibATa | الإنضباط
 and |wa- | و-
 (NP (NP the+fidelity/sincerity |-Al<ixolASa | الإخلاص-
 (PP in |fiy | في
 (NP the+work/action |AlEamal | العمل))))))
 تكافئ الدولة الإنضباط والإخلاص في العمل
 tukAfi}+u + Al+dawolap+u + Al+<inoDibAT+a + wa+Al+<ixolAS+a + fiy + Al+Eamal+i
 reward + the+state + the+discipline + and+the+sincerity + in + the+work
The state rewards for discipline and sincerity at work

In a passive sentence, the passive traces is marked –OBJ, and it is co-indexed with the subject:

(S (VP -nuqil+a:::-نُقِلَ:::be_transferred/be_transported/be_transmitted /it_[verb]
 (NP-SBJ-1 Al+miqoraHiy~:::المقرحي:::the+Megrahi)
 (NP-OBJ-1 *)
 (PP-DIR <ilaY:::إلى:::to/towards
 (NP sijon+K:::سجن:::prison+[indef.gen.]
 <isokutolanodiy~+K:::إِسْكُوتْلَانْدِيَّ:::Scottish+[indef.gen.]))))
 نقل المقرحي الى سجن اسكتلندي
 nuqil+a + Al+miqoraHi + <ilaY + sijon+K + <isokutolanodiy~+K
 transferred + Al+Magraho + to + prison + Scottish
Al-Migrahi was transferred to a Scottish prison

Free relatives can also be direct objects – the –OBJ dashtag should come after the –NOM dashtag:

(S و:::wa-:::and
 (VP (PRT لا:::-la:::no/not/non-)
 نعرف:::na+Eorif+u:::we+know+[ind.]
 (NP-SBJ *)
 (SBAR-NOM-OBJ (WHNP-1 من:::man:::who/whom)
 (S (VP قام:::qAm+a:::undertake/carry_out+he/it_[verb]
 (NP-SBJ-1*T*)
 (PP-CLR ب:::bi-:::with/by
 (NP ها:::-hA:::it/them/her))))))
 ولا نعرف من قام بها
 wa+la+naEorif+u + man + qAm+a + bi+hA
 and+not+know + who + undertake + with+it
We do not know who undertook it

A gerund/masdar or participle with a verbal reading can also be a direct object – the –OBJ dashtag should come after the –NOM dashtag (see section ???? on gerunds and participles):

(S (VP أعلن::>aEolan+a::announce/declare+he/it_[verb]
 (NP-SBJ الاتحاد::Al+{it~iHAd+u::the+union+[def.nom.]})
 (**S-NOM-OBJ** (VP رفض::rafoD+a-::rejection/refusal+[def.acc.]
 (NP-SBJ ه::hu::its/his)
 (NP-OBJ الخصخصة::Al+xaSoxaS+ap+a
 ::the+privatization+[fem.sg.]+[def.acc.]})
 أعلن الإتحاد رفضه الخصخصة
 >aEolan+a + Al+<it~iHAd+u + rafoD+a+hu + Al+xaSoxaSap+a
 announce + the+union + rejection+its + the+privatization
The union announced its rejection for privatization

2.3.2.3 -DTV

Marks the indirect object of a ditransitive verb (like “give”). Every ditransitive verb should have a direct child –OBJ and a direct child –DTV.

- If the verb is passive, either the –OBJ or the –DTV will be on a trace coindexed with the subject.
- Possible indirect objects are NP- DTV, SBAR-NOM-DTV, PP-DTV.

A -DTV argument is an argument that may undergo dative shift (in which the same argument may appear as either an NP or as a PP), whether that argument is a recipient or not. Verbs such as manaHa, Harama, qas~ama are ditransitive verbs that take -DTV arguments, since these arguments can undergo dative shift. All NPs in bold are annotated as -DTV.

منح الدار للرجل/ منح الدار

Need transliteration

Give+he the house **to the man**/give+he **+him** the house
He gave him the house.

حرم الرجل من الميراث/ حرم الرجل الميراث

Need transliteration

Deprive he the man **from the inheritance**/ deprive he the man **the inheritance**
He deprived the man of the inheritance.

قسم الكتاب إلى ثلاثة أقسام/ قسم الكتاب ثلاثة أقسام

Need transliteration

Divided he the book **to three sections**/ Divided he the book **three sections**
He divided the book into three sections.

NB: PP-CLR vs. PP-DTV: when we can turn the PP into NP, it is DTV. But when we cannot turn the PP into NP, it is a PP-CLR.

PP-DTV:

سلب الرجل من حقه/ سلب الرجل حقه

Need transliterationDeny the man **from right his**/ Deny the man **right his***He denied the man what was rightfully his.*

PP-CLR:

اتهم الرجل صديقه بالخيانة *

Need transliterationAccused the man friend his **with betrayal**.*He accused his friend of betrayal.*

(S (VP you+forbid/deprive |taHorimuwna |تَحْرِمُونَا-
 (NP-SBJ *)
 (NP-**OBJ** us |-nA | -نَا)
 (NP-DTV (NP **the+benefit/advantage** |Al<ifAdapa |الإفَادَة
 (PP **from** |min | مِنْ
 (NP (NP **the+law/statutes/regulations** |AlqAnuwni |القَانُونِ
 (NP 161))))))

تحرموننا الإفادة من القانون

taHorimuwna+nA + Al+<ifAdapa+a + min + Al+qAnuwn+i

deprive+us + the+benefit + from + the+law

You deprive us from the benefit of the law

If the indirect object is passivized, the passive trace is marked –DTV:

(S (VP منحت:::munih+at:::be_granted/be_awarded/be_bestowed+it/they/she_[verb]
 (NP-SBJ -1 العائلة:::Al+EA}ilapu:::the+family+[def.nom.]
 (NP-DTV-1 *)
 (NP-**OBJ** التراخيص:::Al+tarAxiyS+a:::the+granting_of_permission/
 licensing+[def.acc.])))

منحت العائلة التراخيص

munih+at + Al+EA}ilap+u + Al+tarAxiyS+a

be_granted + the+family + the+licensing

The family were granted the licening

If there is an alternation for a particular verb between an NP indirect object and a PP alternate, that PP argument is marked –DTV (cf. the English alternation between *I gave Mary a book* and *I gave a book to Mary*).

(S (VP -nu+EoTiy+a:::نُعْطِي::us_to+give/provide+[sub.]
 (NP-SBJ *)
 (PP-DTV li-::ل-::for/to
 (NP -lug+ap+i:::لُغَة::language+[fem.sg.]+[def.gen.]
 Al+faranosiy~+ap+i:::الْفَرَنْسِيَّة:::
 the+French+[fem.sg.]+[def.gen.]))
 (NP-OBJ mawoqiE+a-:::مَوْقِع::position/location/site+[def.acc.]
 (NP -hA:::ها::its/their/her))))
 نُعْطِي- للغة الفرنسية موقعها
 naEoTiy + li+ Al+lugap+i + Al+faranosiy~ap+i + mawoqiE+a+hA
 give + for+language + the+French + position+its
 We give the French language its position

Free relatives can also be indirect objects – the –DTV dashtag should follow the –NOM dashtag:

(S (VP >aEoTaY أعطى)
 (NP-SBJ Almu*iyEu المذيع)
 (SBAR-NOM-DTV (WHNP-1 man من)
 (S (VP AinotaSara انتصر
 (NP-SBJ-1 *T*)
 (PP fiy في
 (NP AlmusAbaqapi المسابقة))))
 (NP-OBJ jA}izapF جائزة)))
 أعطى المذيع من انتصر في المسابقة جائزة
 >aEoTaY + Al+mu*iyE+u + man + <inotasara + fiy + Al+musabaqap+i + jA}izap+F
 gave + the+presenter + who + won + in + the+competition + reward
 The presenter gave a reward for the one who won the competition

2.3.2.4 –CLR

Marks a PP argument of the verb (“closely related”). The use of –CLR depends on the requirement of the verb for a particular prepositional complement – the verb and the preposition should have a particularly close relationship, and the PP-CLR should be obligatory for that sense of the verb.

- Used for the PP alternation with direct object.
- Used for verbs that have an intransitive meaning without the PP and a transitive meaning with the PP-CLR
- Not used for indirect objects – see –DTV above.
- We will eventually have a list of verb-preposition pairs that require –CLR.

(S (VP >adolaY+[null]::أَذْلَى[نُئِل]::provide/present/grant+he/it_[verb]
 (NP-SBJ Al+muHafiz+u::المُحَافِظ::the+governor+[def.nom.]
 Al+mawolaY::المَوْلَى::the+Mawla)
 (PP-CLR bi-::بِ-::by/with
 (NP -taSoriyH+K::تَصْرِيح::declaration/statement+[indef.gen.])))
 أدلى المحافظ المولى بتصريح

>adlaY + Al+muHafiz+u + Al+mawolaY + bi+taSoriyH+K
 provide + the+governor +the+Mawla + with+declaration
The Mawla governor provided a declaration

Verbs can have a PP-CLR when they are intransitive but can have another meaning in which they became “transitive with a preposition” (متعدّ بحرف) **need transliteration here**

جاء الرجل

need transliteration here

Came the man

The man came .

(S (VP جاء
 (NP-SBJ الرجل)
 (PP-CLR ب
 (NP صديق
 (NP ه)))))

جاء الرجل بصديقه

need transliteration here and in the tree

brought the man (with) his friend

The man brought his friend.

Some transitive verbs can be used either with or without the preposition without changing the meaning of the verb; the use with the preposition is kind of a modern usage. So when they are used without the preposition, the argument is NP-OBJ and when they are used with the preposition, the argument is PP-CLR.

(S (NP-SBJ هذا)
 (SBAR-NOM-PRD (WHNP-1 ما)
 (S (VP أكَد
 (NP-SBJ *)
 (PP-CLR علي
 (NP (NP ه)
 (NP-1 *T*)))
 (PP في
 (NP تقرير
 (NP ه))))))))

هذا ما أكد عليه في تقريره

need transliteration here and in the tree

this what insisted he on it in report his

This is what he highlighted on in his report

It is possible to have both a direct object and a PP-CLR complement when it is not possible for the object and the PP to undergo dative shift.

(S (VP اتهم
 (NP-SBJ الرجل)
 (NP-OBJ صديق
 (NP ه))
 (PP-CLR ب
 (NP الخيانة)))))

اتهم الرجل صديقه بالخيانة

need transliteration here and in the tree

accused the man friend his with treason

The man accused his friend with treason.

Certain verbs may have 2 PP-CLRs.

(S (VP سمح
 (NP-SBJ *)
 (PP-CLR ل
 (NP لرجل))
 (PP-CLR ب
 (NP الجلوس)))))

سمح للرجل بالجلوس

need transliteration here and in the tree

allowed he to the man to the sitting

He allowed the man to sit.

2.3.2.4.1 NP-CLR

A noun in NP-OBJ position has the same distribution as any other noun, i.e. it can occur in subject position, object of preposition, it can have a plural form, etc. on the other hand nouns occurring in NP-CLR positions do not have the same distribution, e.g., they do not occur in subject position. They can be a complement a verb, a masdar or a participle with a verbal reading.

The NPs previously annotated as PPs are annotated as NP-CLR whenever they are an essential argument of the verb, eg .

(S (VP تعامل
 (NP-SBJ *)
 (NP-CLR مع
 (NP الأزمة))
 (PP-MNR ب
 (NP كل
 (NP هدوء))))))

تعامل مع الأزمة بكل هدوء

Dealt he with the crisis with all calm

He dealt with the crisis calmly

NB: need transliteration for the full sentence and in the tree

2.3.2.5 -NOM

Marks free relatives (SBAR-NOM), including participles, or verbal gerunds/masdars (S-NOM) in specific nominal positions.

- SBAR-NOM
 - Must have a WHNP as the first child
 - Free relative
 - Functioning in a nominal position (subject of the sentence, object of the sentence, object of a preposition, idafa)
 - -SBJ or -OBJ dashtags can follow the -NOM if appropriate

```
(S (NP-TPC-1 I |>anA | أنا )
  (VP not_be+I |lasotu | لَسْتُ
    (NP-SBJ-1 *T*)
    (PP-PRD from |min- | من-
      (SBAR-NOM (WHNP-2 whom |-man | من-
        (S (VP they_+take/begin |ya>oxu*uwna | يَأْخُذُونَ
          (NP-SBJ-2 *T*)
          أنا لست من من يأخذون
          >anA + lasot+u + min + man + ya>oxu*uwna
          I + not + from + who + take
          I am not from those who take
```

```
(S (NP-SBJ >anA أنا )
  (SBAR-NOM-PRD (WHNP-1 Al~*iy الذي )
    (S (VP naZara نَظَرَ
      (NP-SBJ Al>aEomaY الأعشى )
      (PP <ilaY إلى
        (NP >adab أدب
          (NP (NP iyي )
            (NP-1 *T* ) ) ) ) ) ) )
    أنا الذي نظر الأعشى إلى أدبي
    >anA + Al~*iy + naZara + Al>aEomaY + <ilaY + >adabiY
    I + who + looked + the+blind + to + literature+my
    I am the one to whose literature the blind looked
```

- S-NOM
 - Must have a gerund or masdar as the verb
 - Functioning in a nominal position (subject of the sentence, object of the sentence, object of a preposition, idafa)
 - -SBJ or -OBJ dashtags can follow the -NOM if appropriate

```
(S and |wa- | وَ-
  (VP accuse+he/it |-{it~ahama | تَهَمَّ-
    (NP-SBJ the+president/head/chairman |Alr~a}iysu | الرَّئِيسُ
      the+Venezuelan |Alfinizowiyl~iy~u | الْفِينِزَوِيلِيُّ )
    (NP-OBJ the+striker |AlmuDoribiyna | الْمُضْرِبِينَ )
    (PP-CLR with/by |bi | بِ
      (NP attempt/effort |muHawalapi | مُحَاوَلَةٍ
```


(**S-NOM** (VP robbing/dispossession |salobi | سَلَبَ
 (NP-SBJ *)
 (NP-DTV the+country/countries |AlbilAdi | البلاد)
 (NP-OBJ feast/holiday/festival |Eiyda | عيد
 (NP the+birthday/birth
 |AlmiylAdi | الميلاَد))))))
 واتهم الرئيس الفنزولي المضربين بمحاولة سلب البلاد عيد الميلاد
 wa+<it~aham+a + Al+ra}iys+u + Al+finizowiyliy~+u + Al+muDoribiyina +
 bi+muHAWalap+i + salob+i + Al+bilAd+i + Eiyd+a + Al+miylAd+i
 and + accused + the+president + the+Venezuelan + the+striker + with+attempt +
 dispossession + the+country + feast + the+birthday
*The Venezuelan president accused the stikers of an attempt to deprive the country from
 Christmas*

2.3.2.6 -PRD

Marks the predicate in an equational sentence or clause, and the predicate complements of certain verbs (see section ????? for more on these verbs).

- Must have a -SBJ sister.
- Possible predicates are NP-PRD, ADJP-PRD, PP-PRD, SBAR-NOM-PRD, S-NOM-PRD, ADVP-PRD.

(S (NP-SBJ (NP this/these |-h`*ihi | هـَـذِهِ
 (NP step/stride |xaTowapN | خَطْوَةٌ)
 (ADJP-PRD improbable/unlikely |musotaboEadapN | مُسْتَبَعْدَةٌ))

هذه خطوة مستبعدة

h`*ihi + xaTowap+N + musotaboEadap+N
 this + step + unlikely
This step is unlikely to happen

(S (NP-SBJ it/he |-huwa | هُوَ)
 (NP-PRD matter/issue |>amorN | أَمْرٌ
 rejected/refused |marofuwDN | مَرْفُوضٌ))

هو أمر مرفوض

huwa + >amor+N + marofuwD+N
 it + issue + refused
It is a refused issue

(S (PP-PRD fiy-::في-::in
 (NP (-hA::ها::it/them/her))
 (NP-SBJ Sawot+u::صَوْتُ::voice/sound+[def.nom.]
 (NP tabo\$iy+r+K::تَبْشِير::announcement+[indef.gen.]
 vaqAfiy~+K::ثَقَافِي::cultural/intellectual+[indef.gen.]
 waTaniy~+K::وُطْنِي::national+[indef.gen.]))
 فيها صوت تبشير ثقافي ووطني
 fiy+hA + Sawot+u + tabo\$iy+r+K + vaqAfiy~+K +wa+waTaniy~+K
 in+it + sound + announcement + cultural + and+national
It has a cultural and national announcement sound

If the PP or ADVP predicate also needs an adverbial function dashtag, the –PRD dashtag should come last:

(S wa-::وَ-::and
 (ADVP-LOC-PRD -hunA::هُنَا::here)
 (NP-SBJ (NP Al+\$arikap+u: الشركة : company
 Al+waTaniy~ap+u: الوطنية :national)
 (PP liJ::لِ:for/to
 (NP Al+TAqap+i لطاقة the energy))))
 وهنا الشركة الوطنية للطاقة.
 wa- -hunA Al+\$arikap+u li Al+TAqapi
 and here company national for energy
And here is the National Company of Energy.

-PRD predicates can occur with pro-drop subjects, but an equational sentence can never consist of an overt subject and a null predicate.

(S (S (VP لا la not
 (NP-SBJ غالب gAlib+a victorious)))
 و wa
 (S (VP لا la not
 (NP-SBJ مغلوب magoluwb+a defeated))))
 لا غالب ولا مغلوب
 la gAlib+a wa la magoluwb+a
 not victorious and not defeated.
There exists no winner and no loser.

(S (SBAR-ADV لولا lawola if not
 (S (NP-SBJ *)
 (NP-PRD تشجيعات ta\$ojiyE+At encouragements
 (NP الحكومة Al+Hukuwmap+i the government))))
 (VP (PRT ل la indeed)
 فشل fa\$al+a fail
 (NP-SBJ المشروع Al+ma\$oruwE+u the project)))
 لولا تشجيعات الحكومة لفشل المشروع
 lawola ta\$ojiyE+At Al+Hukuwmap+i la fa\$al+a Al+ma\$oruwE+u
 if not encouragements the government indeed fails the project
If it were not for the government's encouragement, the project would have failed.

NP-PRD as well as ADJP/ADVP/PP-PRD can be topicalized and coindexed with their trace.

(S (ADJP-TPC-1 جميلا
(VP كان
(NP-SBJ وجهها
(ADJP-PRD-1 *T*)))))

جميلا كان وجهها

Beautiful was face her

Her face was beautiful.

NB: need transliteration for the full sentence and in the tree

2.3.2.7 -LOC

Marks a locative modifier. Used for physical location.

- Physical location, not metaphorical.
- Location not direction.
- Not dependent on the variety of physical location (in the kitchen, in the park, in a house, in Paris).

(S (VP -ya+monaE+a:::يمنع:::he/it+prevent/forbid+[sub.]
(NP-SBJ *)
(NP-OBJ (NP Al+Eamaliy~+At+i:::العمليات:::the+practical/
pragmatic+[fem.pl.]+[def.acc.])
(NP-LOC Eabora:::عبر:::across/over/via/be_means_of
(NP Al+Huduwd+i:::الحدود:::the+border/
frontier/limits+[def.gen.])))))

يمنع العمليات عبر الحدود

yamonaE+a + Al+Eamaliy~At+i + Eabora + Al+Hiduwd+i

prevent + the+practical + across + the+borders

He prevents the operations across the borders

(S (VP fight/wage_war+it/they/she |HArabat- | حاربت-
(NP-SBJ Damascus |dima\$oqu | دمشق)
(NP-OBJ (NP (the+fundamentalist |>Al>uSuwliy~apa | الأصولية-
(PP-LOC in |fiy | في
(NP the+areas/zones/territories |AlmanATiqi | المناطق
the+Syrian |Als~uwriy~api | السوريّة)))))

حاربت دمشق الإصولية في المناطق السورية

HArabat + dima\$oq+u + Al+<iSuwliy~ap+a + fiy + Al+manATiq+i + Al+suwriy~ap+i

fight + Damascus + the+fundamentalist + in + the+territories + the+Syrian

Damascus fought the fundamentalism in the Syrian territories

(S (VP تَجَمَّعَ·tajam~aE+a·gather/rally+he/it_[verb]
 (NP-SBJ-1 قُرَابَةُ·qurAb+ap+u·almost/nearly+[fem.sg.]+[def.nom.]
 (NP 70·
 (NP مواطناً·muwATin+AF·citizen+[acc.indef.])))
 (NP-TMP عِنْدَ·Einod+a·with/at+[def.acc.]
 (NP (NP الثامنة·Al+vAmin+ap+i·the+eighth)
 (NP-TMP مساءً·masA'+a·evening+[def.acc.]
 (NP أمس·>amosi·yesterday))))
 (NP-LOC أمام·>amAm+a·in_front_of/facing/vis--vis+[def.acc.]
 (NP مدخل·madoxal+i·entrance/introduction+[def.gen.]
 (NP قصر·qaSor+i·castle/palace+[def.gen.]
 (NP قُرَيْطِمَ·qurayoTim·Qreitim))))))
 تَجَمَّعَ قُرَابَةُ 70 مواطنا عند الثامنة مساء أمس أمام مدخل قصر قريطم
 tajam~aEa + qurAbapu + 70 + muwATinAF + Einoda + Alv~Aminapi + masA'a + >amosi +
 >amAma + madoxali + qasori + qurayoTim
 Gathered + nearly + 70 + citizen + at + eight + evening + yesterday + in front of + entrance +
 castle + Qureitim
 About 70 citizens gathered at eight yesterday evening in front of Qureitim

Metaphorical uses do not receive a –LOC tag.

(S (VP حملت
 (NP-SBJ *)
 (NP-OBJ (ADJP كثيرا
 (PP من
 (NP (الهموم))))))
 (PP في
 (NP قلب
 (NP ها))))))
 حملت الكثير من الهموم في قلبها

Carried she the many of concerns in heart her
 She carried many concerns in her heart

NB: need transliteration for the full sentence and in the tree

vam~apa/ hunA and hunAka are not always referring to a specific location and so the dash tag –LOC is added only in the case they refer to a very specific location. See also section ?????.

(S (ADVP-PRD vam~apa ثَمَّةُ
 (NP-SBJ (NP AlEadydu العديُّ)
 (PP min من
 (NP (NP AlmasA}إلى المسائل)
 (SBAR (WHNP-1 Alaty التي)
 (S (VP (PRT لا)
 taqobalu تقبلُ
 (NP-SBJ-1 *T*)
 (NP-OBJ Alta>jyla التَّأجيلُ))))))))
 ثمة العديد من المسائل التي لا تقبل التأجيل
 Vam~apa + AlEadydu + min + AlmasA}ili + al~aty + lA + taqobalu + Alta>ojiyla
 there + the numerous + of + the issues + that + no + accept + the delaying
 There are so many issues that require immediate attention

2.3.2.8 -DIR

Marks a modifier showing directional movement. Used for movement from or to a physical goal.

- Physical movement, not metaphorical.
- Direction not location – requires movement.
- Not dependent on the variety of physical goal (to the kitchen, to Paris).

(S and |wa- | و-
 (VP be_launched/take_off/be_sent_out/get_started+he/it |-{inoTalaqa | -نُطِّلَقَ-
 (NP-SBJ the+competitor/contestant |AlmutasAbiquwna | المتسابقون)
 (PP-DIR from |min | من
 (NP in_front_of/facing/vis--vis |>amAmi | أمام
 (NP center/headquarters/residence |maqâr~i | مقرّ
 (NP (the+club/association |Aln~Adiy | النادي))))))
 وإطلاق المتسابقون من أمام مقر النادي

wa+<inoTalaq+a + Al+mutasAbiquwna + min + >amAm+i + maqâr~i + Al+nAdiy
 and + start + the+competitors + from + front + headquarter + the+club
The competitors started in the front of the headquarter of the club

(S (VP infiltrate+he/it |tasal~ala | تسلَّلَ
 (NP-SBJ foreigner |>ajonabiy~N | أجنبيّ)
 (PP-DIR from |min | من
 (NP Afghanistan |>afogAnisotAn | أفغانستان))
 (PP-DIR to/towards |<ilaY | إلى
 (NP Iran |<iyrAn | إيران))))
 تسلَّلَ أجنبي من أفغانستان إلى إيران

tasal~al+a + >ajonabiy~+N + min + >afogAnisotAn + <ilaY + <iyrAn
 infiltrate + foreigner + from + Afghanistan + to + Iran
A foreigner infiltrated from Afghanistan to Iran

NB: The annotation of ranges and endpoints follow the English Treebank conventions. The below examples are taken from English PTB guidelines

When the modifier is a range vs. an endpoint: Ranges are contained within a single PP node, and endpoints are separate PPs.

- Range:

```
(VP varied
  (PP (PP from
        (NP 30))
    (PP to
        (NP 53 mg.))))
```

- Endpoints:

```
(VP went
  (PP-DIR from
    (NP Paris)))
```

(PP-DIR to
(NP Dakar)))

2.3.2.9 –MNR

Marks a modifier showing the manner in which the action took place.

- Answers the question “in what manner was xxx done?” or “in what manner did the subject perform the action?” (more specific than “how was xxx done”)
- Instrumental readings are **not** included as –MNR. Instrumental adverbials are marked as –ADV (or plain PP).
- MNR can be used for modifiers of NPs, if the noun is a noun of action, and if the relationship with the modifier is exactly manner.

(S (VP he/it+constitute/form/compose |yu\$ak~ilu | يُشكِّلُ
(NP-SBJ *)
(**PP-MNR** by/with |bi- | بـ
(NP (clarity/plainness | -wuDuwHK | (وُضُوح -
(NP-OBJ (NP call/invitation/supplication |daEowapF | دَعْوَةٌ
(PP to/towards |<ilaY | إلى
(NP sedition/dissension/discord |fitonapK | فِتْنَةٌ
sectarian/factional |TA} ifiy~apK | (طَائِفِيَّةُ |
يشكل بوضوح دعوة الى فتنة طائفية
yu\$ak~il+u + bi+wuDuwH+K + daEowap+F + <ilaY + fitonap+K + TA} ifiy~ap+K
form + with+clarity + call + to + sedition + sectarian
It clearly forms a call for a sectarian sedition

(S (VP affect/influence+he/it |>av~ara | أَثَرَ
(NP-SBJ ذلك|*alika| that
(**NP-MNR** effect/influence/impact |ta>oviyraF | تَأْثِيرًا
serious/grave/dangerous/significant |xaTiyraF | خَطِيرًا
(PP-CLR on/above |EalaY | عَلَى
(NP the+fish |Al>asomAki | (الأسماك |
أثر ذلك تأثيراً خطيراً على الأسماك
>av~ar+a + *`lika + ta>oviyraF + xaTiyraF + EalaY + Al+>asmAk+i
affect + that + impact + serious + on + the fish
That had a serious impact on the fish

(S (VP انقضَّ
(NP-SBJ *)
(PP-CLR على
(NP له))
(S-MNR (VP انقضاض
(NP-SBJ الأسد)
(PP-CLR على
(NP فريسته
(NP له)))))))

انقضَّ عليه انقضاض الأسد على فريسته

Jump he on him jumping the lion on prey his

He attacked him the way a lion would attack its prey

See section ????? on HAl.

See also section ????? on gerunds and participles.

2.3.2.10 –TMP

Marks a temporal modifier, usually a specific reference to a point in time or a span of time. Note that general frequency is not marked as temporal (For example, in “I read that book twice,” *twice* is not -TMP.)

- Possible temporal modifiers are NP- TMP, PP- TMP, S-NOM- TMP, SBAR-NOM- TMP, ADVP- TMP, S- TMP, SBAR-TMP.

```
(S (VP win/be_victorious+he/it |fAza | فاز
  (NP-TMP yesterday |>amosi | أمس)
  (NP-SBJ the+Champville |Al$~Anfoyl | الشانفيل)
  (PP-CLR on/above |EalaY | على
    (NP sons/children |>abonA'i | أبناء
      (NP NOT_IN_LEXICON |AnybAl | انيبال
        Zahle_ |zaHolapF | زحلة))))))
```

فاز أمس الشانفيل على أبناء انيبال زحلة

fAz+a + >amos+a + Al+\$anAfoyl + EalaY + >abonA'i + >aniybAl + zaHolap+F

won + yesterday + the+Champville + on + Children + Aniybal + Zahlah

Yesterday the Champville won against the Children of Anibal zahlah

```
(S (VP وقع::wuq~iE+a::be_signed+he/it_[verb]
  (NP-SBJ-1 الاتفاق::Al+Ait~ifAq+u:: the agreement+[def.nom.])
  (NP-OBJ-1 *)
  (NP-TMP قبل::qabol+a::before+[def.acc.]
    (NP أيام::>ay~Am+K::days+[indef.gen.])))
```

وقع الإتفاق قبل أيام

wuq~iE+a + Al+<it~ifAq+u + qabol+a + >ay~Am+K

be signed + the+agreement + before + days

The agreement was signed days ago

N.B.: Dates are annotated as NPs. Since an NP that is also a date may play many roles in the sentence, dashtags are used with dates as with all other NPs to represent those roles (for example, *January 1st is a good day*, where the date is a subject marked –SBJ; or *My favorite holiday is July 4th*, where the date is a prediate marked –PRD; or *School starts on September 1st*, where the date itself has no dashtag because it is the object of a preposition, but the PP-TMP has the temporal dashtag; or *School starts September 3rd*, where the date is an NP-TMP).

The phrase “the year 2007” is an idafa NP, with 2007 annotated as a complement of year. This NP may be marked with a temporal function (-TMP) if needed, depending on its function in the sentence.

(NP sanap سَنَة
 (NP 2007 2007))
 2007 سَنَة
 sanapa 2007
 (the) year 2007

In a date expression that includes the day, month, and year, the day is annotated as the head of the expression. The month is adjoined to the day, as a temporal modifier marked –TMP. The year is adjoined to the month, as a temporal modifier marked –TMP.

(NP (NP 21)
 (NP-TMP (NP kAnuwn كَانُون
 Al+vAniy الثاني))
 (NP-TMP sanapa سَنَة
 (NP 2006 2006))))
 2006 سَنَة 21 كَانُون الثاني
 21 kAnuwn Al+vAniya sanapa 2006
 21 Kanoun the Second 2006

When the date is written as a numerical expression, the same annotation is followed as when the numbers are separate tokens. If the date is a single token number, it is annotated as an NP.

As usual, whether the top node of the expression is marked –TMP or not depends on its function in the sentence as a whole. The NP complement of a preposition will *never* be marked –TMP, but the PP may be marked –TMP if it plays the role of a temporal modifier.

If the date comes as a single token:

(PP-TMP bi ب
 (NP tAriyxi تاريخ
 (NP 14/01/2002)))
 14/01/2002 بتاريخ
 bi+tAriyxi 14/01/2002
 with (the) date (of) 14/01/2002

If the date comes as separate tokens:

(PP-TMP bi ب
 (NP tAriyxi تاريخ
 (NP (NP 14)
 /
 (NP-TMP (NP 01)
 /
 (NP-TMP 2002))))))
 14/01/2002 بتاريخ
 bi+tAriyxi 14/01/2002
 with (the) date (of) 14/01/2002

2.3.2.11 -PRP

Marks a modifier showing purpose or cause.

- Answers the question “why was xxx done?”
- -PRP can be used for modifiers of NPs, if the noun is a noun of action, and if the relationship with the modifier is exactly purpose.

```
(S (VP (PRT qad::قَدْ::[has/have] )
  taEaT~al+at::تَعَطَّلَتْ::be_hindered/be_interrupted/
                                be_obstructed+it/they/she_[verb]
  (NP-SBJ Al+madAris+u::المَدَارِسُ::the+schools+[def.nom.] )
  (PP-LOC fiy::في::in
    (NP Al+jibAl+i::الجِبَالُ::the+mountains+[def.gen.] ))
  (PP-PRP min::من::from
    (NP jar~A'+i::جَرَاءُ::because_of/as_a_result_of+[def.gen.]
      (NP Al+vuluwj+i::الثلُوجُ::the+snow/ice+[def.gen.]
        wa-::وَ::and
        -Al+jaliyd+i::-الْجَلِيدُ::the+ice+[def.gen.] )))))
  قد تعطلت المدارس في الجبال من جرّاء الثلوج والجليد
qad + taEaT~alat + Al+madAris+u + fiy + Al+jibAl+i + min + jar~A'+i + Al+vuluwj+i +
  wa+Al+jaliyd+i
has + be hindered + the+schools + in + the+mountains + from + because of + the+snow +
  and+the+ice
The schools on the mountains were hindered because of the snow and the ice
```

2.3.2.12 -ADV

Marks ANY OTHER ADVERBIAL MODIFIER.

- Used for any adverbial modifier that is not in one of the more specific categories above.
- PP is assumed to be adverbial by default, so PPs do not need to be marked -ADV if they do not get one of the more specific adverbial function tags.
- ADVP is assumed to be adverbial by default, so ADVPs do not need to be marked -ADV if they do not get one of the more specific adverbial function tags.

```
(S -wa-·and
  (NP-TMP -بَعْدَ·-baEod+a·after+[def.acc.]
    (NP (NP اِذَا·h`*A·this_[masc.sg.]
      (NP الحَادِثِ·Al+HAdiv+i·the+incident/event+[def.gen.])))
  ,
  (VP (VP تَنَاوَلَ·tanAwal+a·deal_with/eat/take/ingest+he/it_[verb]
    (NP-SBJ بُوْش·buw$.Bush)
    (NP-OBJ حَسَاءٌ·HasA'+F·soup+[indef.acc.]
      -wa-·and
      -سَلَطَةٌ·-salaT+ap+F·salad+[fem.sg.]+[indef.acc.]
      (NP-ADV (NP عَوْضًا·EiwaDF·instead)
        (PP (PP غِن·Ean·from/about/of
          (NP الغَشَاءِ·Al+Ea$A'+i·the+dinner/supper+[def.gen.])))
        )
      )
    )
  )
  )
```

وبعد هذا الحادث، تناول بوش حساء وسلطة عوضاً عن العشاء.

wa + baEoda + ha*A + Al+HAdivi + tanAwala + buw\$ + HasA'F + wa+salaTapF +
EiwaDAF + Ean + Al+Ea\$A'i
and + after + this + the+incident + took + Bush + soup + and + salad + instead + on + the +
dinner
And after this incident, Bush ate a soup and salad instead of dinner

Relative clauses modifying verb phrases are annotated as SBAR-ADV children of VP:

(S vum~a ثُمَّ
 (VP bada>a بدأ
 (NP-SBJ Aljumohuwru (الْجُمْهُورُ)
 (PP-CLR bi بِ
 (NP (NP AlhutAfi (الْهُتَافِ)
 wa وَ
 (NP (NP <ilqA'i إلقاء
 (NP AlqAruwrAti (القَارُورَاتِ
 AlfArigapi (الْفَارِغَةِ))
 (NP-DIR wasaTa وَسَطَ
 (NP AlmaloEabi (الْمَلْعَبِ))))))
 (SBAR-ADV (WHNP-1 mA مَا
 (S (VP ADTar~a اضْطَرَّ
 (NP-SBJ-1 *T*)
 (NP-OBJ AlHakama (الْحَكَمَ)
 (PP-CLR <ilaY إِلَى
 (NP <iyqAfi إيقافِ
 (NP AlmubArAti (الْمُبَارَاةِ))))))))))
 ثم بدأ الحمهور بالهتاف وإلقاء القارورات الفارغة وسط الملعب ما اضطر الحكم إلى إيقاف المباراة
 vum~a bada>a Aljumohuwru bi AlhutAfi ou <iloqA'i AlqAruwrAti AlfArighapi mA
 <iDTar~a AlHakama <ilaY <ilgA'I AlmubArAti
 And then the spectators started shouting and throwing empty bottles into the field, which led
 the referee to stop the game

(S-NOM (VP -qiyAm+i-:-قيام-::undertaking/carrying_out/setting_up+[def.gen.]
 (NP-SBJ-2 -hi::ه-::its/his)
 (PP-CLR bi-::ب-::by/with
 (NP (NP ->aEomAl+i-:-أعمال-::actions/activities/work+[def.gen.]
 (NP salob+K::سَلَب-::robbing/dispossession+[indef.gen.]))
 (ADJP Eid~+ap+K::عِدَّة-::several/numerous/
 many+[fem.sg.]+[indef.gen.])))
 (**S-ADV** (VP musotaxodim+AF::مُسْتَخْدِمًا-::employing/using+[acc.indef.]
 (NP-SBJ-2 *)
 (NP-OBJ say~Ar+ap+a::سَيَّارَةً-::car/automobile/
 vehicle+[fem.sg.]+[def.acc.]
 (NP riynuw::رَيْنُو-::Renault))))))
 قيامه باعمال سلبٍ عدةٍ مستخدماً سيارة رينو
 qiyAm+i+hi + bi+>aEomAl+i + salob+K + Eid~ap+K + musotaxodim+AF + say~Arp+a +
 riynuw
 undertaking+his + activities + robbing + numerous + car + Renault
His numerous activities of robbery using a Renault car

2.3.2.13 –HLN

Marks the headline of a news story.

- News story headline, NOT book title.
- Any constituent can be marked –HLN if it is the headline of the story.
- Headlines are often FRAG-HLN
- Used for all news data (newswire and broadcast news, etc.)

(**NP-HLN** man |rajulu | رَجُلُ
 (NP the+religion |Ald~iyni | الدِّينِ
 and |wa- | وَ-
 the+power/authority/rule |-AlsuloTapi | السُّلْطَة -))

رجل الدين والسلطة

rajul+u + Al+diyn+i + wa+Al+suloTap+i
 man + the+religion + and+the+power
The man of religion and power

(**S-HLN** (NP-TPC-1 Al+>asad::الأسد::the+Asad)
 (VP Aisotaqobal+a::اِسْتَقْبَلَ::receive/greet/meet/welcome+he/it_[verb]
 (NP-SBJ-1 *T*)
 (NP-OBJ nA}ib+a::نائب::deputy/delegate/vice-+[def.acc.]
 (NP Al+ra}iys+i::الرئيس::the+president/head/chairman+[def.gen.]
 Al+<iyrAniy~+i::الإيراني::the+Iranian+[def.gen.]))))
 الأسد إستقبل نائب الرئيس الإيراني

Al+>asad + <isotaqobal+a + nA}ib+a + Al+ra}iys+i + Al+<irAniy~+i
 the Asad + receive + vice + the+president + the+Iranian
Al-Asad received the Iranian vice president

See also section ????? on FRAG for the distinction between FRAG and S.

2.3.2.14 –TTL

Marks the title of an artistic work.

- NOT news story headline. –TTL marks the title of books, dissertations, paintings, movies, songs, symphonies, poems, etc.
- Any constituent can be marked –TTL if it is the title of an artistic work.
- –TTL is assumed to be nominal by default. Free relatives or gerunds that are –TTL will also be marked –NOM. But no other constituent marked –TTL should be marked –NOM also (so, do NOT mark PP or ADJP as –NOM when they are also –TTL)

(S (VP carry/bear/transport+he/it | -Hamala | حَمَلَ-
 (NP-SBJ *)
 (NP-TMP (NP this |h`*A | هَذَا |
 (NP the+year |AlEAmā | العام))
 (NP-OBJ (NP title |EunowAna | عُنْوَانُ |
 (NP-TTL "
 (NP the+religions |Al>adoyAni | الأديان
 and |wa- | وَ-
 the+culture/civilization | -AlvaqAfAti | -الثقافات-)
 :
 (NP-ADV between/among |bayona | بَيْنَ
 (NP the+struggle/conflict/fight |AlS~irAEi | المَصْرَاعِ
 and |wa- | وَ-
 the+conversation/dialogue/discussion |
 -AlHiwAri | -الحوار-)
 "))))

حمل هذا العام عنوان الأديان و الثقافات بين الصراع و الحوار

Hamal+a + h`*A + Al+EAm+a + EunowAn+a + Al+>adoyAn+i + wa+Al+vaqAfAt+i +
 bayona + Al+SirAE+i + wa+Al+HiwAr+i
 carry + this + the+year + title + the+religions + and+the+civilizations + between +
 the+conflict + and+the+discussion
This year it is titled "Religions and Civilizations between Clash and Discussion"

(PP لـ·li-·for/to
 (NP (NP ذُور-·-dawor+i·role/part+[def.gen.]
 (NP -ها·-hA·its/their/her))
 (PP في·fiy·in
 (PP-TTL "·
 غلى·EalaY·on/above
 (NP شَفَتِي·-\$af+atayo-·lip+two_[gen.]
 (NP -ي·-ya·my))))))

لدورها في "على شفتي"

li+dawori+hA + fiy + EalaY + \$afatay~a
 for + role + her + in + on + lips + my
For her role in "on my lips"

(S (ADJP-PRD -آخِرُ-·|xir+u·last/end+[def.nom.]
 (NP (NP أَدْوَار-·>adowAr+i·roles/parts+[def.gen.]
 (NP -ها·-hA·its/their/her))
 (NP-SBJ فيلم·fiylm·film/movie
 (S-TTL (VP "
 كُنْتُ·kun+tu·be/was/were+I_[verb]
 (VP أَتَكَلَّمُ·>a+takal~am+u·I+speak/talk/discuss+[ind.]
 (NP-SBJ *)
 (NP-ADV مَعَ·-maE+a-·with+[def.acc.]
 (NP -ها·-hA·it/them/her))
 "))))

آخر أدواوها فيلم "كنت أتكلم معها"

lxiru + >adowAri+hA + fiylm + kunotu + >atakalam~ + maEa+hA
 Latest + role+her + movie + was+I + talk + to+her
Her latest role is in the movie "I was talking to her"

2.3.2.15 –VOC

Marks a vocative constituent, usually a name.

- The vocative should not have any other function in the sentence (they are not subjects or objects or adverbial modifiers, etc.)
- Usually will be NP, because they are usually names.

```
(S (VP qar~+at::قَرَّت::settle_down/remain+it/they/she_[verb]
  (NP-SBJ Eayon+u-::عَيْنُ-::eye+[def.nom]
    (NP -ki::-كِ::your_[fem.sg.] ))
  (NP-VOC (PRT yA::يَا::oh/you )
    kuwayot+u::كُوَيْتُ::Kuwait+[def.nom.] )))
```

قَرَّتْ عَيْنُكَ يَا كُوَيْت

qar~at + Eayon+u+ki + yA+kuwayt+u

settle down + eye+your + oh+Kuwait

Kuwait, congratulations on your security/May thou be rested, Kuwait

The vocative can occur also without the particle yA.

```
(S (VP qar~+at::قَرَّت::settle_down/remain+it/they/she_[verb]
  (NP-SBJ Eayon+u-::عَيْنُ-::eye+[def.nom]
    (NP -ki::-كِ::your_[fem.sg.] ))
  (NP-VOC kuwayot+u::كُوَيْتُ::Kuwait+[def.nom.] )))
```

قَرَّتْ عَيْنُكَ كُوَيْت

qar~at + Eayon+u+ki +kuwayt+u

settle down + eye+your +Kuwait

Kuwait, congratulations on your security/May thou be rested, Kuwait

```
(S (NP-SBJ (NP المَطْلُوبُ·Al+maToluwb+u.the+required +[def.nom.]
  (NP-1 *ICH*))
  (NP-VOC (PRT يَا·yA.oh/you)
    مَعَالِي·maEAliy·His_Excellency
    (NP الوَظِير·Al+waziyr+i.the+minister+[def.gen.])))
  (NP-1 هُوَ·huwa.it/he))
  (SBAR-PRD أَنْ·>an.to
    (S (VP نُعِيدُ·nu+Eiyd+a.we+return/give_back
      (NP-SBJ *)
      (PP إِلَى·<ilaY.to/towards
        (NP المُوَاطِن·Al+muwATin+i.the+citizen))
      (NP-OBJ كَرَامَتُ·karAm+at+a.dignity/generosity
        (NP -هُ·-hu.his/its))))))
```

المطلوب يا معالي الوزير هو أن نعيد إلى المواطن كرامته

Al+maToluwb+u + yA + maEAliy + Al+waziyr+ huwa + >an + nuEiyda + <ilaY +

AlmuwATini + karAmata+hu

the+needed + Oh + honor + the+minister + he + that + return we + to + the+citizen +
dignity+his

What is needed, dear Minister, is to give back to the citizen his dignity

2.3.2.16 –TPC

Marks the fronted element in a sentence with topicalization or in a specific “topic” construction (see below). The goal is to use the trace of the –TPC element to represent where the element originated, and what its role in its “home” position is.

Note that this is a productive construction, and many types of topicalization are possible. It is not possible to list every combination as an example below, but some general guidelines about topicalization are

- In Arabic, these are usually the subject in a sentence where the subject precedes the verb.
- Can also be the object or the predicate, or any argument of the local clause.
- Can also be a fronted argument in an equational sentence (a sentence without a VP). In this case, the pronoun in the surface subject position (or the surface possessive pronoun, etc., etc.) is annotated as a resumptive pronoun. (See section ??????? for more on resumptive pronouns.)
- Can also be a constituent from a lower clause, including fronted modifiers of a lower clause. **Any** fronted element from a lower clause, any long-distance fronting, is marked with –TPC and a trace to the correct location.

(S (NP-TPC-1 the+winds/odors+[def.nom.] Al+riyAH+u | الرياح)
 (VP be/was/were+it/they/she_[verb] kAn+at | كانت
 (NP-SBJ-1 *T*)
 (ADJP-PRD (ADJP north/northern \$amAliy~+ap+F | شمالية
 west/western/westerly garobiy~+ap+F | غربية
 or >aw
 (ADJP north/northern \$amAliy~+ap+F | شمالية
 East/Eastern/Oriental \$aroqiy~+ap+F | شرقية
 الرياح كانت شمالية غربية أو شمالية شرقية
 Al+riyAH+u + kAnat + \$amAliy~+ap+F + garobiy~+ap+F + aw + \$amAliy~+ap+F +
 \$arqiy~+ap+F
 the+wind + was + northern + western + or + northern + eastern
The wind was northern western or northern eastern

(S (NP-TPC-1 results/consequences |natA}ija | نتائج
 (NP actions/activities/work |>aEomAli | أعمال
 (NP institution/organization |mu&as~asapi | مؤسسة
 (NP casino/nightclub |kAziynuw | كازينو
 (NP Lebanon |lubonAna | (لبنان))))
 (VP start/begin+it/they/she |bada>at | بدأت
 (NP-SBJ-1 *T*)
 (PP in |fiy | في
 (NP the+improvement |Alt~aHas~uni | (التحسين))))
 نتائج أعمال مؤسسة كازينو لبنان بدأت في التحسين
 natA}ij+u + >aEomAl+i + mu&as~asap+i + kAziynuw + lubnAn+a + bada>at + fiy +
 Al+taHas~un+i
 results + activities + institute + Casino + Lebanon + started + in + the+improvement
The results of Lebanon Casino activities started to improve

NP-PRD as well as ADJP/ADVP/PP-PRD can be topicalized and be traced back

Topicalization inside SBAR and over S is possible:

(S (NP-SBJ هي)
 (ADJP-PRD حريصة
 (PP على
 (SBAR (NP-TPC-1 الموقف)
 أن
 (S (VP يكون
 (NP-SBJ-1 *T*)
 (ADJP-PRD (موحداً)))))))
 هي حريصة على الموقف أن يكون موحداً

She eager on the attitude that it is united.

She desires that the views are the same.

NB: need transliteration for the full sentence and in the tree

(S (VP (PRT لم)
 نعرف
 (NP-SBJ *)
 (SBAR (NP-TPC-1 غير
 (NP لنا))
 (WHADV-2 كيف)
 (S (VP يفكر
 (NP-SBJ-1 *T*)
 (ADVP-MNR-2 *T*)))))))

لم نعرف غيرنا كيف يفكر

Not know we other us how thinks he

We don't know how other people think.

NB: need transliteration for the full sentence and in the tree

If there is a resumptive pronoun in the VP, the trace of the topicalized constituent is adjoined to the resumptive pronoun. See section ??? on resumptive pronouns.

(S (NP-TPC-1 Albayodapu البَيْضَةُ)
 (VP >akala أَكَلَ)
 (NP-OBJ (NP hA ها)
 (NP-1 *T*))
 (NP-SBJ zayodN زَيْدُ)))

البيضة أكلها زيد

Al+bayodap+u + >akal+a+hA + zayod+N

the+egg + ate+it + Zaid

Zaid ate the egg

Note that separating pronouns/ Damiyr Al-faSol are no longer annotated as resumptive pronouns of topicalization. See section ??? on separating pronouns/ Damiyr Al-faSol.

(S (NP-SBJ (NP (NP -المُعْدَلُ·-Al+muEad~al+u·the+average/rate/mean+[def.nom.]
 العام·Al+EAm~+u·the+general/common/public+[def.nom.]
 (PP-TMP في·fiy·in
 (NP مثل·mivol+i·like/such_as+[def.gen.]
 (NP (NP اذْه·h`A·this_[masc.sg.]
 (NP التاريخ·Al+tAriyx+i·the+date))))))
 (NP هو·huwa·it/he))
 (NP-PRD 462
 (NP ملليمترأ·mllymtrAF·nogloss)))

المعدل العام في مثل هذا التاريخ هو 462 ملليمترا

Al+muEad~alu + Al+EAmu + fiy + mivoli + ha*A + Al+tAiyxi + huwa + 462 +
 millimitorAF

the + average + the + general + in + similar + this + date + he + 462 + ml

The average for the season is 462 ml

-TPC can be used to mark the topic in a sentence with an anticipatory pronoun. See section
 3.5.5 on anticipatory pronouns/Damir al-\$ayon. In this construction, the -TPC function
 tag marks the *topic* not a moved *topicalized* constituent, and as such the topic is not coindexed
 with a trace.

(S (VP قال qAla
 (NP-SBJ-1 *)
 (SBAR أن>an~a-
 (S (NP-TPC هـ-hu)
 (VP سيلقي sayuloqiy
 (NP-SBJ الرئيس Alra{iysu)
 (NP-OBJ خطابا xiTAbAF)
 (PP بـ -bi
 (NP المناسبة AlmunAsabapi))))))

قال أنه سيلقي الرئيس خطابا بالمناسبة

qAl+a + >an~a+hu + sa+yaloqiy + Al+ra{iys+u + xiTAb+AF + bi+Al+munAsabap+i
 said + that + will+deliver + the+president + oration + in+the+occasion
He said that the president will deliver an oration for the occasion

(S (NP-TPC-1 احلام)
 (VP كان
 (NP-SBJ اسم
 (NP ها))
 (NP-PRD-1 *T*)))

اسمها كان احلام

Ahlam was name her

Her name is Ahlam.

NB: need transliteration for the full sentence and in the tree

(S (ADJP-TPC-1 أبرز
 (NP ملامح
 (NP اللقاء)))
 (VP كان
 (NP-SBJ (NP الدعم
 الإيراني
 القوي)
 (PP ل
 (NP اتفاق
 (NP مكة))))
 (ADJP-PRD-1 *T*))))

أبرز ملامح اللقاء كان الدعم الإيراني القوي لاتفاق مكة

NB: need transliteration for the full sentence and in the tree

Most prominent characteristics of the meeting was the support the Iranian the strong to the agreement Mecca

One of the most important features of this meeting was the strong Iranian support to Mecca agreement

(S (PP-TPC-1 على
 (NP كل
 (NP (NP هذه
 (NP النقاط))))
 (VP ركز
 (NP-SBJ *)
 (PP-CLR-1 *T*)
 (PP في
 (NP خطاب
 (NP ه))))))

على كل هذه النقاط ركز في خطابه

On all these points focused in speech his.

He focused on all these points in his speech

NB: need transliteration for the full sentence and in the tree

2.3.2.17 Dashtags not used in ATB

The following dashtags are not used in the ATB. These dashtags are used in the English Treebank, but do not apply for Arabic. Due to technical issues with the annotation tool, a few labels of this kind exist in earlier releases of the Arabic Treebank; they will not be part of any future release.

-LGS
 -BNF
 -EXT
 -CLF

2.3.3 Dashtag decisions:

- PP and ADVP are regarded as adverbial by default (as if they have an invisible –ADV).

- No dashtag at all if they are regular PPs, modifying either the verb or a noun
- If they have a more specific adverbial function, use that dashtag (-TMP, -LOC, -MNR, -PRP)
- If they are actually arguments, use that dashtag to show what the argument role is (-CLR or -PRD)
- SBAR and S are regarded as arguments by default (i.e., no dashtag at all when they are complements of the verb or the noun)
 - No dashtag at all when they are complements of a verb or a noun
 - No dashtag (but in an adjunction structure) when an SBAR is a relative clause modifying a noun
 - Use -NOM for an SBAR that is a free relative or an S that is a gerund/masdar in a nominal position
 - Also add -SBJ or -OBJ if appropriate, following the -NOM
 - So, it should look like SBAR-NOM-SBJ or S-NOM-OBJ, for example
 - SBAR-NOM
 - Must have a WHNP as the first child
 - Free relative
 - Functioning in a nominal position (subject of the sentence, object of the sentence, object of a preposition, idafa)
 - -SBJ or -OBJ dashtags can follow the -NOM if appropriate
 - S-NOM
 - Must have a gerund or masdar as the verb
 - Functioning in a nominal position (subject of the sentence, object of the sentence, object of a preposition, idafa)
 - -SBJ or -OBJ dashtags can follow the -NOM if appropriate
 - Use adverbial dashtags when the S or SBAR has an adverbial function in the VP
 - Use the more specific tags if it is a more specific function (-TMP, -LOC, etc.)
 - Use -ADV if it is not one of the more specific functions
- NP in the S or in the VP must always have a dashtag
 - Use -SBJ, -OBJ, or -PRD for subject, objects and predicates
 - Use adverbial dashtags when the NP has an adverbial function in the VP
 - Use the more specific tags if it is a more specific function (-TMP, -LOC, etc.)
 - Use -ADV if it is not one of the more specific functions
 - The same adverbial tags can also be used if the NP has a truly adverbial function modifying another NP
- Specific dashtags:
 - Every sentence must have a subject labeled -SBJ
 - Every sentence must have a predicate, either a VP or something labeled -PRD
 - PP arguments of the verb should be labeled -CLR or -DTV

2.4 VP arguments and adjuncts

As in the Penn English Treebank, the distinction between arguments and adjuncts of the verb or verb phrase is made through the use of functional dashtags rather than with a structural difference. Both arguments and adjuncts are children of the VP node. No distinction is made between VP-level modification and S-level modification. All constituents that appear before the verb are children of S and sisters of VP; all constituents that appear after the verb are children of VP.

ARGUMENTS of the verb are: NP-SBJ, NP-OBJ, SBAR (no dashtag or -NOM-SBJ/OBJ), S (no dashtag or -NOM-SBJ/OBJ), PP-DTV, PP-CLR (closely/clearly related – a PP the annotator's intuition says is an argument, though it doesn't fall into one of the official argument categories).

ADJUNCTS are: any XP with any other adverbial dashtag, PP (no dashtag), ADVP (no dashtag).

In this example, the NP-SBJ is the subject, NP-OBJ is the object of the verb, and NP-TMP is an adverbial (temporal) NP:

(S (VP سجّلت·saj~al+at·register/record/inscribe+it/they/she_[verb]
 (NP-SBJ (NP "·
 (NP (NP (NP مُنْظَمَةٌ·munaZ~am+ap+u·organization+[fem.sg.]
 (NP حُقُوق·Huquwq+i·rights/law+[def.gen.]
 (NP إنْسان·<inosAn+i·human_being+[def.gen.])))
 (PP مِن·min·from
 (NP أَجْل·>ajol+i·for_sake_of/because_of+[def.gen.]
 (NP لُبنان·lubonAn+a·Lebanon+[indef.gen.])))
 ")
 (NP -LRB-
 سيدر·sydr·NOT_IN_LEXICON
 واتش·wAt\$·NOT_IN_LEXICON
 -RRB-))
 (NP-OBJ (NP {نْتِهَآكَاتٌ·{inotihAk+At+K·violation/contravention
 عِدَّةٌ·Eid~+ap+F·several/numerous/many)
 (PP لـ·li-·for/to
 (NP (NP -الْحُرِّيَّاتِ·-Al+Hur~iy~+At+i·the+freedom)
 وَ·-wa-·and
 (NP -حُقُوق·-Huquwq+i·rights/law+[def.gen.]
 (NP الْإِنْسَانِ·Al+<inosAn+i·the+human_being))))
 (PP-LOC فِي·fiy·in
 (NP لُبنان·lubonAn+a·Lebanon+[indef.gen.])))
 (NP-TMP مَطْلَع·maTolaE+a·beginning/rise+[def.acc.]
 (NP السَّنَةِ·Al+san+ap+i·the+year
 الْجَدِيدَةِ·Al+jadiyd+ap+i·the+new/modern))))
 سجّلت منظمة حقوق انسان من أجل لبنان سيدر واتش انتهاكات عدّة للحريات وحقوق الإنسان في لبنان مطلع السنة الجديدة
 Saj~alat + munaZ~amapu + Huquwqi + <insAni + min + Ajoli + lubonAn + <inotihAkAtK +
 Eid~apK + li+Al+Hur~iyAti + wa+Huquwqi + Al+>inosAni + fiy + lubonAn + maTlaEa
 + Alsanapi + Aljadiydapi
 Register + organization + rights + human being + for + sake + Lebanon + Cedar + Watch +
 violations + numerous + freedoms + and + rights + human beings + in + Lebanon +
 beginning + the+year + the+new
*The Human Rights organization, Cedar Watch, registered a number of human rights and
 liberties violation in Lebanon in the beginning of the new year.*

2.5 NP arguments and adjuncts

The argument/adjunct distinction is shown structurally inside NPs. Argument constituents are children of NP, sister to the head noun: (NP head (NP argument)). Adjunct constituents are sister to the NP that contains the head noun, child of the NP that contains both: (NP (NP head) (NP adjunct)).

Arguments are genitive, possessive, or (for deverbal head nouns) clausal constituents that would be arguments of the verb that the noun derived from.

Adjuncts are all other modifiers of the NP, and include ALL NP-internal PPs.

NP with NP argument – the NP argument (NP *tadoriyb* "training" is a sister of the head noun *TA}irapu* "aircraft" itself:

(NP-SBJ (NP طائرة::TA}ir+ap+u::aircraft/airplane+[fem.sg.]+[def.nom.]
 (NP تدريب::tadoriyb+K::training/coaching/practice+[indef.gen.]))
 (ADJP عسكرية::Easokariy~+ap+K::military/army+[fem.sg.]+[indef.gen.]))
 طائرة تدريب عسكرية

TA}irap+u + tadoriyb+K + Easokariy~ap+K
 aircraft + training + military
Military training aircraft

NP with PP adjunct – the NP containing the head noun (NP *Al+mu\$ar~adi+iyona*) "the homeless" and the PP adjunct (PP-LOC *fiy...*) "in..." are sisters, both children of a containing NP:

(NP-OBJ (NP مجمعات::mujam~aE+At+i::compounds/buildings+[fem.pl.]+[def.acc.]
 (NP كهوف::kuhuwf+K::caves/cavities+[indef.gen.]))
 (PP في::fiy::in
 (NP شرق::\$aroq+i::east/East+[def.gen.]
 (NP أفغانستان::>afogAnisotAn+a::Afghanistan+[indef.gen.]))))
 مجمعات كهوف في شرق أفغانستان

mujam~aEAt+i + kuhuwf+K + fiy + \$aroq+i + >afogAnisotAn
 compounds + caves + in + east + Afghanistan
Compounds of caves in the east of Afghanistan

2.6 Clause combinations

As in other languages, there are three ways in which clauses can combine in Arabic:

- subordination
- coordination with a conjunction
- coordination without a conjunction

2.6.1 Coordination

Coordination is done as adjunction (Z (Z) and (Z)); coordination has the same structure at all phrase levels.

2.6.1.1 Coordination with a conjunction

See also section ?????? on Coordination in the Treebank in general.

We consider that two or more phrases are coordinated when there is a conjunction between them (either a word with the POS tag CONJ or a multi-word conjunction forming a CONJP (see section ?????? for a list of conjunctions).

Coordinated Ss:

(S (S (NP-TMP قبل::qabola::before
 (NP بضعة::biDoE+ap+i::some/several+[fem.sg.]+[def.gen.]
 (NP أيام::>ay~Am+K::days+[indef.gen.])))
 (VP استولت::{isotawol+at::overpower/capture+it/they/she_[verb]
 (NP-SBJ اسرائيل::<isorA}iyl+u::Israel+[def.nom.]
 (PP-CLR على::EalaY::on/above
 (NP (NP سفينة::safiyn+ap+K::ship/vessel+[fem.sg.]+[indef.gen.]
 (ADJP محملة::muHam~al+ap+K::loaded/burdened+
 [fem.sg.]+[indef.gen.]
 (PP ب::bi-::with/by
 (NP الاسلحة::-Al+>asoliH+ap+i::the+weapons+
 [fem.sg.]+[def.gen.]))))))
 و::wa-::and
 (S (VP زعمت::-zaEam+at::allege/claim+it/they/she_[verb]
 (NP-SBJ *)
 (SBAR ان::>an~a-::that
 (S (NP-SBJ ها::-hA::it/they/she)
 (ADJP-PRD ايرانية::<iyrAniy~+ap+u::Iranian+[fem.sg.]+[def.nom.]
 (NP المصدر::Al+maSodar+i::the+source+[def.gen.])))
 قبل بضعة أيام استولت إسرائيل على سفينة محملة بالأسلحة و زعمت أنها إيرانية المصدر
 qabol+a + biDoEap+i + >ay~Am+K + <isotawlat + <isorA}iyl+u + EalaY + safiynap+K +
 muHam~alap+K + bi+Al+>asoliHap+i + wa+zaEamat + >an~a+hA + <iyrAniy~ap+u +
 Al+maSodar+i
 before + several + days + capture + Isreal + on + ship + loaded + with+the+weapons +
 and+claim + that+it + Iranian + the+source
Several days ago, Isreal captured a ship loaded with weapons and claimed she was from Iran

Complement clauses coordinated under a single complementizer (eg >an~a) should be annotated as coordinated Ss under a single SBAR, i.e., (SBAR >an~a (S (S...) wa (S...)))

(S و wa and
 (VP نقلت naqal+at reported
 (NP-SBJ إذاعة محلية <i*AE+ap+N maHal~iy~+ap+N local radio station)
 (PP عن Ean from
 (NP (NP مسئولين maso&uwl+iyina officials)
 (PP في fiy in
 (NP الشرطة Al+\$uroT+ap+i the police))))
 (SBAR أن >an~a that
 (S (S (NP-TPC-1 21
 (NP منزلا manozil+AF house))
 (VP احترقت {iHotaraq+at burnt
 (NP-SBJ-1 *T*)))
 و wa and
 (S (VP لحقت laHiq+at happened
 (PP-CLR ب bi to
 (NP ثلاثة أخرى valAv+ap+F
 >uxoraY three others))
 (NP-SBJ اضرار >aDorAr+N damage))))))
 ونقلت إذاعة محلية عن مسئولين في الشرطة أن ٢١ منزلا احترقت و لحقت بثلاثة أخرى اضرار

NB: need transliteration for the full sentence

And reported radio station local from officials in the police that 21 house burnt and happened to three others damage.

A local radio station reported from officials in the police that 21 houses burnt down and three others were damaged.

NP coordination:

Single-word nouns are coordinated flat.

(NP qiTAE+At+i::قطاعات::sectors/sections+[fem.pl.]+[def.acc.]
 (NP Al+AiqotiSAd+i::الاقتصاد::the+economy/saving+[def.gen.]
 wa-::و-::and
 -Al+siyAH+ap+i::السياحة::the+tourism+[fem.sg.]+[def.gen.]))
 قطاعات الإقتصاد والسياحة
 qitAEAt+i + Al+<iqotiSAd+i + wa+Al+siyAHap+i
 sectors + the+economy + and+the+tourism
Sectors of economy and tourism

Nouns with modifiers or complements must be NP constituents. NPs are coordinated as adjunction.

(NP (NP مَدِينَة·madiyn+ap+u·city+[fem.sg.]+[def.nom.]
 (NP زَامْبُوَانْغَا·zAmobuwAnogA·Zamboanga))
 وَ·wa·and
 (NP جَزِيرَة·-jaziyar+ap+u·island+[fem.sg.]+[def.nom.]
 (NP بَاسِيلَان·bAsiylAn·Basilan)))

مَدِينَة زَامْبُوَانْغَا وَ جَزِيرَة بَاسِيلَان

madiynapu + zAmobuwAnogA + wa+jaziyrapu + bAsiylAn
 city + Zamboanga + and + island + Basilan
The city of Zamboanga and the island of Basilan

This is an example of coordinated SBARs:

(SBAR (SBAR (WHNP-3 التي::Al~atiy::which/who/whom_[fem.sg.])
 (S (VP تعمل::ta+Eomal+u::it/they/she+work/function/act+[ind.]
 (NP-SBJ امريكا::>amiyrokA::America)
 (PP-CLR على::EalaY::on/above
 (NP اشاعت::<i\$AE+at+i::spreading/
 circulation+[fem.sg.]+[def.gen.]
 (NP (NP ها::-hA::its/their/her)
 (NP-3 *T*)))))
 وَ::wa::and
 (SBAR (WHNP-4 التي::-Al~atiy::which/who/whom_[fem.sg.])
 (S (VP تبنا::taban~A+-LRB-null-RRB--::adopt+he/it_[verb]
 (NP-OBJ (NP ها::-hA::it/them/her)
 (NP-4 *T*))
 (NP-SBJ (NP رئيس::ra}iys+u::president/head/chairman+[def.nom.]
 (NP السلطة::Al+suloT+ap+i::the+power/
 authority/rule+[fem.sg.]+[def.gen.])
 (NP السيد::Al+say~id+u:: the+Sir/Mr./Mister+[def.nom.]
 ياسر::yAsir::Yasser/Yasir
 عرفات::EarafAt::Arafat))))
 التي تعمل أمريكا على إشاعتها والتي تبناها رئيس السلطة السيد ياسر عرفات

Al~atiy + taEomal+u + >amoriyK + EalaY + <i\$AEat+i+hA + wa+Al~atiy + taban~A+hA
 + ra}iys+u + Al+suloTap+i + Al+say~id+i + yAsir + EarafAt
 which + work + America + on + spreading+it + and+which + adopt+it + president +
 the+authority + the+Mr. + Yasser + Arafat
Which America is working on to spread and is adopted by the President of the authority, Mr. Yasser Arafat.

The coordination of relative clauses where the second conjunct does not have an overt relative pronoun is nonetheless treated as a case of SBAR coordination, with a (WHNP 0) in the second conjunct.

(NP (NP السياسة need transliteration the policy)


```

(SBAR (SBAR (WHNP-3 التي:::Al~atiy::which/who/whom_[fem.sg.]
  (S (VP تبنا:::taban~A+-LRB-null-RRB--::adopt+he/it_[verb]
    (NP-OBJ (NP ها:::-hA::it/them/her)
      (NP-3 *T*))
    (NP-SBJ (NP رئيس:::ra}iys+u::president/head +[def.nom.]
      (NP السلطة:::Al+suloT+ap+i: authority))
    (NP السيد:::Al+say~id+u:: the+Sir/Mr.+[def.nom.]
      ياسر:::yAsir::Yasser/Yasir
      عرفات:::EarafAt::Arafat))))
  و:::wa-::and
  (SBAR (WHNP-4 0 )
    (S (VP تعمل:::ta+Eomal+u::it/they/she+work/function/act+[ind.]
      (NP-SBJ اميركا:::>amiyrokA::America)
      (PP-CLR على:::EalaY::on/above
        (NP اشاعت:::<i$AE+at+i-::spreading
          (NP (NP ها:::-hA::its/their/her)
            (NP-4 *T*))))))))

```

السياسة التي تبناها رئيس السلطة السيد ياسر عرفات والتي تعمل أمريكا على إشاعتها

???need transliteration -Al~atiy taban~A -hA ra}iys+u Al+suloT+ap+i Al+say~id+u yAsir
EarafAt wa ta+Eomal+u >amiyrokA EalaY <i\$AE+at+i- -hA

The policy that adopted it president authority Mr. Yasser Arafat and that works USA on spreading it

The policy that is adopted by Yassar Arafat, President of the Palestinian Authority and that USA is working on propagating.

When constituents of different types are coordinated, the outer coordination-level node label is UCP (Unlike Coordinated Phrase). Any shared function tags are put on the UCP label, and not on the lower labels.

```

(SBAR (WHNP-1 التي·Al~atiy·which/who/whom_[fem.sg.]
  (S (VP (PRT س·sa·will)
    -تُقام·-tu+qAm+u·it/they/she+take_place/be_established+[ind.]
    (NP-SBJ-1 *T*)
    (NP-OBJ-1 *)
    (PP-LOC في·fiy·in
      (NP القاهرة·Al+qAhir+ap+i·the+Cairo))
    (UCP-TMP (NP (NP {عُتِبَ رَأً{iEotibAr+AF·beginning_[on])
      (PP من·min·from
        (NP الأحد·Al+>aHad+i·the+Sunday+[def.gen.])))
      و·wa·and
      (PP لـ·-li·for/to
        (NP مُدَّة·-mud~+ap+i·period_of_time+[fem.sg.]
          (NP أُسْبُوعٍ·>usobuwE+K·week+[indef.gen.]))))))
    التي ستقام في القاهرة اعتباراً من الأحد ولمدة أسبوع

```

Al~atiy + sa + tuqAmu + fiy + Al+qAhirapi + {iEotibArAF + min + Al+>aHadi + wa+
li+mud~api + >usobuwEK

That + will + be held + in + the+Cairo + beginning + from + the+Sunday + and + for +
period + week

Which will be held in Cairo beginning this Sunday and will go on for a week

If different types of S (i.e., S and SQ) are coordinated, they are coordinated under S.

If different types of SBARs (i.e., SBAR and SBARQ) are coordinated, they are coordinated under SBAR.

However, if an S type is coordinated with an SBAR type, they are coordinated under UCP.

2.6.1.1.1 Coordination with only “>amo IA” (*or not*)

The collocation >amo IA 'or not' will be done as 'or (YP (FRAG not))' conjoined to whatever YP is being potentially negated by 'not'.

```
(S (S (VP خرج
      (NP-SBJ *)))
  أم
  (S (FRAG (PRT لا))))
```

خرج أم لا

NB: need transliteration for the full sentence and in the tree

Go out or not

He went out or not

```
(ADJP (ADJP شجاع
        أم
        (ADJP (FRAG (PRT لا)))))
```

شجاع أم لا

NB: need transliteration and gloss for the full sentence and in the tree

Courageous or not

2.6.1.2 Coordination without a conjunction

Coordination of non-clausal constituents can occur without a conjunction in some constructions.

This kind of coordination between sentences is quite rare in Arabic. Most of the time, when there is no conjunction between sentences, they should be separate sentences except when the second sentence is kind of repetition and explanation of the first sentence

```

(S (S (S و::wa-::and
      (VP ظل::-Zal~+a::remain/continue+he/it_[verb]
        (NP-SBJ جان::jAn::Jean/Jan
          غانم::gAnim::Ghanim/Ghanem)
        (PP-PRD ك::ka::as/like
          (SBAR-NOM (WHNP-1 ما::mA::that/which)
            (S (NP-SBJ هو::huwa::it/he)
              (NP-PRD-1 *T*)))))
        ,::nogloss
      (S (VP (PRT لم::lam::did_not)
        يتغير::ya+tagay~ar+o::he/it+be_changed/be_modified+[jus.]
        (NP-SBJ *) )))

```

وظلَّ جان غانم كما هو لم يتغيَّر

wa+Zal~+a + jAn + gAnim + ka+mA + huwa + lam + yatagay~ar+o
 and+remain + Jean + Ghanim + as+which + he + did not + be changed
Jean Ghanim remained as he was; he did not change

```

(S (S vum~a::ثم::then/thereupon
      (VP taHad~av+a::تحدَّث::speak/discuss+he/it_[verb] 6
        (NP-SBJ Al+nA}ib+u::النائب::the+deputy/delegate/vice-+[def.nom.]
          junoblAT::جُنْبَلَاط::Jumblatt )))
      (S (VP qAl+a::قال::said+he/it_[verb]
        (NP-SBJ *)

```

ثمَّ تحدث النائب جنبلاط قال

vum~a + taHad~av+a + Al+nA}ib+u + junoblAT + qAl+a
 then + spoke + the+deputy + Jumblatt + said
Then, the deputy Jumblatt spoke, he said

This construction is quite rare. In most cases, Ss should be annotated as coordination (with a conjunction) or as separate sentences (with no punctuation):

1. If there is a conjunction between Ss, the Ss must be annotated with a coordination structure.
2. If there is no conjunction and no final punctuation between Ss, they should be annotated as independent sentences (as if there were final punctuation between them) – unless they are an instance of the very rare coordination without conjunction repetition structure above. This
 - a. Note that the annotation of independent Ss with no final punctuation between them will result in multiple S-expressions on a single line (from a single “paragraph” or TOP node) in the .tree files.

2.6.1.2.1 When sentences are NOT coordinated

Personal pronouns or demonstratives in the second sentence that are related to the previous sentence are not indications of coordination. Sentences with only this type of connection should be annotated as separate sentences, NOT as coordination without a conjunction.

(S (wa-::and
 (PP-LOC في::-fiy::in
 (NP (duwal+i::states/countries+[def.gen.]
 (NP (العلم::Al+EAlam+i::the+world+[def.gen.])))
 (VP (tasA'al+a::ask/wonder+he/it_[verb]
 (NP-SBJ كثيرون::kaviyr+uwna::many/much/numerous+[masc.pl.]
 (SQ (PRT ما::mA::what/which)
 (NP-SBJ هو::huwa::it/he)
 (NP-PRD "::::nogloss
 الپرتزل::Al+brytzl::the+NOT_IN_LEXICON
 "::::nogloss)
 ,::nogloss)

وفي دول العالم تساءل كثيرون ما هو البريتزل

wa+fiy + duwal+i + Al+EAlam+i + tasA'al+a + kaviyruwna + mA + huwa + Al+boriyyoziyl
 and+in + countries + the+world + wonder + many + what + it + the+pretzel
Many people around the world wondered what the pretzel was

(S (NP-SBJ ذلك::*`lika::that)
 (SBAR-PRD ان::>an~a::that
 (S (NP-TPC-3 الاميركيين::Al+>amiyrokiy~+iyna::
 the+American+[masc.pl.acc.]
 كافة::kAf~+ap+F::all+[fem.sg.]+[indef.acc.]
 (VP (يعرفون::ya+Eorif+uwna-::they_[people]+know+[masc.pl.]
 (NP-SBJ-3 *T*)
 (NP-OBJ ه::hu::it/him))))

ذلك أنَّ الأميركيين كافة يعرفونه

*`lika + >an~a + Al+>amiyrokiy~iyna + kAf~ap+F + yaEorifuwna+hu
 like that + that + the+Americans + altogether + know+it
that is because all the Americans knew it

(S (wa-::and
 (NP-SBJ رحمة::raHom+ap+u::compassion/mercy+[fem.sg.]+[def.nom.]
 (NP (اسرائيل::<isorA}iyl+a::Israel+[indef.gen.])))
 (ADJP-PRD واسعة::wAsiE+ap+N::wide/extensive/broad+[fem.sg.]+[indef.nom.]
 ::,::nogloss)

ورحمة إسرائيل واسعة

wa+raHomap+u + <isorA}iyl+a + wAsiEap+N
 and+compassion + Israel + extensive
Israel's compassion is extensive

(S (VP (يعرف::ya+Eorif+u::he/it+know+[ind.]
 (NP-OBJ ذلك::*`lika::that)
 (NP-ADV جيداً::jay~id+AF::well+[acc.indef.]
 (NP-SBJ لبنان::lubonAn+u::Lebanon+[def.nom.]
 و::wa-::and
 العرب::Al+Earab+u::the+Arabs+[def.nom.])))
nogloss)

يعرف ذلك جيّداً لبنان و العرب

yaEorif+u + *`lika + jay~id+AF + lubonAn+a wa+Al+Earab+u
 know + that + well + Lebanon + and+the+Arabs
Lebanon and the Arabs know that well

A syntactically related adverbial clause should be annotated as S-ADV (or S-MNR, if appropriate). It should not be annotated as coordination.

```
(S (VP -تَظَاهَرَ-taZahar+a.manifest/demonstrate+he/it_[verb]
  (NP-SBJ-1 (NP سُكَّانٌ.suk~An+N.residents/inhabitants+[indef.nom.])
    (PP مِن.min.from
      (NP التبت .Al+tybt.the+NOT_IN_LEXICON)))
  (NP-TMP الْاَحَد .Al+>aHad+a.the+one+[def.acc.] وَ
    -wa-and
    {ثَنَيْنِ}-Al+{ivonayoni.the+Monday}
  (PP-LOC فِي.fiy.in
    (NP نِيُودِلْهِى.niyuwdilohiy.New_Delhi))
  (S-ADV (VP مُطَالِبِينَ.muTAlib+iyna.demanding+[masc.pl.acc.]
    (NP-SBJ-1 *)
    (NP-OBJ فَاجِبَايِي.fAjobAyiy.Vajpayee)
    (PP-CLR بِ-bi-by/with
      (NP (NP حَضْرًا-HaD~+i.incitement/instigation
        (NP الْحُكُومَةُ .Al+Hukuwm+ap+i.the+government
          الصَّيْنِيَّة .Al+Siyniy~+ap+i.the+Chinese))
        (PP عَلَى .EalaY.on/above
          (NP فَتْح .fatoH+i.opening/beginning
            (NP (NP حِوَار .HiwAr+K.dialogue)
              (NP-ADV مَعَ.maE+a.with+[def.acc.]
                (NP الدالاي .Al+dAlAy.the+NOT_IN_LEXICON
                  لاَما.lAmA.the+NOT_IN_LEXICON
                  ))))))))))))
```

تظاهر سكان من التبت الأحد والإثنين في نيودلهي مطالبين فاجبايي بحضّ الحكومة الصينية على فتح حوار مع الدالاي لاما
 taZAhara + suk~AnN + min + Al+tiybAt + Al+>aHada + wa+Al+>ivonayoni + fiy + niyuw
 + delhiy + muTAlibiyina + fAjobAyiy + bi+HaD~i + Al+Hukuwmapi + Al+Siyniy~api +
 EalaY + fatoHi + HiwArK + maEa + Al+dAlAy + lAmA
 demonstrated + inhabitants + from + the + Tibet + the + Sunday + and + the + Monday + in
 + New + Delhi + asking + Vajpayee + with + encouraging + the + government + the +
 Chinese + with + opening + dialog + with + the + Dalai + Lama
*A number of Tibet inhabitants demonstrated Sunday and Monday in New Delhi, asking
 Vajpayee to encourage the Chinese government to open a dialog with Dalai Lama*

Relative clauses with zero/null relative pronouns should also not be confused with coordinated sentences. Relative clauses should be adjoined to the NP they modify.

(S (NP-TPC-1 وزارة::wizAr+ap+a::ministry+[fem.sg.]+[def.acc.]
 (NP التربية::Al+tarobiy+ap+i::the+education/
 pedagogy/breeding+[fem.sg.]+[def.gen.]))
 (VP ستنظم::sa+tu+naZ~im+u::will+it/they/she+
 arrange/organize/regulate+[ind.]
 (NP-SBJ-1 *T*)
 (NP-OBJ (NP احتفالاً::iHotifAl+AF::celebration/ceremony+[acc.indef.])
 (SBAR (WHNP-2 0)
 (S (VP يشارك::yu+\$Arik+u::he/it+participate_with/
 share_with+[ind.]
 (PP-CLR في::fiy-::in
 (NP (NP ه::-hi::it/him)
 (NP-2 *T*))
 (NP-SBJ نحو::naHowa::towards/approximately
 (NP 25::25::nogloss
 (NP الف ::>alof+a::thousand+[def.acc.]
 (NP تلميذ::tilomiy*+K::
 student/pupil+[indef.gen.]
))))))))
 وزارة التربية ستنظم إحتفالاً يشارك فيه نحو ألف تلميذ
 wizArap+a + Al+tarobiyap+i + sa+tunaZ~im+u + <iHtifAl+AF + yu\$Arik+u + fiy+hi +
 naHowa + >alof+a + tilomiy*+K
 ministry + the+education + will+organize + celebration + participate + in+it + approximately
 + thousand + pupil
*The Ministry of Education will organize a celebration in which approximately a thousands of
 students participate*

Headlines should be marked –HLN, and they should not be annotated as coordination unless there is an overt conjunction. See section ??? for more on –HLN.

2.6.1.3 Initial conjunction

Sentence-initial conjunction (POS tag = CONJ; the most frequent is the wa) is treated as having a discourse rather than coordinating function, and as such is put inside the S. However, all other instances of conjunctions are treated as true coordination. See section ???? for the prepositional use of wa.

(S و::wa-::and
 (VP يشكل::-yu+\$ak~il+u::he/it+constitute/form/compose+[ind.]
 (NP-SBJ الرفض::Al+rafoD+u::the+rejection/refusal+[def.nom.])
 (NP-OBJ (NP أمراً::>amor+AF::matter/issue+[acc.indef.])
 (ADJP طبيعياً::TabiyEiy~+AF::natural/normal+[acc.indef.] .
 و::wa-::and
 متوقعاً::-mutawaq~aE+AF::expected/
 anticipated+[acc.indef.])))
 ويشكل الرفض أمراً طبيعياً ومتوقعاً
 wa+yu\$ak~il+u + Al+rafoD+u + >amor+AF + TabiyEiy~+AF + wa+mutawaq~aE+AF
 and+form + the+refusal + matter + normal + and+expected
Refusal is a normal expected matter

(S fa-::ف-::and/so
 (NP-TPC-1 (NP -Al+Eabod+u::-العبد::the+Abd+[def.nom.])
 (PP fiy::في::in
 (NP Al+suloT+ap+i::السُلْطَة::the+power/authority/
 rule+[fem.sg.]+[def.gen.])))
 (VP (PRT lA::لا::no/not/non-)
 ya+sotaTiyE+u::يَسْتَطِيعُ::he/it+be_able/be_capable+[ind.]
 (NP-SBJ-1 *T*)
 (NP-OBJ samAE+a::سَمَاعَ::hearing/listening+[def.acc.]
 (NP -Al+Haq~+i::-الحَقَّ::the+truth/right+[def.gen.]))))
 فالعبد في السلطة لا يستطيع سماع الحقّ
 fa+Al+Eabod+u + fiy + Al+suloTap+i + lA+yasotaTiyE+u + samAE+a + Al+Haq~+i
 so+the+person + in + the+power + no+be able + hearing + the+truth
 So the person in power can not hear the truth.

(S vum~a::ثُمَّ::then/thereupon
 (VP gAdar+at::غَادَرَتْ::leave/depart+it/they/she_[verb]
 (NP-SBJ *)
 (PP-DIR fiy::في::in
 (NP Ait~ijAh+i::إِتْجَاهَ::direction/course+[def.gen.]
 (NP Al+>arADiy::الأَرَاضِي::the+territories/land
 Al+muHotal~+ap+i::الْمُحْتَلَّةَ::the+occupied+[fem.sg.]+[def.gen.])))
 (NP-TMP Einoda::عِنْدَ::with/at
 (NP Al+sAE+ap+i::السَّاعَةِ::the+hour/time+[fem.sg.]+[def.gen.]
 (NP 11.25::11.25::))))
 11:25 ثم غادرت في إتجاه الأراضي المحتلة عند الساعة
 vum~a + gAdarat + fiy + <it~ijAi+i + Al+>arADiy + Al+muHotal~ap+i + Einoda +
 Al+sAEap+i + 11:25
 thereupon + depart + in + direction + the+territories + the+occupied + at + the+time + 11:25
 Thereupon, at 11:25, they departed towards the occupied territoruies

2.6.2 Subordination: The use of SBAR

Complement clauses of verbs that begin with a complementizer (overt or null) are annotated as SBAR.

(S و::wa-::and
 (VP يعتقد:::-ya+Eotaqid+u::he/it+believe+[ind.]
 (NP-SBJ (NP (NP الكاتب::Al+kAtib+u::the+writer/author+[def.nom.]
 (PP في::fiy::in
 (NP صحيفة::SaHiyf+ap+i::newspaper+[fem.sg.]+[def.gen.]
 (NP ":::::nogloss
 النيويورك::Al+niyuwyuwrok::the+New_York
 تايمس::tAyoms::Times
 ":::::nogloss))))
 (SBAR أن::>an~a::that
 (S (NP-TPC-1 (NP المعسكر::Al+muEasokar+a::the+camp/
 encampment+[def.acc.]
 الموالي::Al+muwAliy+a::the+partisan/
 sympathizer+[def.acc.]
 (PP ل::li-::for/to
 (NP اسرايل::-<isorA}iyl+a::
 Israel+[indef.gen.])))
 (VP تجاوز::tajAwaz+a::exceed/disregard+he/it_[verb]
 (NP-SBJ-1 *T*)
 (NP-OBJ الحد::Al+Had~+a::the+extent/limit/level+[def.acc.]
 المقبول::Al+maqobuwl+a::the+accepted/welcome/
 admitted+[def.acc.])))
 .:::::nogloss)
 ويعتقد الكاتب في صحيفة النيويورك تايمس أنَّ المعسكر الموالي لاسرائيل تجاوز الحد المقبول
 wa+yaEotaqid+u + Al+kAtib+u + fiy + SaHiyfap+i + Al+niyuwyuwrok + tAyomos + >an~a
 + Al+muEasokar + Al+muwAliy+a + li+<isorA}iyl+a + tajAwaz+a + Al+Had~+a +
 Al+maqobuwl+a
 and+think + the+writer + in + newspaper + the+New York + Times + that + the+camp +
 the+partisan + to+Israel + exceeded + the+limit + the+admitted
*In the newspaper of The New York Times, the writer believes that the Israeli partisan camp
 exceeded the admitted limit*

Adverbial clauses headed by complementizers are also annotated as SBAR, but they must also have an adverbial dashtag reflecting the function they serve.

Subjects can also be SBAR (see section ???? on clausal subjects):


```

(S (VP yu+mokin+u::يُمْكِنُ::he/it+be_possible/make_possible_for+[ind.]
  (SBAR-SBJ >an::أَنْ::to
    (S (VP ta+tataHaw~al+a::تَتَحَوَّلُ::it/they/she+be_changed/
      be_transformed+[sub.]
      (NP-SBJ (NP هذه ::ha*ihi:: this)
        (NP الحرب Al+Harob+u:: the+war))
      (PP fiy::في::in
        (NP suroE+ap+K::سُرْعَة::speed/velocity/
          promptness+[fem.sg.]+
          [indef.gen.] )))
      (NP-ADV (NP Harob+AF::حَرْباً::war/warfare+[acc.indef.])
        (NP-ADV gayor+a::غَيْرَ::not/other+[def.acc.]
          (NP taqoliydiy~+ap+K::تَقْلِيدِيَّة::
            traditional/conventional+
            [fem.sg.]+[indef.gen.] [ADJ]
            )))))))
    يمكن أن تتحول هذه الحرب في سرعة حرباً غير تقليدية
yumokin+u + >an + tataHaw~al+a + h`*ihi + Al+Harob+u + fiy + suroEap+K + Harob+AF
+ gayor+a + taqliydoy~ap+K
be possible + to + be changed + this +war + in + promptness + war + not + traditional
This war can promptly change into a non-traditional one
eg to be checked with the tree

```

2.6.3 Gapping (VP Template Gapping)

Template gapping is done as in the Penn English Treebank, with the exception that all gapping indexing is shown with an = and is, like all indices in the Arabic Treebank, on the node label itself.

Gapping is the only construction in the Arabic Treebank that allows coordinated VPs.

The intended interpretation of the template and the gapping is that

- The first conjunct sets up the template. Everything in the first conjunct that is not marked with an = index is to be interpreted in the same way in all following conjuncts (the verb, for example).
- Successive conjuncts are interpreted by repeating the non-indexed portions of the first conjunct, and replacing all = indexed constituents with the co-indexed constituent in the new conjunct.

So, in the example below, the template is “fAza fiy hi.” When interpreting the second conjunct, the NP-SBJ=1 in the first conjunct is replaced in the template by the NP-SBJ=1 in the second conjunct; the NP-TMP=2 in the first conjunct is also replaced in the template by the NP-TMP=2 in the second conjunct. This gives a reconstructed second conjunct of “fAza fiy hi \$uwmaAxar EAm+ayo 2000 wa 2001.”

(S (VP (VP -fAz+a::فاز::win/be_victorious+he/it_[verb]
 (PP fiy-::في::in
 (NP -hi::هو::it/him))
 (NP-SBJ=1 (NP Al+biriyTAniy~+u::البريطاني::the+British+[def.nom.])
 (NP Ady::إدي::Ed
 <iyrfAyn::إيرفاين::Irvine))
 (NP-TMP=2 EAm+a::عام::year+[def.acc.]
 ,::,::nogloss
 (NP 1999::1999::nogloss)))
 wa-::و::and
 (VP (NP-SBJ=1 -\$uwmAaxar::شوماخر::Schumacher)
 (NP-TMP=2 EAm+ayo::عامي::year+two_[acc.]
 (NP 2000::2000::nogloss
 wa::و::and
 .2001::.2001::nogloss))))
 فاز فيه البريطاني ادي إيرفاين عام 1999 وشوماخر عامي 2000 و 2001
 fAz+a + fiy+hi + Al+biriyTAniy~+u + >adiy + <iyrfAyn + EAm+a + 1999 + wa+\$uwmAaxar
 + EAmay+o + 2000 + wa+2001
 won + in+it + the+British + Eddie + Irvine + year + 1999 + and+Schumacher + two years +
 2000 + and+2001
In which the Brit, Eddie Irvine, won in 1999 and Schumacher in 2000 and 2001

2.6.4 Separate sentences

Sentences that do not have any syntactic relation (subordination, coordination) with the previous sentence are annotated as separate sentences.

(S (NP-TPC-1 الحرب ::-Al+Harob+u::the+war/warfare+[def.nom.]
 الاهلية ::Al+>aholiy~+ap+u::the+civil/domestic/
 family+[fem.sg.]+[def.nom.]
 الفلسطينية ::Al+filasoTiyniy~+ap+u::the+Palestinian+
 [fem.sg.]+[def.nom.])
 (VP (PRT لم::lam::did_not)
 تقع::ta+qaE+o::it/they/she+fall_down/take_place/be_located+[jus.]
 (NP-SBJ-1 *T*))
nogloss)
 الحرب الأهلية الفلسطينية لم تقع
 Al+Harob+u + Al+>aholiy~+ap+u + Al+filasoTiyniy~+ap+u + lam + taqaE+o
 the+war + the+civil + the+Palestinian + did not + take place
The Palestinian civil war did not happen

(S (PRT لكن::l`kin~a::however)
 (NP-TPC-1 صورة::Suwr+ap+a::manner/way/form+[fem.sg.]+[def.acc.]
 (NP إسرائيل::<isorA}iyl+a::Israel+[indef.gen.]
 "::~":nogloss
 الديمقراطية::Al+diymuwqrATiy~+ap+a::the+
 democratic+[fem.sg.]+[def.acc.]
 "::~":nogloss))
 (VP اهتزت::{ihotaz~+at::tremble/quake/be_shaken+it/they/she_[verb]
 (NP-SBJ-1 *T*)
 (PP في::fiy::in
 (NP العمق::Al+Eumoq+i::the+depth/bottom+[def.gen.])))
 لكن صورة إسرائيل الديمقراطية اهتزت في العمق
 l`kin~a + Suwr+ap+a + <isorA}iyl+a + Al+diymuwqrATiy~+ap+a + <ihotaz~+at + fiy +
 Al+Eumoq+i
 however + image + Isreal + the+democratic + be shaken + in + the+depth
However, the image of a democrattic Israel suffered deeply

2.6.5 S-ADV versus SBAR zero relative clause

جاء الرجال من بينهم محمد
 jA'a AlrijAlu min baynihim Muham~ad
 came the men of between them Mohamed
The men came having Mohamed with them

جاء رجال من بينهم محمد
 jA'a rijAlN min baynihim Muham~ad
 came men of between them Mohamed
Men that have Mohamed among them came

In the first example, the NP “الرجال” (Al+rijAl, the men) is definite, so the clause following it cannot be a relative clause without an overt WHNP. This is considered as an S-ADV which is HAl (i.e., giving the state of the NP-SBJ or NP-OBJ of the main clause when the event took place) in Arabic grammar, and as such it is attached at the VP level.

(S (VP جاء
 (NP-SBJ الرجال)
 (S-ADV (PP-PRD من
 (NP بين
 (NP هم)))
 (NP-SBJ محمد)))

جاء الرجال من بينهم محمد

jA'a AlrijAlu min baynihim Muham~ad
 came the men of between them Mohamed
The men came having Mohamed with them

NB: please check transliteration, and add to the tree also

The absence of an overt WHNP is possible in relative clauses in MSA. The second clause above is a relative clause.

NB: This needs to have an expanded explanation (describing how to know that it is a relative clause and not S-ADV).

```
(S (VP جاء
  (NP-SBJ (NP رجال)
    (SBAR (WHNP-1 *0*)
      (S (PP-PRD من
        (NP بين
          (NP (NP هم)
            (NP-1 *T*)) ) ) )
        (NP-SBJ محمد) ) ) ) ) )
```

جاء رجال من بينهم محمد

jA'a rijAlN min baynihim Muham~ad
came men of between them Mohamed
Men that have Mohamed among them came

NB: please check transliteration, and add to the tree also

2.6.6 SBAR-ADV versus separate S's

جاء الرجال و من بينهم محمد
jA'a AlrijAlu wa min baynihim Muham~ad
came the men and of between them Mohamed
The men came while having Mohamed with them

جاء رجال و من بينهم محمد
jA'a rijAlN wa min baynihim Muham~ad
came men and of between them Mohamed
Men that have Mohamed among them came

In the first example, the NP “الرجال” (Al+rijAl, the men) is definite and the clause following it gives the state of the NP-SBJ of the main clause when the event took place. It is therefore annotated as an SBAR-ADV modifying the VP.

```
(S (VP جاء
  (NP-SBJ الرجال)
  (SBAR-ADV و
    (S (PP-PRD من
      (NP بين
        (NP (NP هم)
          (NP-1 *T*)) ) ) )
      (NP-SBJ محمد) ) ) ) )
```

جاء الرجال و من بينهم محمد

jA'a AlrijAlu wa min baynihim Muham~ad
came the men and of between them Mohamed
The men came while having Mohamed with them

NB: please check transliteration, and add to the tree also

In the second example, the NP “رجال” (rijAl+N, men) is indefinite, so the clause following it has to be a new sentence coordinated with the first one. It cannot be an SBAR-ADV (i.e., HAI clause) because it does not modify an indefinite noun.

2.7 Empty categories

The empty categories are essentially the same as in the Penn English Treebank. The most common being

- * Pro-drop subjects and passive traces
- ***T*** WH-traces, NP-TPC trace to subject
- ***ICH*** Rightward movement (for the most part, also *RNR*, etc.)
- ***0*** Null complementizer or zero WH- pronoun
- *?* Placeholder for ellipsed material

As in the Penn Treebank, we are not showing any pronominal coreference. Coreference will be indicated only for empty categories and exceptional cases such as VP gapping structures.

Empty categories are complete constituents in themselves – they can never have modifiers or complements. Nothing additional can be inside an empty category's node. (See also section ??? on the treatment of resumptive pronoun traces.)

A simple sentence with pro-drop:

```
(S (S (VP qAma قام ) )
    (NP-SBJ zayodN زَيْدُ )
    wa وَ
    (S (VP tawaj~aha تَوَجَّهَ )
        (NP-SBJ * )
        (PP <ilaY إِلَى
            (NP AlbAbi الباب ) ) ) ) ) )
```

قام زيد وتوجَّه إلى الباب

qAm+a + zayod+N + wa+tawaj~ah+a + <ilaY + Al+bAb+i
 stood up + Zaid + and+move towards + to + the+door
Zaid stood up and moved towards the door

A topicalized NP subject trace:

(S-HLN (NP-TPC-1 majolis+u::مَجْلِس::council/board+[def.nom.]
 (NP ">::":::nogloss
 <inoruwn::إِنْزُون::Enron
 ">::":::nogloss))
 (VP >aqAl+a::أَقَالَ::dismiss/discharge+he/it_[verb]
 (NP-SBJ-1 *T*)
 (NP-OBJ (NP nA}ib+ayoni::نَائِبَيْن::deputy/delegate/vice-+two_[acc.]
 (PP li-::ل-::to/for
 (NP -Āl+ra}iys+i::الرَّئِيس::the+president/head/
 chairman+[def.gen.])))))
 مجلس إنرون اقال نائبين للرئيس
 majolis+u + <inoruwn + >aqAl+a + nA}ib+ayoni + li+Al+ra}iys+i
 council + Enron + dismissed + deputies+two + for+the+chairman
Enron Council dismissed two of the chairman's deputies

2.7.1 Discontinuous Constituents/Rightward Movement

Rightward-moved constituents (usually complements or modifiers of NPs) are coindexed with an empty element *ICH* (Interpret Constituent Here) at the location where they originate.

(S (VP أبدى::>abodaY+-LRB-null-RRB-::express/demonstrate/show+he/it_[verb]
 (NP-SBJ المستشار::Al+musota\$Ar+u::the+counselor+[def.nom.]
 الألماني::Al+>alomAniy~+u::the+German+[def.nom.]
 (NP-OBJ (NP رغبة::ragob+at+a::desire/wish/will+[fem.sg.]+[def.acc.]
 (NP ه::hu::his/its))
 (PP-2 *ICH*))
 (NP-TMP الاثنين::Al+{ivonayoni::the+Monday
 الماضي::Al+mADiy::the+past/bygone)
 (PP-2 في::fiy::in
 (NP (NP إرسال::<irosAl+i::sending/deploying_[troops]+[def.gen.]
 (NP قوة::quw~+ap+K::force/violence+
 [fem.sg.]+[indef.gen.]
 عسكرية::Easokariy~+ap+K::military/army+
 [fem.sg.]+[indef.gen.])))
 (PP-DIR الى::<ilaY::to/towards
 (NP الشرق::Al+\$aroq+i::the+east/East+[def.gen.]
 الاوسط::Al+>awosaT+i::the+middle+
 [def.gen.])))
 أبدى المستشار الألماني رغبته الإثنين الماضي في إرسال قوة عسكرية الى الشرق الأوسط
 >abodaY + Al+musota\$Ar+u + Al+>alomAniy~+u + ragobat+a+hu + Al+<ivonayon+i +
 Al+mADiy + fiy + <irosAl+i + quw~ap+K + Easokariy~ap+K + <ilaY + Al+\$aroq+i +
 Al+>awosaT+i
 expressed + the+counselor + the+German + will+his + the+Monday + the+last + in +
 sending + force + military + to + the+east + the+middle
Last Monday, the German counselor expressed his will to send a military force to the Middle East

Some examples of constructions that use *ICH*:

- If an NP has an SBAR complement but it is also modified by a multi word ADJP

```

(S (VP إن
  (NP-SBJ هـ)
  (NP-PRD (NP اعتراف
    (SBAR-1 *ICH* ) )
    (ADJP واضح
      و
      (و واقع ) )
    (SBAR-1 أن
      (S (NP-SBJ إسرائيل )
        (PP-PRD في
          (NP إنحدار
            شديد ) ) ) ) ) ) ) ) )

```

إنه اعتراف واضح و واقع أنّ إسرائيل في انحدار شديد

Indeed it confession clear and real that Israel in decline strong

Indeed it is clearly and really admitted that Israel is in strong decline

NB: needs transliteration, and add to the tree also

- If a construction involving a judgment verb has constituents in the following order:
judgment verb + subject of embedded clause + subject of main clause + predicate of
embedded clause

```

(S (NP-TPC-1 (NP هذا )
  (NP الأمر ) )
  (VP تعتبر
    (S (NP-SBJ (NP هـ)
      (NP-1 *T*) )
      (ADJP-PRD-2 *ICH*) )
      (NP-SBJ المنظمة الحقوقية )
      (ADJP-2 ضروريا ) ) ) )

```

هذا الامر تعتبره المنظمة الحقوقية ضروريا

This matter considered it the organization the civil rights necessary

The civil rights consider this matter important.

NB: please check transliteration, and add to the tree also

Right Node Raising: Right node raised constituents are similarly coindexed with an empty element *RNR* (for **Right Node Raising**) in each of the positions where the constituent is interpreted. Right Node Raising usually occurs in coordination contexts, and the raised (overt) constituent is placed at the conjunction level.

```

(S (S (VP -katab+a-::كَتَبَ::write+it/he/they _[verb]
      (NP-OBJ -hu::هُ::it/him )
      (NP-SBJ-3 *RNR*)) )
  wa-::وَ::and
  (S (VP -naf~a*a-::نَفَّذَ::implement/carry_out/accomplish+he/it_[verb]
      (NP-OBJ -hu::هُ::it/him )
      (NP-SBJ-3 *RNR*)) )
  (NP-SBJ-3 Huk~Am+u::حُكَّامُ::rulers/governors+[def.nom.]
    (NP (Al+balad+ayoni::الْبَلَدَيْنِ::the+country+two_[gen.] ) )
    كَتَبَهُ وَنَفَّذَهُ حُكَّامُ الْبَلَدَيْنِ
  katab+a+hu + wa+naf~a*a+hu + Huk~Am+u + Al+balad+ayoni
  wrote+it + and+accomplish+it + governors + the+country+two
  It was written and accomplished by the governors of the two countries

```

Some examples of constructions using RNR are below. (Note that right node raising is a productive construction, so there cannot be an exhaustive list of possible contexts for RNR.)

VP arguments:

```

(S (S (VP شجب
      (NP-SBJ *)
      (NP-OBJ-1 *RNR*)) )
  و
  (S (VP اذان
      (NP-SBJ *)
      (NP-OBJ-1 *RNR*)) )
  (NP -1 (NP جميع
      (NP (التدخلات) )
      (PP في
        (NP (NP الشؤون
              الداخلية
              ل
              (NP (للعراق) ) ) ) )
        شجب و اذان جميع التدخلات في الشؤون الداخلية للعراق

```

Denounced he and condemned he all interventions in the issues the internal to Iraq
He denounced and condemned all the interventions in the internal affairs of Iraq

NB: needs the transliteration, and add to the tree also

NP complements, when the NP conjuncts are more than one token (for single token conjuncts, no RNR is needed because the complement would be a child of the coordinated NP):

(NP (NP وزير
(NP خارجية
(NP-1 *RNR*)))
و
(NP وزير
(NP داخلية
(NP-1 *RNR*)))
(NP-1 لبنان))

وزير خارجية ووزير داخلية لبنان

Minister foreign and minister interior Lebanon

The Lebanese Ministers of Foreign Affairs and the Interior

NB: needs the transliteration, and add to the tree also

(NP بين
(NP (NP واحد
و
عشرين
(NP-1 *RNR*))
و
(NP خمسة
و
عشرين
(NP-1 *RNR*))
(NP-1 رجلا)))

بين واحد وعشرين وخمسة وعشرين رجلا

Between one and twenty and five and twenty men

Between twenty-one and twenty-five men

NB: needs the transliteration, and add to the tree also

(NP بين
(NP عشرة
و
عشرين
(NP رجلا)))

بين عشرة وعشرين رجلا

Between ten and twenty man

Between ten and twenty men

NB: needs the transliteration, and add to the tree also

Modifiers of VP:

(S (S (VP استشهد
(NP-SBJ فلسطيني
(PP-2 *RNR*)))
و
(S (VP جرح
(NP-SBJ-1 آخر)
(NP-OBJ-1 *)
(PP-2 *RNR*)))
(PP-2 ب
(NP نيران
(NP الاحتلال))))

استشهد فلسطيني و جرح آخر بنيران الاحتلال

Died Palestinian and was injured another with fires the colonizers

One Palestinian died and another was injured with the fires of the colonizers

NB: needs the transliteration, and add to the tree also

Modifiers of NP if they need to be interpreted at two different levels in the two conjuncts (otherwise modifiers can just be adjoined to the conjunction level), and if the conjuncts are more than one token (if both conjuncts are single tokens, no RNR is needed because the modifier will be adjoined to the coordinated NP):

NB: Need an example here, if this is going to be included in the list of example constructions.

Occasionally something which is not exactly a constituent has been moved rightward. Usually this happens with second conjuncts, where both the conjunction and the second conjunct are moved (as in "I ate lunch on Tuesday and dinner" or "I drink milk in the morning and juice").

When this happens, the entire moved portion is given the node label NAC (for Not A Constituent) and then coindexed with an empty *ICH* adjoined to the first conjunct.

NAC: discontinuous coordination

```
(S (VP $aributu شربت
  (NP-SBJ * )
  (NP-OBJ (NP HaliybAF حليباً
    (NAC-1 *ICH* ) )
  (PP-TMP fiy في
    (NP AlSabAHi الصباح ) )
  (NAC-1 wa و
    (NP EaSiyrAF عصيراً ) ) ) ) )
```

شربت حليباً في الصباح وعصيراً

\$aribot+u + Haliyb+AF + fiy + Al+SabAH+i + wa+EaSiyr+AF

drink + milk + in + the+morning + and+juice

I drank milk and juice in the morning

A parallel example of normal, unmoved coordination:

```
(S (VP $aributu شربت
  (NP-SBJ * )
  (NP-OBJ HaliybAF حليباً
    wa و
    EaSiyrAF عصيراً )
  (PP-TMP fiy في
    (NP AlSabAHi الصباح ) ) ) ) )
```

شربت حليباً وعصيراً في الصباح

\$aribot+u + Haliyb+AF + wa+EaSiyr+AF + fiy + Al+SabAH+i

drink + milk + and+juice + in + the+morning

I drank milk and juice in the morning

2.7.2 Ellipsis

? (placeholder for elided material)

? acts as a placeholder for a missing predicate or piece thereof, in environments where predicate deletion occurs. Although the missing material represented by *?* is often identical to another constituent in the same sentence, the two are never coindexed.

```
(S (S (NP-ADV TabE+AF طبعا )
      (NP-TMP Al+Ay~Am الأيام
        Al+mADiy~+ap الماضية )
      (VP kAn+at كانت
        (NP-SBJ Al+{it~iSAl+At الاتصالات )
        (ADJP-PRD say~i}+ap سيئة
          jid~+AF جدا)))
```

wa و

```
(S (VP -ya+bduw يبدو
      (SBAR-SBJ >an~a- أن
        (S (NP-TPC-1 -hA ها )
          (VP (PRT mA- ما )
            -zAl+at زالت
            (NP-SBJ-1 *T*)
            (ADJP-PRD *?*) ) ) ) ) ) )
```

طبعا الأيام الماضية كانت الاتصالات سيئة جدًا ويبدو أنها مازالت

Naturally the days the passed were the communications bad very and seems that it not ceased
Naturally the past days the network was very bad and it seems tha it still is

NB: needs the transliteration, and add to the tree also

2.8 Clitics and Tokenization

2.8.1 Clitics

The prevalence of cliticization in Arabic sentences of determiners, prepositions, conjunctions, and pronouns led to a necessary difference in tokenization between the POS files and the treebanking files. Clitics that play a role in the syntactic structure are split off into separate tokens (e.g., object pronouns cliticized to verbs, subject pronouns cliticized to complementizers, cliticized prepositions, etc.). Clitics that do not affect the structure are not separated (e.g., determiners). Since the word boundaries necessary to separate the clitics are taken from the POS tags, and since it is not possible to show the syntactic structure unless the clitics are separated, clitics necessary for the trees are automatically separated based on the POS selection in order to create the segmentation necessary for treebanking. The edges of these clitic separations are shown with “-“ in the transliteration.

PP cliticized to an object NP, split apart so that the NP can be shown:

```
(PP by/with bi-
  (NP any+[def.gen.] ->ay~+i
    (NP activity+[indef.gen.] na$AT+K
      political+[indef.gen.] siyAsiy~+K)))
```

بأي نشاطٍ سياسي

bi+>aiy~+i + na\$AT+K + siyAsap+K
 in+any + activity + policy
There is policy in every activity

PP with a cliticized object pronoun, split apart so that the NP can be shown:

```
(PP من::-min-::from
  (NP هم::-hum::them_[masc.pl.])))
```

من هم

min + hum
 from + they
from them

PP with a clausal complement – the preposition is separated, so that the SBAR can be shown:

```
(S (VP يعرف::yu+Ear~if+u::he/it+acquaint+[ind.]
  (NP-SBJ *)
    (PP ب::bi-::with
      (SBAR أن::->an~a-::that
        (S (NP-SBJ هو::-hu::it/he)
          (NP-PRD (NP المعجب::Al+muEojab+u::the+admirer
            proud+[def.nom.]
            الأول::Al+>aw~al+u::the+first+[def.nom.]
            (PP ب::bi-::by/with
              (NP أعمال::->aEomAl+i::actions/activities/
                work+[def.gen.]
                (NP والد::wAlid+i-::father+[def.gen.]
                  (NP هو::-hi::its/his))))))))))
        يعرف بأنه المعجب الأول بأعمال والده
        yuEar~if+u + bi+>an~a+hu + Al+muEojab+u + Al+>aw~al+u + bi+>aEomAl+i +
        wAlid+i+hi
        acquaint + with+that+he + the+admirer + the+first + by+works + father+his
        He tells that he is the first admirer of his father's works
```

Subject pronoun cliticized to a complementizer, split so that the structure can be shown:

(SBAR ان::>an~a-:::that
 (S (NP-TPC-3 ه:::-hu:::it/he)
 (VP سيلتقي::sa+ya+lotaqiy+-LRB-null-RRB-:::will+he/it+meet/
 encounter+[ind.]
 (NP-SBJ-3 *T*)
 (NP-OBJ (NP رئيس::ra}iys+a:::president/head/chairman+[def.acc.]
 (NP الوزراء::Al+wuzarA'+i:::the+
 ministers+[def.gen.])))
 (NP رفيق::rafiyq::Rafiq/Rafeek
 الحريري::Al+Hariyriy~:::the+Hariri))))
 أنه سيلتقي رئيس الوزراء رفيق الحريري
 >an~a+hu + sa+yalotaqiy + ra}iys+a + Al+wuzarA'+i + rafiyy + Al+Hariyriy
 that+he + will+meet + chairman + ministers + Rafeek + the+Hariri
 That he will meet the prime minister Rafeek Hariri

(SBAR >an~a-:::أَنْ-:::that
 (S (NP-TPC-3 -hA:::-ها:::it/they/she)
 (VP (PRT لا:::لا:::no/not/non-)
 tu+mAniE+u:::تُمانع:::it/they/she+oppose+[ind.]
 (NP-SBJ-3 *T*)
 (NP-OBJ (NP Eaod+a:::عقد:::holding/concluding/convening+[def.acc.]
 (NP Al+qim~+ap+i:::القمة:::the+summit+
 [fem.sg.]+[def.gen.]))
 (PP-LOC في:::في:::in
 (NP Al+qAhir+ap+i:::القاهرة:::the+Cairo+
 [fem.sg.]+[def.gen.]))))))
 أنها لا تمنع عقد القمة في القاهرة
 >an~a+hA + lA + tumAniE+u + Eaod+a + Al+qim~ap+i fiy + Al+qAHirap+i
 that + not + oppose + holding + the+summit + in + Cairo
 That it would not mind holding the summit in Cairo

Conjunction cliticized to a verb, split so that the verbal structure can be shown:

(S و:::wa-:::and
 (VP (PRT س will)
 تستكمل : ta+sotakomil+u::: +it/they/she+complete/fulfill+[ind.]FUT_PRT
 (NP-SBJ-1 (NP هذه::h`*ihi:::this/these)
 (NP المحادثات::Al+muHADav+At+u:::the+discussions/talks/
 negotiations+[fem.pl.]+[def.nom.])))
 (NP-OBJ-1 *)
 (PP-TMP في:::fiy:::in
 (NP (NP النصف::Al+niSof+i:::the+half/middle/semi-+[def.gen.]
 الثاني::Al+vAniy:::the+second/next)
 (PP من:::min:::from
 (NP نيسان::niysAn+a:::April+[indef.gen.]
 المقبل::Al+muqobil+i:::the+next/coming/
 approaching+[def.gen.]))))))
 وستكمل هذه المحادثات في النصف الثاني من نيسان المقبل
 wa+satasotakomil+u + h`*ihi + Al+muHADavAt+u + fiy + Al+niSof+i + Al+vAniy + min +
 niysAn + Al+muqobil+i
 and+complete + this + the+discussions + in + the+half + the+second + from + April +
 the+next
 These discussions will complete in the last two weeks of next April

Pronoun object cliticized to verb, split so that the object structure can be shown:

```
(S (NP-TPC-1 (NP h`*ihi::هـَـذِهِ::this/these )
  (NP Al+>aroqAm+u::الأَرْقَامُ::the+numbers/numerals+[def.nom.] ))
  (VP >aEolan+a-::أَعْلَنَ-::announce/declare+he/it_[verb]
    (NP-OBJ (NP -hA::هـَا::it/them/her )
      (NP-1 *T*))
    (NP-TMP (>amosi::أَمْسَ::yesterday ))
    (NP-SBJ ra}iys+u::رَئِيسُ::president/head/chairman+[def.nom.]
      (NP majolis+i::مَجْلِسُ::council/board+[def.gen.]
        (NP <idAr+ap+i::إِدَارَةُ::administration/
          management/bureau+[fem.sg.]+[def.gen.]
            (NP Al+baronAmaj+i::الْبَرْنَامَجُ::the+
              program+[def.gen.] ))))))
      هذه الأرقام أعلنها أمس رئيس مجلس إدارة البرامج
h`*ihi + Al+>aroqAm+u + >aEolan+a+hA + ra}iys+u + majolis+i + <idArap+i +
  Al+barAmij+i
these + the+numbers + announce+it + yesterday + chairman + council + administration +
  the+program
yesterday these numbers were announced by the council chairman of the program's
  administration
```

Conjunction cliticized to a noun, split so that the noun phrase structure can be shown:

```
(S (VP -ya+doEam+u::يَدْعُمُ-::he/it+support/strengthen+[ind.]
  (NP-SBJ *)
  (NP-OBJ (NP SAdir+At+i::صَادِرَاتُ-::exports+[fem.pl.]+[def.acc.]
    (NP -hi::هِيَ::its/his ))
    wa::وَ::and
    (NP -AiqotiSAd+a-::اِقْتِمَادُ-::economy/saving+[def.acc.]
      (NP -hu::هُوَ::its/his )))))
  يدعم صادراته واقتصاده
yadoEam+u + SAdirAt+i+hi + wa+<iqotiSAd+a+hu
support + exports+its + and+economy+its
Supports its exports and economics
```

Cliticized determiners are left attached to the noun/adjective:

```
(NP الدَّم Al+dam+i + the+blood+[def.gen.] )
```

In the above example, technically “the/Al-” is a clitic as well, but we don’t split them in the treebank because we never annotate the determiner separately from the noun in the Arabic Treebank.

Possessive pronoun clitics are split from the noun, so that the pronoun can be annotated as a complement of the noun:

```
(PP by/with bi-
  (NP hand+two_[gen.] -yad+ayo-
    (NP his/its -hi)))
```

بيديه

bi+yad+ay+o+hi
 with+hand+two+his
With his hands

2.8.2 Tokenization

Clitics that play a role in the Arabic Treebank trees are split off into separate tokens. For the sake of the annotators, clitics that do not affect the structure are not separated (e.g., determiners). Since the word boundaries necessary to separate the clitics are taken from the POS tags, and since it is not possible to show the syntactic structure unless the clitics are separated, clitics necessary for the trees are automatically separated based on the POS selection in order to create the segmentation necessary for treebanking. The edges of these clitic separations are shown with “-“ in the transliteration.

The following POS tags indicate that a clitic will be separated. This separation is accomplished via automatic means – with rare exceptions, Treebank annotators do not split clitics manually. Clitics are never separated from case marking information.

- **Conjunctions:** CONJ
- **Prepositions:** PREP
- **Subordinating conjunctions, complementizers, and relative pronouns** (including any inflectional variations): SUB_CONJ, REL_PRON, REL_ADV, INTERROG_PRON, INTERROG_ADV, EXCLAM_PRON, INTERROG_PRON+CASE_DEF_GEN, INTERROG_PRON+CASE_DEF_NOM, INTERROG_PRON+NSUFF_FEM_SG+CASE_DEF_GEN, REL_PRON+CASE_DEF_ACC, REL_PRON+CASE_DEF_GEN, REL_PRON+CASE_DEF_NOM, REL_PRON+CASE_INDEF_ACC, REL_PRON+NSUFF_FEM_SG+CASE_DEF_GEN, REL_PRON+NSUFF_MASC_SG_ACC_INDEF
- **Particles:** CONNEC_PART, EMPHATIC_PART, FOCUS_PART, FUT_PART, INTERROG_PART, JUS_PART, NEG_PART, PART, RC_PART, RESTRIC_PART, VERB_PART, VOC_PART
- **Pronouns:** PRON, POSS_PRON, CVSUFF_DO:1P, CVSUFF_DO:1S, CVSUFF_DO:3FS, CVSUFF_DO:3MP, CVSUFF_DO:3MS, IVSUFF_DO:1P, IVSUFF_DO:1S, IVSUFF_DO:2FS, IVSUFF_DO:2MP, IVSUFF_DO:2MS, IVSUFF_DO:3D, IVSUFF_DO:3FS, IVSUFF_DO:3MP, IVSUFF_DO:3MS, PRON_1P, PRON_1S, PRON_2D, PRON_2FP, PRON_2FS, PRON_2MP, PRON_2MS, PRON_3D, PRON_3FP, PRON_3FS, PRON_3MP, PRON_3MS, PVSUFF_DO:1P, PVSUFF_DO:1S, PVSUFF_DO:2FS, PVSUFF_DO:2MP, PVSUFF_DO:2MS, PVSUFF_DO:3D, PVSUFF_DO:3FS, PVSUFF_DO:3MP, PVSUFF_DO:3MS, POSS_PRON_1P, POSS_PRON_1S, POSS_PRON_2FP, POSS_PRON_2FS,

POSS_PRON_2MP, POSS_PRON_2MS, POSS_PRON_3D, POSS_PRON_3FP,
POSS_PRON_3FS, POSS_PRON_3MP, POSS_PRON_3MS

The above tags are based on the ATB3-v3.1 release Arabic Treebank Part 3 v 3.1 (LDC catalog number: LDC2008E22). Specific additional inflectional variants may arise in other corpora.

2.9. Attachment decisions:

- Make the attachment decision for every single constituent. Every PP. Every ADJP. Every SBAR. Every NP. Etc.
- Get the full interpretation of the sentence *in context* first, then
- If you know from the context where the constituent should be attached (what it belongs with), attach it in the right place
- If it is truly ambiguous between two possible attachments, even in the context of the article, attach it with the higher one

TEST #1: Adverbial modifier with NP or with VP?

- Move the modifier to before the NP. If the meaning of the sentence stays the same, the modifier should be attached at the VP level. If the meaning of the sentences changes when the modifier is moved, then the modifier should be attached to the NP.

```
(S (VP سمعت
  (NP-SBJ *)
  (NP-OBJ (NPالكثير )
    (PP من
      (NP الرجال ) ) ) ) )
```

سمعت الكثير من الرجال

Heard the many of men

I heard many men

NB: needs the transliteration, and add to the tree also

```
(S (VP سمعت
  (NP-SBJ *)
  (NP-OBJ الكثير )
  (PP من
    (NP الرجال ) ) ) )
```

سمعت الكثير من الرجال

I heard a lot from men

NB: needs the transliteration and the gloss, and add to the tree also

If we can say with the same meaning, “Occur in the year 1995 the election” then the temporal modifier is related to the VP.

If moving the temporal modifier changes the meaning, then the temporal modifier should be attached to the NP.

Example: “I saw a man with binoculars”

2.9 Punctuation

Punctuation marks of all kinds ***except for quotation marks and parenthetical marks*** should be put as high as they can be in the tree.

Sentence-final punctuation should be a child of S, or of the highest level syntactic node:

S with period "."

```
(S (NP-SBJ وضع::waDoE+u::situation/status/condition+[def.nom.]
    (NP مؤسسة::mu&as~as+ap+i::institution/
        organization+[fem.sg.]+[def.gen.]
        (NP الكهرباء::Al+kahorabA'+i::the+electricity+[def.gen.])))
    (ADJP-PRD مفلس::mufolis+N::bankrupt/insolvent+[indef.nom.]
    .:::nogloss)
```

وضع مؤسسة الكهرباء مفلس.

waDoE+u + mu&as~as+ap+i + Al+kahorabA'+i + mufolis+N
situation + institution + the+electricity + insolvent
the situation of the electricity institution is insolvent.

S with question mark "?"

```
(SQ (PRT hal::هل::does/do?/did?/is/are? )
    (VP sa+ya+fuk~+u::سيفك::will+he/it+separate/dismantle/
        detach/disengage+[ind.]
        (NP-SBJ $Aruwn::شارون::Sharon )
        (NP-OBJ >asor+a-::أسر::capture/captivity+[def.acc.]
        (NP -hu::هو::its/his )))
    ?:::nogloss )
```

هل سيفك شارون أسره

hal + sa+yafuk~+u + \$Aruwn + >asor+a+hu
does/do?/is? + will+detach + Sharon + captivity
Will Sharon release him?

The final punctuation for any stand-alone phrase should be the child of the highest level syntactic node:

FRAG with final punctuation:

```
(FRAG (NP bAriys::باريس::Paris )
    -:::nogloss
    (PP min::من::from
        (NP biyAr::بيار::Pierre
            EaTAAll~`h::عطا الله::Attallah ))
    :::::nogloss )
```

باريس من بيار عطا الله

bAriys + min + biyAr + EaTAAll~`h
Paris + from + Pierre + Attallah
Paris from Pierre Attallah

Top level NP with final punctuation:

```
(NP-HLN (NP 13::13::nogloss
  (NP niysAn+u::نيسان::April+[def.nom.] ))
  wa::و::and
  ">::"::nogloss
  (NP mu$otaq~+At+u-::مُشتَقَاتْ-::derivatives+[fem.pl.]+[def.nom.]
    (NP -hA::ها::its/their/her ))
  ">::"::nogloss )
```

نيسان ومشتقاتها

niysAn+u + wa+mu\$otaq~+At+u+hA
 April + and+derivatives+its
Nissan and its derivatives

Commas should be put high in the tree, not inside a lower constituent:

```
(S (VP (PRT -qad::قد::[has/have] )
  >avobat+a::أثبت::ascertain/establish+he/it_[verb]
  (NP-SBJ *)
  (NP-ADV sAbiq+AF::سابقاً::formerly/earlier+[acc.indef.] )
  (NP-OBJ jadAr+at+a-::جدارته::worthiness/aptitude/merit+
    [fem.sg.]+[def.acc.]
    (NP -hu::هو::his/its ))
  , :: , ::nogloss
  (NAC wa-::و-::and
    (PP (NP-ADV -xuSuwS+AF::خُصُوصاً::especially/in_particular+[acc.indef.])
      bi-::بِ-::by/with
      (NP -<izA'+i::إزاء::towards/facing/regarding/
        vis--vis+[def.gen.]
        (NP Al+muEAriD+iyna::المعارضين::the+opponent/adversary/
          opposition+[masc.pl.gen.] )))))
  قد أثبت سابقاً جدارته وخصوصاً بإزاء المعارضين
```

qad + >avobat+a + sAbiq+AF + jadArat+a+hu + wa+xuSuwS+AF + bi+<izA'+i +
 Al+muEAriDiyna
 has + establish + earlier + wptitude+his + and+especially + with+facing + the+opponents
Earlier he has established his aptitude especially in facing the opponents

The only exception to the “punctuation high” rule is for quotation marks (“”) or parentheses (()) or other balanced punctuation marking parenthetical material (dashes –, for example).

Quotation marks should be placed inside the highest node of the quotation. Punctuation marking parentheticals should be placed inside the PRN. *All other punctuation marks are placed high in the tree.*

Quotation marks should be put inside the quoted constituent (the closing quotation mark should be at the same level as the opening quotation mark, if possible – otherwise, it should be as high as possible):

```
(PP li-::ل-::to/for
  (NP (NP -Al+qaDA'+i::-القضاء::the+extermination/annihilation+[def.gen.])
    (PP EalaY::عَلَى::on/above
      (NP "::::nogloss
        >abuw::أَبُو::Abu
        say~Af::سَيَّاف::Sayyaf
        "::::nogloss ))))
```

للقضاء على أبو سيّاف

li+Al+qaDA'+i + EalaY + a>buw + say~Af
 for+the+extermination + on + Abu + Sayyaf
To exterminate Abu Sayyaf

```
(S wa
  (VP qAla
    (NP-SBJ whoever)
    (S "
      blahblah blab blah
      ")))
```

Quotation marks should be placed as high as they can be inside the quotation.

Parentheses and other matching brackets should be put inside the PRN (the closing parenthesis should be at the same level as the opening parenthesis).

```
(S blah blah blah
  (PRN -LRB-
    blah blah blah
    -RRB-)
  blah blah blah
  .)
```

Note that the parentheses marks in the .tree files are represented as –RRB- (Right Round Bracket) and –LRB- (Left Round Bracket) in the .tree files so that they do not interfere with the brackets delimiting the tree nodes.

See also section ????? on the use of PRN for parentheticals.

3 Noun Phrase and Adjective Phrase Structure

In traditional Arabic grammar, only three parts of speech are recognized: noun, verb and particle. The distinction between nouns that are nominal vs. adjectival is made at a slightly different level in the traditional grammar. However, for the purpose of Treebank annotation, it is desirable to have this distinction at the level of the core part of speech. For this reason, we include Noun Phrase and Adjective Phrase Structure together in this section.

The traditional Arabic grammar treats nouns and adjectives as part of a single larger category that has the following characteristics in common:

- having the three possible case endings (nominative, accusative, genitive)
- being definite (with the determiner *Al-*) or indefinite (with the tanwin *sign of infiniteness/making-an-n* case ending of N, F, or K)
- being the complement of a preposition in some contexts

This larger category can be divided into “invariable” and “variable” (Al>asomA' Almu3rabap *inflected nouns*) nouns. “Invariable” nouns do not overtly show the different case endings. “Variable” nouns (Al>asomA' Almu3rabap) do overtly show the three case endings.

The subcategory of “variable” nouns (Al>asomA' Almu3rabap) has two subcategories itself:

- non-adjectival nouns (Aism gayr Sifap *noun-not-adjective*)
- adjectival nouns (Aism Sifap *noun-adjective*)

This is where nouns and adjectives are distinguished in the traditional Arabic grammar.

The non-adjectival nouns (Aism gayr Sifap) include non-event nouns and event nouns (masdar). All nouns in this category have the core part of speech “NOUN” in the Treebank, and all should head NPs.

The adjectival nouns (Aism Sifap) include adjectives (Sifap mu\$ab~ahap, *stative adjective/ adjectives-made-like-participles* and Sigap mubAlagap, *intensifier form/form-of-exaggeration*), active and passive participles, elatives and ordinal numerals – along with participles that are lexicalized as nouns. Most of the words in this category will have the core part of speech “ADJ” in the Treebank, although the lexicalized ones will have the core part of speech “NOUN.” The guidelines for distinguishing these core parts of speech are in the Part of Speech Guidelines (section ???). For the structure of phrases headed by adjectives, see section ????? below.

The argument/adjunct distinction is shown structurally inside NPs for NP and clausal complements. All PPs, ADJPs and other modifiers are shown as adjuncts. Argument/complement constituents are children of NP, sister to the head noun: (NP head (NP argument)).

3.1 Heads

In general, noun phrases, NPs, will be headed by nouns (with *NOUN* in their part of speech tag) and adjective phrases, ADJPs, will be headed by adjectives (with *ADJ* in their part of speech tag). However, constructions that differ from this general rule in one way or another, and these deserve particular mention.

3.1.1 Numbers as heads

See the section on Numbers, Quantifiers and Comparatives below, section number ???.

nb: QP (Quantity Phrase) is no longer used in the Arabic Treebank for the constructions in which it is used in the English Treebank (or in early releases of the Arabic Treebank). It is in principle being held in reserve, in case future data arises with multi-word numbers that require its use.

See section ?????????? below for a more detailed discussion of the treatment of numbers.

```
(PP-LOC EalaY::على::on/above
  (NP (NP masAf+ap+i::مسافة::distance/interval+[fem.sg.]+[def.gen.]
    (NP >arobaE+ap+i::أربعة::four+[fem.sg.]+[def.gen.]
      (NP kiyluwmitr+At+K::كيلومتر::kilometer[s].) )))
  (NP-LOC dAxila::داخل::inside_of
    (NP Al+>arADiy::الأراضي::the+territories/land
      Al+<isorA}iyliy~+ap+i::الإسرائيلية::the+Israeli.]
      )))
  على مسافة أربعة كيلومترات داخل الأراضي الإسرائيلية
EalaY + masAfap+i + >arobaEap+i + kilowmitrAt+K + dAxil+a + Al+>arADiy +
Al+<isorA}iy~ap+i
On + distance + four + kilometers + inside + the+territories + the+Israeli
Four kilometers inside the Israeli territories
```

```
(S (VP kAna كَانَ
  (NP-SBJ * )
  (ADJP-PRD >aw~ala أَوَّل
    (NP AlmugAdiriyna الْمُغَادِرِينَ ) ) ) )
  كان أول المغادرين
kAn+a + >aw~al+a + Al+mugAdiriyna
was + fist + the+leavers
He was the first one who left
```

3.1.2 Quantifiers as heads:

See the section on Numbers, Quantifiers and Comparatives below, section number ????

Quantifiers are annotated as the head of an NP that must have a complement in idafa.

(S (VP (PRT lA::لا::no/not/non-)
 ta+Eotamid+u::تَعْتَمِدُ::it/they/she+depend/rely_[on]+[ind.]
 (NP-SBJ *)
 (PP-CLR EalaY::عَلَى::on/above
 (NP Tariyq+ap+i::طَرِيقَةً::method/procedure+[fem.sg.]+[def.gen.]
 (NP (NP naZar+i::نَظَرٍ::view/look/seeing/
 eye-sight+[def.gen.]
 (NP (NP kul~+K::كُلٌّ::every/all/
 each_one+[indef.gen.])
 (PP min::مِنْ::from
 (NP Al+dawol+atayoni::الدَوْلَتَيْنِ::
 the+state/country+two_[gen.])))
 (PP <ilaY::إِلَى::to/towards
 (NP bAkisotAn+a::بَاقِشْتَانِ::
 Pakistan+[indef.gen.]))))))))
 لا تعتمد على طريقة نظر كل من الدولتين إلى باكستان

lA+taEotamid+u + EalaY + Tariyqap+i + naZar+i + kul~+K + min + Al+dawolat+ayoni +
 <ilaY + bAkisotAn+a
 not + depend + on + method + view + each + from + the+countries+two + towards +
 Pakistan

Does not depend on the view of the two countries to Pakistan

(S (VP <in~a::إِنْ:: indeed
 (NP-SBJ-1 Al+jalos+ap+a::الْجُلْسَةُ::the+session/meeting+
 [fem.sg.]+[def.acc.])
 (S (VP rufiE+at::رُفِعَتْ::be_lifted/be_raised/be_increased+
 it/they/she_[verb]
 (NP-SBJ-1 *)
 (NP-OBJ-1 *)
 (NP-TMP baEoda::بَعْدَ::after
 (NP biDoE+i::بَعْضَ::some/several+[def.gen.]
 (NP daqA}iq+a::دَقَائِقَ::minutes+[indef.gen.]))))))))
 إِنَّ الْجُلْسَةَ رُفِعَتْ بَعْدَ بَعْضِ دَقَائِقَ

<in~a + Al+jalosap+a + rufiEat + baEoda + biDoE+i + daqA}iq+a
 indeed + the+session + was ended + after + few + minutes

The session was ended after few minutes

3.1.3 Adjectives as heads

See the section on False Idafa below, section ?????

See the section on Numbers, Quantifiers and Comparatives below, section number ????.

Adjectives can also be the head of an ADJP with complement PPs or NPs.

ADJP with a complement PP:

(S (PRT >am~A::أَمَّا::as_for/concerning)
 (NP-SBJ Al+bayAn+u::الْبَيَانُ::the+communiqu/statement/declaration+[def.nom.]
 Al+vAniy::الثَّانِي::the+second/next)
 (PRT fa-::فَـ::and/so)
 (ADJP-PRD -muwaj~ah+N::مُوجَّهَ::directed/aimed+[indef.nom.]
 (PP <ilaY::إِلَى::to/towards
 (NP Al+Tul~Ab+i::الطُّلَابُ::the+students+[def.gen.]))))
 أَمَّا الْبَيَانُ الثَّانِي فَمُوجَّهٌ إِلَى الطُّلَابِ
 >am~A + Al+bayAn+u + Al+vAniy + fa+muwaj~ah+N + <ilaY +Al+Tul~Ab+i
 as for + the+communiqu + the+second + so aimed + to + the+students
The second communiqué was aimed to the students

ADJP with a complement NP (false idafa):

(S (ADJP-PRD >aboraz+u-::أَبْرَزَ-::more/most_prominent+[def.nom]
 (NP -hA::هَـ::its/their/her))
 (NP-SBJ >aroqAm+u::أَرْقَامُ::numbers/numerals+[def.nom.]
 (NP Al+<inotAj+i::الْإِنْتَاجُ::the+production/output+[def.gen.]
 Al+SinAEiy~+i::الصَّنَاعِيَّ::the+industrial/
 artificial+[def.gen.])))
 أَبْرَزَهَا أَرْقَامُ الْإِنْتَاجِ الصَّنَاعِيِّ
 >aboraz+u+hA + >aroqAm+u + Al+<inotAj+i + Al+S~inAEiy~+i
 most prominent+its + numbers + the+production + the+industrial
Its most prominent are the numbers of the industrial production

3.1.3.1 NP's headed by adjectives

3.1.3.1.1 Single-word NP's headed by adjectives

NPs that are headed by single word adjectives are treebanked as NP. The adjectival nature of the head is represented in the part-of-speech tag.

(S (VP jA'a جَاءَ
 (NP-SBJ AlTawiyLu الطَّوِيلُ)))

جاء الطويل

jA'+a + Al+TawiyL+u
 came + the+tall
The tall came

Single word example with POS:

(S (VP jA'a جَاءَ)
 (NP-SBJ (ADJ AlTawiyLu الطَّوِيلُ))))

جاء الطويل

jA'+a + Al+TawiyL+u
 came + the+tall
The tall came

(S (VP fataH+a::فَتَحَ::open/conquer+he/it_[verb]
 (NP-SBJ *)
 (NP-OBJ Harob+AF::حَرْبًا::war/warfare+[acc.indef.]
 (NP gayor+a::غَيْرَ::not/other+[def.acc.]
 (NP ma\$oruwE+ap+K::مَشْرُوعَةٍ::lawful/legal+[fem.sg.] +
 [indef.gen.]))))
 فَتَحَ حَرْبًا غَيْرَ مَشْرُوعَةٍ
 fataH+a + Harob+AF + gayor+a + ma\$oruwEap+K
 Open + war + not + legal
He waged an illegal war

3.1.3.1.2 Multi-word NP's headed by adjectives

NPs that are headed by adjectives that are multiword (i.e., modified in any way, in a false idafa construction, etc.) are annotated as NPs with a single ADJP child.

(PP fiy::في::in
 (NP (ADJP muxotalif+i::مُخْتَلِفٍ::different/various+[def.gen.]
 (NP Al+majAl+At+i::الْمَجَالَاتِ::the+areas/fields/arenas/contexts))))
 فِي مَخْتَلَفِ الْمَجَالَاتِ
 Fiy + muxotalaf+i + Al+majAlAt+i
 In + various + the+areas
In various areas

(S (VP >awoqaE+at::أَوْقَعَتْ::inflict/bring_about+it/they/she_[verb]
 (NP-SBJ Al+Harob+i::الْحَرْبِ::the+war/warfare+[def.gen.]
 Al+>aholiy~+ap+i::الْأَهْلِيَّةِ::the+civil/domestic/
 family+[fem.sg.] + [def.gen.])
 (NP-OBJ (ADJP >akovar+a::أَكْثَرَ::more/most+[def.acc.]
 (PP min::مِنْ::from
 (NP miloyuwn+ayo::مِلْيُونِي::million+two_[gen.]
 (NP qatiyl+K::قَتِيلٍ::casualty/dead person/
 killed_person))))))
 أَوْقَعَتْ الْحَرْبَ الْأَهْلِيَّةَ أَكْثَرَ مِنْ مِلْيُونِي قَتِيلٍ
 >awqaEat + Al+Harob+i + Al+>aholiy~ap+i + >akovar+a + min + maloyuwn+ayo +
 qatiyl+K
 inflict + the+war + the+civil + more + than + million+two + killed person
The civil war caused the killing of more than two million people

(S (VP jA'a جَاءَ
 (NP-SBJ (ADJP >akovaru أَكْثَرُ
 (NP AlrijAli الرجال)
 (NP-ADV SaborAF صَبْرًا)))))
 جَاءَ أَكْثَرُ الرِّجَالِ صَبْرًا
 jA'+a + >akovar+u + Al+rijAl+i + Sabor+AF
 came + most + the+men + patience
The most patient man came

(S (VP jA'a جاء
 (NP-SBJ (ADJP Al>akovarو الأكثر
 (NP-ADV Saborأف صبراً ()))))

جاء الأكثر صبراً

jA'a + Al>akovar+u + Sabor+AF

came + the+most + patience

The most patient came

(S (VP qAbalotu قابلت
 (NP-SBJ *)
 (NP-OBJ (ADJP AlmaHoZuwZa المَحْظُوط
 jid~AF جداً ())))

قابلت المحظوظ جداً

qAbalot+u + Al+maHoZuwZ+a + jid~+AF

met + the+lucky + very

I met the very lucky person

3.1.3.1.3 Modification of NP's headed by adjectives

Whether the NP headed by the adjective is a single word or multi word, when that NP is further modified by an additional PP, the following tests will help to decide if the PP should be adjoined to the NP or a child of ADJP.

Difference between the attachment of the PP either NP or ADJP

الكثير من الرجال
 The many of the men
 Many men

الأكرم من الجميع
 The most generous of the all
 The most generous of all

2 consecutive tests:

1. Insert the deleted noun

العدد الكثير من الرجال
 The number the many of the
 men
 The big number of men

2. remove the adjective and see if it is
 grammatical

عدد من الرجال

2 consecutive tests:

1. Insert the deleted noun

الرجل الاكرم من الجميع
 The man the most generous
 of the all
 The most generous man of
 all

2. remove the adjective and see if it is
 grammatical

* الرجل من الجميع

The man of the all

Note: Use with extra care when using this test especially for the inserted nouns. They should be of a matching semantic class.

3.2 Complements

This section will be divided into a description of our treatment of complements of nouns (primarily the construct state, or “true” idafa: <iDAfap Haqiyqiy~ap *annexation-true*; and also tamiyz) and a description of our treatment of complements of adjectives (also the construct state, or “false” idafa: <iDAfap gayr Haqiyqiy~ap *annexation-not-true*).

The common traits of all construct state/idafa constructions is that:

- the complement must be genitive
- the head must be without tanwiyn.

Complements of numbers from 11 to 99 are accusative and are called in Arabic grammar "tamoyyz."

3.2.1 Complements of nouns

Complements/arguments are :

- genitive, possessive, obligatory, or (for deverbal head nouns) clausal constituents that would be arguments of the verb that the noun derived from.
- Accusative when complement of cardinals from 11-99

3.2.1.1 Genitive complement :the construct state (“true” idafa)

The **construct state headed by a noun** consists of two parts :

1. muDAf *annexed* or the Determined Noun

2. muDAf <ilayh *annexed to-it* or the Determining Noun (includes nouns with possessive pronouns)

```
(NP noun1/muDAf annexed
  (NP noun2/muDAf <ilayh annexed to-it))
```

The first part, which is the syntactic head, is often also the semantic head. However, in some constructions, the noun that is the syntactic head is not the semantic head of the construction. So we have :

a. muDAf *annexed* is

+ semantic head

+ syntactic head

+ noun indefinite and without tanween

```
(NP noun1/muDAf annexed
  (NP noun2/muDAf <ilayh annexed to-it))
```

b. muDAf annexed is

- semantic head
- + syntactic head
- + noun

(NP noun1/muDAf *annexed*
 (NP noun2/muDAf <ilayh *annexed to-it*))

The following categories are in this group of nouns that function syntactically as the head of the idafa/construct state construction, although they are not the semantic heads:

- Numbers (cardinal pronominal)
- Quantifiers (see POS guidelines for a list of NOUN_QUANT)
- Nouns of similitude (mivl *sameness*, \$iboh *similarity*, \$abiyh *resemblance*, naZiyr *counterpart*, maviyl *equal*)
- Nouns of exception (gayr *otherness*, siwaY *otherness*?????/)

(NP mivola مِثْلُ
 (NP SanamK صَنْمِ))

mivol+a + Sanam+K
 like + idle
Like an idle

مِثْلَ صَنْمٍ

(NP gayora غَيْرُ
 (NP Sabiy~K صَبِيٍّ))

gayor+a + Sabiy+K
 another + boy
Another boy

غَيْرَ صَبِيٍّ

(NP siwaY سِوَى
 (NP Aimora>apK اِمْرَاةٍ))

siwaY + <imora>ap+K
 except + woman
Except a woman

سِوَى اِمْرَاةٍ

3.2.1.1.1 Noun not obligatory in construct state

The NP which is the syntactic head, is also the semantic head. And this is the case especially of common nouns like كتاب kitAb (book) in the following example:

(NP kitaAbu book
(NP naHowK grammar))

كتاب نحو

kitAb+u + naHow+K
book + grammar
A grammar book

The theta role of the genitive complement in Arabic, as in German and probably in other languages is not necessarily that of the possessor (Kremers 2003). The construct state can be based on other semantic relation other than possession.

(NP xAtamu خاتمُ
(NP *ahabK ذهب))

خاتم ذهب

xAtam+u + *ahab+K
ring + gold
A gold ring

(NP madiynapu مَدِينَةُ
(NP dima\$ق دمشق))

مدينة دمشق

madiynap+u + dima\$oq
city + Damascus
The city of Damascus

The noun head of a true <iDAfap can be a proper noun with a metaphorical meaning of possession between the proper noun and its complement.

(NP jamiylu جميل
(NP bu*aynata بئينة))

جَمِيلُ بُئِينَةٍ

jamiyl+u + bu*aynat+a
Jamil + Butheina
Butheina's Jamil

The proper noun can also occur as a complement of a regular noun.

(NP Sadiyqu صديقُ
(NP zaydK زَيْد))

صَدِيقُ زَيْدٍ

Sadiyq+u + zayd+K
Friend + Zayd
Zayd's friend

NB: The possessive pronoun bound to a noun should always be treated as the complement of the noun.

(NP فرسه farasu
(NP هـ hu))

فَرَسُهُ

faras+u+hu
horse+his
His horse

3.2.1.1.2 Nouns with obligatory complements

In some constructions, the noun that is the syntactic head is not the semantic head of the construction. this is the case of the following category of nouns. These are referred to as >asmA' lAzimap Al<iDAfap *noun obligatory of annexation*.

3.2.1.1.2.1 Numbers (1-10 and hundreds, thousands, millions ...)

When these numbers are cardinal and pre-nominal, they are treated as the head of an idafa/complement construction.

(NP >arobaEapu أَرْبَعَةُ
(NP kutubK كُتُبُ))

أَرْبَعَةُ كُتُبٍ

>arobaEap+u + kutub+K
four + books
Four books

(NP >alofu أَلْفُ
(NP layolapK لَيْلَةٍ))

أَلْفُ لَيْلَةٍ

>alof+u + layolap+K
one thousand + night
One thousand nights

(NP رَغْمَ::ragoma::in spite of/despite
(NP (NP مَضَى::muDiy~+i::expiration+[def.gen.]
(NP ثلاثة::valAv+ap+i::three+[fem.sg.]+[def.gen.]
(NP أعوام::>aEowAm+K::years+[indef.gen.]))
(PP على::EalaY::on/above
(NP اغتيال::igotiyAl+i-::assassination/murder+[def.gen.]
(NP هم::-him::their))))))

رَغْمَ مَضَى ثَلَاثَةَ أَعْوَامٍ عَلَى إِغْتِيَالِهِمْ

ragoma + muDiy~+i + valAvap+i + >aEowAm+K + Ealay + <igotiyAl+i+him
despite + expiration + three + years + on + assassination+their
Despite three years have passed since they were assassinated

(S (VP participate_with/share_with+he/it |\$Araka | شارك
 (PP-CLR in |fiy- | في-
 (NP it/them/her |-hA | -ها))
 (NP-SBJ towards/approximately |naHowa | نحو
 (NP 200
 (NP child/boy |TifolK | طفل
 and |wa- | و-
 young_girl |-fatAapK | فتاة))))))
 شارك فيها نحو 200 طفل وفتاة
 \$Arak+a + fiy+hA + naHowa + 200 + Tifol+K + wa+fatAp+K
 participate + in+it + approximately + 200 + child + and+girl
Approximately 200 boys and girls have participated in it

(S (VP it/they/she+be |takuwna | تكون
 (NP-SBJ value/worth |qiymapu | قيمة
 (NP the+contract/agreement |AlEaqodi | العقد))
 (NP-PRD 900
 (NP thousand Alf ألف
 (NP dollar |duwlArK | دولار))))))
 تكون قيمة العقد 900 الف دولار
 takuwn+u + qiymapu+u + Al+Eaqod+i + 900 + >alof + duwlAr+K
 be + worth + the+contract + 900 + thousand + dollar
The contract is worth 900 dollars

N.B.: for more example on numbers constructions see section ?????

3.2.1.1.2.2 Quantifiers

This is the list of quantifiers formerly treated as prepositions:

حوالي HawAlaY
 حوالي HawAlay
 زهاء zuhA'a
 قرابة qurAbapa
 نحو naHowa
 قاب qAba
 قيس qayosa
 قيد qayoda

(NP بعض baEoDu
 (NP العلماء Al+EulamA'i))

بعض العلماء

baEoD+u + Al+EulamA'i
 some + the+scholars
some scholars

(NP kilaA كِلَا
(NP Alxabarayoni الْخَبْرَيْنِ))

كِلَا الْخَبْرَيْنِ

kilaA + Al+xabar+ayoni
both + the+news+two
Both pieces of news

(NP >ay~u أَيُّ
(NP Al+rijAli الرِّجَالِ))

أَيُّ الرِّجَالِ

>ay~+u + Al+rijAl+i
any + the+men
Any of the men

(S and |wa- | و-
(VP he/it+insist/assert |~yuSir~u | يُصِرُّ-
(NP-SBJ Denktash |dinokoTA\$ | دِنْكُطَاشْ)
(PP-CLR on/above |EalaY | عَلَى
(NP (NP sovereignty/supremacy |siyAdapK | سِيَادَة
complete/full/integral |kAmilapK | كَامِلَة)
(PP for/to |li- | ل-
(NP every/all/each_one |-kul~i | كُلَّ-
(NP collection/group/bloc |
majomuwEapK | (مَجْمُوعَة)))))))
ويصر دنكطاش على سيادة كاملة لكل مجموعة
wa+yuSir~+u + danokoTA\$ + EalaY + siyAdap+K + kAmilap+K + li+kul+i +
majomuwEap+K
and+insist + Denktash + on + sovereignty + complete + for+every + group
Denktash insists on a complete sovereignty on every group

(S-ADV (VP |UNmlthmAF |UTF8UN ملتهماً
(NP-SBJ-2 *)
(NP-OBJ approximately/around |HawAlaY | حِوَالِي
(NP 250
(NP (NP dunum_ |duwnumAF | دُونُماً)
(PP from |min | مِنْ
(NP the+trees |Al>a\$ojAri | الْأَشْجَارِ
(NP the+trees |Al>a\$ojAri | الْأَشْجَارِ)))))))
ملتهماً حوالي 250 دونماً من الأشجار
multahim+AF + HawAlay + 250 + duwnam+AF + min + Al+>a\$ojAr+i
gormandizing + approximately + 250 + dunum + from + the+trees
Gormandizing approximately 250 dunums of trees

3.2.1.1.2.3 Nouns formerly treated as prepositions or particle

The following nouns were annotated as prepositions in earlier versions of the Arabic Treebank. They are now annotated as nouns, at both the POS level (with the core POS tag NOUN) and at the treebank level (heading an NP):

>amAma, إثرَ <ivora, إزاء <izA'a, بَعْدَ baEoda, بَيْنَ bayona, تُجَاهَ tujAha, تَحْتَ taHota, تِلْوَا tilowa, حَذُوَ Ha*owa, حَوْلَ Hawola, حِينَ Hiyna, خَلْفَ xalofa, ضِمْنَ Dimona, عَقِبَ Eaqiba, عَبْرَ Eabora, عِنْدَ Einoda, فَوْرَ fawora, فَوْقَ fawoqa, قَبْلَ qabola, قُبَايَلَا qubayola, قُبَالَا qubAlapa, قُرْبَ quroba, مَعَ maEa, حِينَ Hiyna, أَمْتَال >amovAl, أَثْنَاء >avonA'a, ضِدَّ Did~a, طَوَالَ TiwAla, عَوْضَ EiwaDa, حَسَبَ Hasaba, وَفَقَ wifoqa.

The following nouns were annotated as particles in earlier versions of the Arabic Treebank. They are now annotated as nouns, at both the POS level (with the core POS tag NOUN) and at the treebank level (heading an NP):

غير gayor,
سوى siwaY

3.2.1.1.2.3.1 Nouns expressing possession or accompaniment:

لدى/ladAY, عند/ Einda, مع/ maEa

```
(SBARQ fa-::ف-::and/so
  (WHADVP-3 -kayofa::كَيْف-::how )
  (S (VP ya+kuwn+u::يَكُونُ::he/it+be+[ind.]
    (NP-PRD laday-::لَدَي-::with/by
      (NP -hi-::ه-::it/him ))
    (NP-SBJ >asoraY::أَسْرَى::prisoners/captives )
    (ADVP-MNR-3 *T*)) ) )
```

فكيف يكون لديه أسرى

fa+kayfa + yakuwn+u + laday+hi + >asoraY

so+how + be + having + prisoners of war

So how can he have prisoners of war?

```
(S (VP -ya+taEAwan+u::-يَتَعَاوَنُ::he/it+cooperate+[ind.]
  (NP-SBJ *)
  (NP-CLR maEa::مَعَ::with
    (NP (NP Al+>amiyn+i::الأمين::the+secretary+[def.gen.]
      Al+EAm~+i::العام::the+general/common/public+[def.gen.] )
    (PP li-::ل-::for/to
      (NP -jAmiE+ap+i::-جَامِعَة::league+[fem.sg.]+[def.gen.]
        (NP Al+duwal+i::الدُول::the+states/countries+[def.gen.]
          Al+Earabiy~+ap+i::العَرَبِيَّة::
            the+Arab/Arabic+[fem.sg.]+[def.gen.]
          ) ) ) ) ) ) ) )
```

يتعاون مع الأمين العام لجامعة الدول العربية

yataEAwan+u + maEa + Al+>amiyn+i + Al+EAm+i + li+jAmiEap+i + Al+duwal+i
+ Al+Earabiy~+ap+i

cooperate + with + the+secretary + the+general + for+league + the+states +
the+Arab

He cooperate with the general secretary of the Arab countries' league

3.2.1.1.2.3.2 Nouns of similitude

mivl *sameness*, \$iboh *similarity*, \$abiyh *resemblance*, naZiyr *counterpart*, maviyl *equal*

(S (VP arrive/come/occur+he/it |jA'a | جاء
 (NP-SBJ some/several |baEoDu- | بَعْضُ-
 (NP their |-hum | -هُم))
 (PP-LOC from |min | مِنْ
 (NP (NP places/locations/positions |>amAkina | أَمَاكِنَ
 remote/distant/far |baEiydapK | بَعِيدَة
 (NP-ADV like/such_as |mivola | مِثْلُ
 (NP California |kAliyfuwroniyA | كاليفُورْنِيَا))))))
 جاء بعضهم من أماكن بعيدة مثل كاليفورنيا
 jA'+a + baEoD+u+hum + min + >amAkin+a + baEiydap+K + mivola +
 kAliyfuwroniyA
 came + some+them + from + places + remote + such as + California
Some of them came from remote places like California

(NP مثل mivola
 (SBAR-NOM (WHNP-1 الذي Al~a*iy)
 (S (VP رأيت ra>ayota
 (NP-SBJ *)
 (NP-OBJ-1 *T*)))))

مثل الذي رأيت

mivol+a + Al~a*iy + ra>ayot+a
 Like + what + saw
As you saw

3.2.1.1.2.3.3 Nouns of exception

gayr *otherness*, siwaY *otherness*

(S l`kin | ل`كن | however
 (ADVP-LOC there_ |vam~apa | ثَمَّة)
 (NP-TPC-1 currents/streams |tay~ArAtN | تيارات
 strong/powerful |qawiy~apN | قويَّة)
 (VP it/they/she+ happen/take_place |taHoduvu | تَحْدُثُ
 (NP-SBJ-1 *T*)
 (NP-OBJ records/traces/effects | |vAru- | آثارُ-
 (NP its/their/her |-hA | -ها))
 (PP-DIR in |fiy | في
 (NP not/other |gayori | غَيْرُ
 (NP that |*`lika | ذَٰلِكَ
 the+direction/course |AlAit~ijzAhi | (الاتِّجَاهُ |))))))
 لكن ثمة تيارات قوية تحدث آثارها في غير ذلك الإتجاه

A case of relative clause modifying indefinite noun

l`kin + vam~apa + tay~ArAt+N + qawiy~ap+N + taHoduv+u + |vAr+u+hA + fiy +
 gayor+i + *`lika + Al+<it~ijAh+i
 but + there are + currents + strong + happen + effects+its + in + other + that +
 the+direction

But there are strong currents whose effects happen in another direction

gayor can also be followed by an adjective in nominal position:

(S (VP he/it+mean/signify/concern |yaEoniy | يَعْني
 (NP-SBJ-2 *T*)
 (SBAR that |>an~a | أَنَّ
 (S (NP-SBJ the+list/table/schedule+two |All~A}iHatayoni | اللَّائِحَتَيْنِ
 the+competing/opposing+two |AlmutanAfisatayoni |
 (الْمُتَنَافِسَتَيْنِ)
 (NP-PRD not/other |gayoru | غَيْرُ
 (NP completed/finished+two | mukotamalatayoni |
 (مُكْتَمَلَتَيْنِ))))))
 يعني أن اللائحتين المتنافستين غير مكتملتين

yaEoniy + >an~a + Al+|A}iHat+ayoni + Al+mutanAfisat+ayni + gayor+u +
 mukotamalat+ayni

mean + that + the+list+two + the+competing+two + not + completed+two

It means that the two competing lists are not complete

Gayor and siwaY can also have an SBAR or an SBAR-NOM complement.

(S (VP {it~aSal+tu::تُصَلِّتُ::contact/get_in_touch_[with]+I_[verb])
 (NP-SBJ *)
 (NP-ADV mirAr+AF::مِرَاراً::often+[acc.indef.])
 (PP-CLR bi-::بِ-::with/by
 (NP (NP -Al+maso&uwl+iyana::-المُسْؤُولِينَ::the+official/
 functionary+[masc.pl.gen.])
 (PP fiy::فِي::in
 (NP Al+maSolaH+ap+i::المَصْلَحَةِ::the+department/
 agency+[fem.sg.]+[def.gen.]))))
 ,::,::nogloss
 (NP-ADV gayor+a::غَيْرَ::not/other+[def.acc.]
 (SBAR >an~a-::أَنَّ-::that
 (S (NP-TPC-1 -hum::هم::they_[masc.pl.])
 (VP (PRT lam::لَمْ::did_not
 yu+naf~i*+uwA::يُنْفِذُوا::they_[people]+implement/
 carry_out/accomplish+[masc.pl.]
 (NP-SBJ-1 *T*)
 (NP-OBJ \$ayo}+AF::شَيْئاً::something/thing+
 [acc.indef.])
 (NP-ADV ragoma::رَغْمَ::in spite of/despite
 (NP (NP wuEuwd+i-::وَعُودٍ-::promises+
 [def.gen.]
 (NP -him::هم::their)
 (PP li-::لِ-::to/for
 (NP -y::ي-::me))))))))
 إتصلت مراراً بالمسؤولين في المصلحة غير أنهم لم ينفذوا شيئاً رغم وعودهم لي
 <it~aSalot+u + mirAr+AF + bi+Al+mas&uwl+iyana + fiy + Al+maSolaHap+i + gayora +
 >an~a+hum + lam + yunaf~i*uwA + \$ayo}+AF + ragoma + wuEuwd+i+him + l+iy
 contact + often + with+the+officials + in + the+department + other + that+they + did not +
 accomplish + thing + despite + promises+their + for+me
*I often contacted the department officials but they did not accomplish anything despite their
 promises for me*

(NP gayora غَيْرَ)
 (SBAR-NOM (WHNP-1 Al~a*iyana الَّذِينَ)
 (S (VP mar~uwA مَرُّوا
 (NP-SBJ-1 *T*)))))

غير الذين مرُّوا

gayor+a + Al~a*iyana + mar~uwA]
 except + who + came
Except those who came

(NP siwaY سِوَى)
 (SBAR-NOM (WHNP-1 man مَنْ)
 (S (VP Zalama ظَلَمَ
 (NP-SBJ-1 *T*)))))

سوى من ظلم

siwaY + man + Zalam+a
 except + who + maltreated
Except the one who maltreated

3.2.1.1.2.3.4 Nouns expressing locative or temporal relationships

These nouns were formerly treated as prepositions (they now have NOUN POS tags, not PREP POS tags).

أمام >amAma, إزاء <ivora, إزاء <izA'a, بعد baEoda, بين bayona, تجاه tujAha, تحت taHota, عقب عقِب Dimona, خلف خَلَفَ xalofa, حين Hiyna, حول حَوَلَ Hawola, حذو حَذَوَ Ha*owa, تلو تَلَوَ tilowa, عابراً Eabora, عندَ Einoda, فوقَ fawora, فوقَ fawoqa, قبلَ qabola, قبيلَ qubayola, قبالة qubAlapa, قربَ quroba, معَ maEa, حين Hiyna,

(S (VP stop/halt/depend_on+it/they/she |tawaq~afat | تَوَقَّفت |
(NP-SBJ-2 *)
(NP-ADV with/at |Einoda | عندَ
(NP (NP the+commemoration/remembrance|Al*~ikoraY | الذِّكْرَى |
(ADJP the+fourth |Alr~AbiEapi | الرَّابِعَةَ
and |wa- | وَ-
the+fifty/fiftieth |-Alxamosiyna | الخمسين-)
(PP for/to |li- | لـ-
(NP rape/usurpation |-{igotiSAbi | اغْتِصَاب-
(NP Palestine |filasoTiyna | (فِلَسْطِينَ))))))
توقفت عند الذكرى الرابعة والخمسين لإغتصاب فلسطين

tawaq~afat + Einoda + Al+*ikoraY + Al+rAbiEap+i + wa+Al+xamsiyna +
li+<igotiSAb+i + filasoTiyna
stop + at + the+remembrance + the+fourth + and+the+fiftieth + li+usurpation +
Palestine

It stopped at the fifty fourth remembrance of the usurpation of Palestine.

Some nouns of this category can have an SBAR complement headed by >an or mA, called in traditional Arabic grammar "AlHarofAn AlmaSodariy~An" (the two particles of MaSodar).

(NP baEoda بَعْدَ
(SBAR >an أَنْ
(S (VP taroHala تَرَحَّلَ)
(NP-SBJ *)))))

بعد أن ترحل

baEoda + >an + taroHal+a
after + that + leave
After you leave

(NP baEoda بَعْدَ
(SBAR mA مَا
(S (VP taroHalu تَرَحَّلَ)
(NP-SBJ *)))))

بعد ما ترحل

baEoda + mA + taroHal+a
after + that + leave
After you leave

(S (NP-TPC-1 Bush |buw\$ | بُوش)
 (VP he/it+sleep/lie_down |yanAmu | ينامُ
 (NP-SBJ-1 *T*)
 (NP-TMP after |baEoda | بَعْدَ
 (SBAR to |>an | أَنْ
 (S (VP he/it+watch/observe/witness |yu\$Ahida | يُشَاهِدُ
 (NP-SBJ *)
 (NP-OBJ the+news/reports |Al>axobAra | الْأَخْبَارُ
 the+local |AlmaHal~iy~apa | الْمَحَلِّيَّةُ))))))
 بوش ينام بعد أن يشاهد الأخبار المحلية

buw\$ + yanAm+u + baEoda + >an + yu\$Ahid+a + Al+>axobAr+a +

Al+maHal~iy~ap+a

Bush + sleep + after + that + watch + the+news + the+local

Bush sleeps after he watches the local news

(S (NP-ADV at the time/ near/ upon|Einoda | عِنْدَ
 (SBAR to /that |mA| مَا)
 (S (VP enter+he/it |daxala | دَخَلَ
 (NP-SBJ the+Arabs |AlEarabu | الْعَرَبُ
 the+Muslim |Almusolimuwna | الْمُسْلِمُونَ)
 (PP-DIR to/towards |<ilaY | إِلَى
 (NP Spain |<isobAniyA | إِسبَانِيَا))))))
 (VP be/was/were+he/it |kAna | كَانَ
 (NP-SBJ (NP the+cover/blanket |AlgiTA'u | الْغِطَاءُ
 the+ideological |Al<iydiyuwluwjy~u | الْإِيدِيُولُوجِيّ)
 (PP for/to |li- | لِ-
 (NP (NP this |~h`*A | هَذَا)
 (NP the+entrance/penetration |Ald~uxuwli | الدُّخُولُ))))
 (NP-PRD spreading/propagation |na\$ora | نَشْرُ
 (NP the+missionary_work/Da'wah |Ald~aEowapi | الدَّعْوَةُ
 the+Islamic/Islamist/Muslim |Al<isolAmiy~api | الْإِسْلَامِيَّةُ))))
 عندما دخل العرب المسلمون الى إسبانيا كان الغطاء الأديولوجي لهذا الدخول نشر الدعوة الإسلامية

Einda+mA + daxal+a + Al+Earab+u + Al+musolimuwna + <ilaY + <isobAniyA +

kAn+a + Al+giTA'u + Al+>adiyuwluwjy~+u + li+h`*A + Al+duxuwl+i +

na\$or+a + Al+daEowap+i + Al+<isolAmiy~ap

when+that + enter + the+Arabs + the+Muslims + Spain + was + the+cover +

the+idiologic + for+this + the+entrance + propagation + the+missionary +

the+Islamic

*When the Muslim Arabs entered Spain, their entrance was covered by an ideology
 of propagation the Islamic missionary*

It is also possible to have the NOUN حين |Hiynal time, as well as other NOUNs indicating temporal unit as يوم |yawm |day, ساعة |sAEap| hour, اللحظة |laHoZap| instant, or NOUNs referring to period of time in general like وقت |waqot| , عصر |EaSor| , زمن |zaman| with an S complement.

(NP حين Hyna
(S (VP ترحل taroHalu
(NP-SBJ *))))

حين ترحل

Hiyna + taroHal+u

when + leave

When you leave

(S (VP >afoD+at::أَفَضْتُ::inform/notify/lead/take+it/they/she_[verb]
(PP-CLR <ilay-::إِلَى::to/towards
(NP (NP -hi::هِ::it/him)
(NP-2 *T*)))
(NP-SBJ juhuwd+u::جُهُودُ::efforts+[def.nom.]
(NP -hu ::هُ :: him)))
(NP-TMP Hiyna::حِينَ::when
(S (VP EAd+a::عَادَ::he/it+return/go_back+[ind.]
(NP-SBJ *)
(PP-MNR bi-::بِ-::by/with
(NP -xuf~+ayo::-خُفِّي::slippers+two_[acc.]
(NP Hnyn::حَنِين::NOT_IN_LEXICON))))))))
أَفَضْتُ إِلَيْهِ جُهُودَهُ حِينَ عَادَ بِخُفِّي حَنِينَ

>afoDat + <ilay+hi + juhuwd+a+hu + Hiyna + EAd+a + bi+xuf~+ayo + Haniyn

inform + to+him + with+slippers+two + Haneen

Informed him about Haneen's slippers

(NP-TMP yawoma يَوْمَ
(S (VP yanofaEu يَنْفَعُ
(NP-OBJ AlnAsa النَّاسَ)
(NP-SBJ Sidoqu صِدْقُ
(NP hum هُمْ)))))

يَوْمَ يَنْفَعُ النَّاسَ صِدْقُهُمْ

yawom+a + yanofaE+u + Al+nAs+a + Sidoq+u+hum

day + benefits + the+people + honesty+their

A day when people get benefit from their honesty

(S (VP him/it_to+remain/last |ybqY |UTF8UNيبقى
(NP-SBJ *)
(ADJP-PRD holding_together/cohesive/calm |mutamAsikAF |مُتَمَاسِكاً
(NP-TMP day |yawoma |يَوْمَ
(S (VP it/they/she+increase/exceed/add |taziydu |تَزِيدُ
(NP-SBJ the+pressures/stresses | AlD~uguwTu |الضُّغُوطُ
(PP on/above |Ealay- |عَلَى-
(NP it/him |-hi | -هِ))))))

يبقى متماسكاً يوم تزيد الضغوط عليه

yaboqaY + mutamAsik+AF + yawom+a + taziyd+u + Al+DuguwT+u + Ealay+hi

remain + cohesive + day + exceed + the+pressures + on+him

He remains calm when the pressures exceed on him

NB: muno*u, which is in the current list of true prepositions, can also be tagged as a noun heading an NP when it is completed by S or SBAR because it is a noun indicating a specific time (*an instant*).

(NP since/starting_from | -muno*u | مُنْذُ -
 (S (VP be/was/were+it/they/she | kAnat | كَانَتْ
 (NP-SBJ the+league | AljAmiEapu | الجامعة
 the+Arab/Arabic | AlEarabiy~apu | الْعَرَبِيَّةُ))))
 منذ كانت الجامعة العربية

muno*+u + kAnat + Al+jAmiEap+u + Al+Earabiy~ap+u
 since + was + the+league + the+Arab
Since the Arab league was there

(NP since/starting_from | muno*u | مُنْذُ
 (SBAR to | >an | أَنْ
 (S (VP announce/declare+it/they/she | >aEolanat | أَعْلَنْتُ
 (NP-SBJ the+nations/peoples | Al>umamu | الأمم
 the+united | Almut~aHidapu | الْمُتَّحِدَةُ
 (NP-OBJ the+line | AlxaT~a | الخط
 (NP (the+blue | Al>azoraqa | الْأَزْرَقُ)))))))
 منذ ان أعلنت الأمم المتحدة الخط الأزرق

muno*+u + >an + >aEolanat + Al+>umam+u + Al+mut~aHidap+u + Al+xaT+a +
 Al+>azraq+a
 since + that + announced + the+nations + the+united + the+line + the+blue
Since the United Nations announced the blue line

3.2.1.1.2.4 Idiomatic annexation/idafa phrases

*uw and its inflection in gender and number (see list below) 'possessor of' are always the head of <iDAfap (annexation), and they function attributively (in apposition) or predicatively: *uw , *aAt, *awaA, *awuw, *awAt, >uwlw, >uwlaAt (Badawi, 139).

(NP (NP Aimora>apN (إِمْرَأَةٌ)
 (NP *Atu ذات)
 (NP jamAlK جمال)))
 امرأة ذات جمال

<imora>ap+N + *At+u + jamAl+K
 woman + possessor + beauty
A woman of beauty

(S (NP-SBJ zayodN زَيْدٌ)
 (NP-PRD *uw ذُو
 (NP maAlK مال)))
 زيد ذو مال

zayod+N + *uw + maAl+K
 Zayd + possessor + money
Zayd has money

(S (NP-TPC-1 (NP the+human_being |Al<inosAnu | الإنسان)
 (NP of/with |*uw | ذو
 (NP the+desire/will |Al<irAdapi | الإرادة)))
 (NP-PRD with/by |laday- | لذي-
 (NP (NP it/him |-hi | -ه)
 (NP-1 *T*)))
 (NP-SBJ (NP the+jug/pot |Alqidorapu | القِذْرَة)
 (PP on/above |EalaY | على
 (NP resistance |Al+taHam~ul+i | التحمل))))
 الإنسان ذو الإرادة لديه القدرة على التحمل
 Al+<inosAn+u + *uw + Al+<irAdap+i + laday+hi + Al+qudorap+u + EalaY +
 Al+taHam~ul+i
 the+human being + with + the+will + with+him + the+ability + on + resistance
 Human being has the will; he is able to resist

(PP on/above |EalaY | على
 (NP (NP portion/section |qisomK | قِسم
 (PP from |min | من
 (NP (NP Kashmir |ka\$omiyra | كَشْمِير
 (NP same/self/essence |*Ati | ذات
 (NP (NP the+majority |AlgAlibiy~api | الغالبية
 (PP from |min | من
 (NP the+Muslim |Almusolimiyna | المسلمين))))))))
 على قسم كبير من كشمير ذات الغالبية من المسلمين
 Ealay + qisom+K + min + ka\$omiyr + *Ata + Al+gAlibiy~ap+i + min + Al+muslimiyna
 on + part + from + Kashmir + same + the+majority + from + the+Muslims
 On a part of Kashmir of the Muslim majority

3.2.1.1.2.5 Prayer with gerund (masdar) completed by noun or bound pronoun

(list provided by Ibn Yaish (sp???)):

suboHAna سبحان
 maEA*a معاذ
 lab~ayo +ka لبي+ك
 HanAnayo+ka حناني+ك
 suEodayo+ka سعدي+ك
 Eamoru+ka عمر+ك

(NP suboHAna سُبحان
 (NP All~ahi الله))

سبحان الله

suboHAN+a + All~ah+i
 glory + God
 Glory be to God

(PP-DIR to/towards |<ilaY | إلى
 (NP (NP Allah/God |All~`hi | اللّٰه
 (NP **praise** |suboHana- | سُبْحَانَ-
 (NP its/his |-hu | -هُ)))

إلى الله سبحانه

<ilaY + >all~`h+i + subHAn+a+hu

to + God + glory+his

To the glorious God

NB: is the level and tag for this suboHana NP correct? Is it in apposition to Allah?

3.2.1.2 Accusative complements: tamiyz complements of numbers (11-99)

Pre-nominal cardinal numbers (11-99) are completed by an accusative, indefinite NP in a construction called tamoyyiz (see section ??? about Numbers for more examples of numerical expressions).

(NP (NP debts |-duyuwnu | دُيُونُ-
 (NP the+Hariri |AlHaryriy~ | الحريري))
 (SBAR (WHNP-3 which/who/whom |Al~atiy | الَّتِي)
 (S (VP arrive_at/reach/connect+it/they/she |waSalat | وصلت
 (NP-SBJ-3 *T*)
 (PP-CLR to/towards |<ilaY | إلى
 (NP (NP vlAvyn ثلاثين
 (NP **billion** |miloyArAF | (ملياراً))
 (PP from |min | من
 (NP the+dollars |Ald~uwlArAti | الدُولارات
 the+American |Al>amiyrokiy~api | الأَمِيرَكِيَّة
)))))))
 ديون الحريري التي وصلت إلى ثلاثين ملياراً من الدولارات الأمريكية
 duyuwn+u + Al+Haryriy~ + Al~atiy + waSalat + <ilaY + valAvyn + miloyAr+AF + min +
 Al+duwlArAt + Al+>amiyrokiy~ap+i
 debts + the+Hariri + which + reached + to + thirty + billion + from + the+dollars +
 the+American
Al-Hariri's debts that reached thirty billion American dollars

3.2.2 Complements of adjectives

Complements/arguments are genitive, obligatory, and for deverbals head participles clausal constituents that would be arguments of the verb that the participle derived from.

Most of the time the genitive complement of the adjective carries the meaning that the subject or the object of the verb that the adjective is derived from would have.

False idafa, where the first term is an adjective (ADJ) that modifies the second term (the adjectival first term includes adjectives, active participles, passive participles, comparative adjectives and numerical/ordinal adjectives)

- semantic head
- + syntactic head
- + adj
- noun

The adjective head should be indefinite/definite without tanween.

(ADJP muDaf or the annexed ADJ
(NP muDaf <ilayh or the annexed-to-it Noun))

False idafa will *always* be an ADJP.

3.2.2.1 Semantic specificity of the non-true construct state

This kind of complementation is called in traditional Arabic grammar <iDAfap gayr Haqiyqiy~ap (annexation not true) because the semantic relation between the adjective head and the complement is similar to the relationship between the verb and its argument (SBJ or OBJ) as we have explained above. But given that the adjective assigns to its argument the genitive case, it is considered as a head of construct state.

In the true construct state, the relationship between the head noun and its complement is one of possession, locality, temporality, origin, material or containment. However, in the non-true construct state (or “false idafa”), in addition to the POS category being different (the syntactic head is an adjective rather than a noun) the semantic relationship between the head and its complement is one of patient or agent. This difference in semantic relationship leads to the “non-true” or “false” label for this kind of complementation. Semantically, this kind of relationship does not meet the definition of idafa/construct state (hence it is “non-true” or “false”); syntactically the relationship follows the case assignment and obligatoriness parts of the idafa/construct state definition (hence it is still “idafa”). Therefore, this kind of complementation is called “false idafa” or the “non-true construct state.”

Besides the difference in the semantic relation between the head and the complement between the true and the false construct state, there is another important point of distinction which is that the complement in the true construct state doesn't allow the NP head to have neither the tanwin (sign of indefiniteness) nor the DET Al (sign of definiteness). The adjective head in a false idafa construction, however, can have the determiner Al-, to agree for example with the external noun that the adjective modifies.

Examples of true idafa: ungrammatical with Al, ungrammatical with tanwiyn.

True idafa:

كتاب زيد

kitAb+u + zayod+K
book + Zayd
Zayd's book

Ungrammatical with Al-:

الكتاب زيد

*Al+kitAb+u + zayod+K
*the+book + Zayd
**The book Zayd*

Ungrammatical with tanwiyn:

كِتَابُ زَيْدٍ

*kitAb+N + zayodK
*book + Zayd
**A book Zayd*

However, the adjective head in the false idafa construct agrees in definiteness with the noun that it modifies, as all adjectives do. So it can be indefinite without Al- (but also without tanwiyn because it is the head of a construct state).

Examples of false idafa: grammatical without and also with Al-, ungrammatical in any case with tanwiyn.

False idafa, grammatical without Al-:

(S (VP jA'a جاء
(NP-SBJ (NP rajulN رَجُلُ)
(ADJP TawiyLu طَوِيلُ)
(NP AlqAmapI الْقَامَةِ))))

جاء رجل طويل القامة

jA'+a + rajul+N + TawiyL+u + Al+qAmap+i
came + man + tall + the+stature
A man with a tall stature came

Grammatical also definite with Al-:

(S (VP jA'a جاء
(NP-SBJ (NP Alrajulu الرَّجُلُ)
(ADJP AlTawiyLu الطَّوِيلُ)
(NP AlqAmapI الْقَامَةِ))))

جاء الرجل الطويل القامة

jA'+a + Al+rajul+u + Al+TawiyL+u + Al+qAmap+i
came + the+man + the+tall + the+stature
The man with a tall stature came

Ungrammatical with tanwiyn:

جاء رجل طويل القامة

*jA'a + rajul+N + Tawiyul+N + Al+qAmap+i

*came + man + tall + the+stature

*A man a tall stature came

Note that a false idafa ADJP can occur in nominal position (see section ???? on adjectival heads). When this is the case, the ADJP heads an NP (just as any other adjective or ADJP can head an NP when it occurs in nominal position, see section ??? for a description of possible nominal positions and more examples of adjectives heading NPs).

(S (VP jA'a جاء
(NP-SBJ (ADJP Tawiyul طویل
(NP AlqAmap (القامة))))))

جاء طويل القامة

jA'a + Tawiyul+u + Al+qAmap+i

came + tall + the+stature

The one with a tall stature came

3.2.2.2 Adjectival head of non-true construct state (false idafa)

An adjective head of the non-true construct state or false idafa can be an adjective that is what is called in Arabic grammar "adjectival noun" (>asomA' SifAt) or what Wrigth called "verbal adjective," which includes the following:

- **participles (active and passive)**

(ADJP \$Aribu شارب
(NP AlHalybi الحليب))

شارب الحليب

\$Arib+u + Al+Haliyb+i

drinker + milk

The milk drinker

(ADJP Al+\$Aribu الشارب
(NP Al+Haliyb (الحليب)))

الشارب الحليب

Al+\$Arib+u + Al+Haliyb+i

the + drinker + the+milk

The one who drinks the milk

(NP appointing/assignment |taEoyiyni | تَعْيِين
 (NP (ADJP working/functioning |\$Agiliy | شاغلي
 (NP (NP these_two |hAtayoni | هَاتَيْنِ)
 (NP the+job/position/function+two |AlwaZiyfatayoni | الوظيفَتَيْنِ
 the+important/serious+two |Almuhim~atayoni | المُهِمَّتَيْنِ
))))

تعيين شاغلي هاتين الوظيفتين المهمتين

taEoyiyn+i + \$Agiliy + hAtayoni + Al+waZiyfat+ayoni + Al+muhim~at+ayni
 assignment + functioning + these two + the+position+two + the+important+two
Assignment of functioning for these two important positions

(ADJP makosuwru مَكْسُورُ
 (NP janAHiK جَنَاحِ))

مَكْسُورُ جَنَاحِ

makosuwru+u + janAH+K
 broken + wing
Someone with a broken wing

(ADJP Almakosuwru المَكْسُورُ
 (NP AljanaAHi الجَنَاحِ))

المَكْسُورُ الجَنَاحِ

Al+makosuwru + Al+janaAHi
 the+broken + the+wing
The one with the broken wing

- **static adjective (sifAt mu\$ab~hap bi Aism Al fAEil / adjectives....to the active participles, Aisom nisobap and Sigap mubAlagap/ form of exaggeration):**

(ADJP Jamiylu جَمِيلُ
 (NP Alwajhi الوَجْهِ))

جَمِيلُ الوَجْهِ

Jamiyl+u + Al+wajoh+i
 beautiful + the+face
The one with a beautiful face

(ADJP Al+qaliylu القَلِيلُ
 (NP Al+maAli المَالِ))

القَلِيلُ المَالِ

Al+qaliyl+u + Al+mAl+i
 the+little + the+money
The one with little money

- **form of exaggeration (Sigap Al mubAlagap)**

(ADJP Ham~Alu حَمَّالُ
 (NP Al+>lowiyapi الأَلْوِيَّةِ))

حَمَّالُ الأَلْوِيَّةِ

Ham~Al+u + Al+>alowiyap+i
 holder + the+flag

The holder of the flag

- **adjective expressing belonging to a space, a trend, religion ... (Aism Al nisobap/ noun of belonging or relationship)**

(ADJP filasoTiyniy~u فلسطيني
(NP Al+mawlid المولد))

فلسطيني المولد

filasoTiyniy~u + Al+mawolid+i
Palestinian + the+birth
Palestinian by birth

- **comparative adjective (with the core POS tag ADJ_COMP) when it means superlative**

(ADJP >ajmalu أجمل
(NP hum هم))

أجملهم

>ajomal+u+hum
most beautiful+their
The most beautiful one among them

(ADJP xayoru خَيْرُ
(NP Al+n~aAsi الناس))

خير الناس

xayor+u + Al+nAs+i
best + the+people
The best one among people

(ADJP >asoraEu أَسْرَعُ
(NP Al+mutasaAbiqiyina المتسابقين))

أَسْرَعُ المتسابقين

>asoraE+u + Al+mutasaAbiqiyina
fastest + the+runners
The fastest runner

(S (VP he/it+participate_with/share_with |yu\$Ariku | يُشارك
(PP-CLR in |fiy- | في-
(NP it/them/her |-hA | -ها))
(NP-SBJ (NP vmAnyp)
(PP from |min | من
(NP (NP (ADJP more/most_famous |>a\$ohari | أشهر
(NP the+player/athlete |All~AEibiyna | اللَّاعِبِينَ))
(PP-LOC in |fiy | في
(NP the+world |AlEAlami | العالم))))))
يشارك فيها ثمانية من أشهر اللاعبين في العالم

yu\$Arik+u + fiy+hA + vamAniyap + min + >a\$ohar+i + Al+lAEibiyna + fiy + Al+EAlam+i
participate + in+it + eight + from + the+most famous + the+players + in + the+world
Eight of the most famous players in the world participate in it

When the comparative adjective is in a nominal position, as in the example above (complement of a preposition), we assume that this comparative adjective is modifying a missing noun, and we show this in the tree by adding an extra NP level before the ADJP node.

- **Pre-nominal numerical/cardinal adjective**

(ADJP >aw~alu أول
(NP Al>um~ahaAti ()))

أَوَّلُ الْأُمّهَاتِ

>aw~al+u + Al+>um~ahaAt+i
first + the+mothers
The first one among the mothers

(ADJP sAbiEu سابع
(NP bAbK باب))

سَابِعُ بَابٍ

sAbiE+u + bAb+K
seventh + door
Seventh door

(ADJP vaAlivu ثالِثُ
(NP hum هُم))

ثَالِثُهُمْ

vAliv+u + hum
third+their
Their third

(S (NP-SBJ it/they/she |-hiya | هِيَ-)
(ADJP-PRD fifth |xAmisapu | خَامِيسَةُ
(NP championship/tournament |daworAti | دَوْرَاتِ
(NP the+Masters |AlmAsotiroz | (الْمَاسْتِرْزُ)))

هي خامسة دورات الماسترز

hiya + xAmisap+u + daworAt+i + Al+mAsotiroz
it + fifth + tournament + the+Masters
It is the fifth Masters tournament

3.3 Adjuncts

Adjuncts are descriptive, not possessive, not obligatory. In addition, *all* PPs, ADJPs and other modifiers of NP are shown as adjuncts.

Adjunct constituents are sister to the NP that contains the head noun, child of the NP that contains both: (NP (NP head) (NP adjunct)). For the most part, we do not distinguish among levels or "scope" of modification – all adjuncts are at the same level, sisters of the head NP, with the exception of relative clauses and appositive NPs. An extra level of NP adjunction is added for the relative clause SBAR and/or for the appositive NP.

3.3.1 Adjectival modification

The agreement between the head noun and a modifying adjective is in number, gender, definiteness, and case.

3.3.1.1 Single token noun, modified by a single token adjective

A single token noun modified by a single token adjective will be annotated as a flat NP:

```
(S (VP turn/attract+he/it | -lafata | لَفَتَ-
  (NP-SBJ *)
  (PP-CLR to/towards | <ilaY | إِلَى
    (NP necessity/need/imperative | Daruwrapi | ضُرُورَة
      (NP discovery/finding | <iyjAdi | إِجَاد
        (NP (NP projects/enterprises | ma$AriyEa | مَشَارِيع
          vital/lively/biological | Hayawiy~apK | حَيَوِيَّة
            (PP-LOC in | fiy | فِي
              (NP the+area/zone/territory |
                AlminoTaqapi | (الْمِنْطَقَة))))))
                لفت الى ضرورة إيجاد مشاريع حيوية في المنطقة
lafat+a + <ilaY + Daruwrap+i + <iyjAd+i + ma$AriyE+a + Hayawiy~ap+K
attracted + to + necessity + finding + projects + vital + in + the+area
He derived the attention to the necessity of finding vital projects in the area
```

3.3.1.2 Multi-token NP, modified by a single token adjective

When either the head NP or the modifying adjective phrase is more than one token, the adjunction structure must be shown.

It is particularly important to annotate the modification level correctly – the adjective or ADJP must be with the noun or NP that it modifies. The level of modification will affect the tree structure as follows.

- ADJP modifying the NP head of complementation structure

In this example, the NP adjunction structure must be shown because the head NP is more than one word, even though the adjective is only one word.

```
(S (VP HadAra حَضَرَ
  (NP-OBJ Aljalosapa الْجَلْسَة )
  (NP-SBJ (NP waziyr وزير
    (NP Al<iElAmi الإعلام ) )
    (ADJP AlsAbiqu السابق ) ) )
    حضر الجلسة وزير الإعلام السابق
Hadar+a + Al+jalosap+a + waziyr+u + Al+<iEolAm+i + Al+sAbiq+u
attend + the+session + minister + the+mass information + the+former
The former mass information minister attended the session
```


- Adjective modifying the NP complement

In the example, the NP modified by the adjective is flat, because both the noun being modified and the adjective are single words, even though that NP is in a complement position.

(S (VP \$Araka شارك
(PP-CLR fiy في
(NP AlHafolapi الحفلة))
(NP-SBJ ra}iysu رئيس
(NP AlHizbi الحزب
AldiymoqrATiy~i الديمقراطي))))
شارك في الحفلة رئيس الحزب الديمقراطي
\$Arak+a + fiy + Al+Hafolap+i + ra{iys+u + Al+Hizob+i + Al+diymuqorATiy~+i
participated + in + the+celebration + chairman + the+party + the+democratic
The chairman of the democratic party participated in the celebration

- ADJP modifying the first complement in a complex complementation structure

(S (NP-SBJ huwa هو)
(NP-PRD Aibnu ابن
(NP (NP zawjapi زوجة
(NP Alra}iysi الرئيس))
(ADJP Al>uwlaY الأولى))))
هو ابن زوجة الرئيس الأولى
huwa + ibn+u + zawjap+i + Alra}iys+i + Al+>uwlaY
he + son + wife + the+president + the+first
He is the son of the president's first wife

- with a quantifier

(S (NP-SBJ (NP baEoDu بعض
(NP Al>aTofAlI الأطفال))
(ADJP AlSigAri الصغار))
(ADJP-PRD \$adiyduw شديد
(NP Al*akA'i الذكاء))))
بعض الأطفال الصغار شديدو الذكاء
baEoD+u + Al+>aTofAl+i + Al+SigAr+i + \$adiyduw + Al+dakA'+i
some + the+children + the+young + intense + the+intelligence
Some of the young children are intensely intelligent

3.3.1.3 Multi-token adjective modifying a single token NP

When the adjective is the head of a multi-token ADJP, that ADJP is adjoined to the head NP (even if the head NP is a single token noun).

Coordinated adjectives:

All coordinated adjectives must be multi-word, and therefore must be adjoined to the NP they modify (even if the NP is a single word noun).

```
(S (VP will+he/it+continue/last_ |sayasotamir~u | سَيَسْتَمِرُّ
  (NP-SBJ *)
  (PP-CLR in |fiy | في
    (NP (NP the+way/method/procedure |Aln~ahoji | النُّهْجُ)
      (ADJP the+extremist/fanatic/strict/zealot |
        Almuta$ad~idi | المُتَشَدِّدُ
          and |wa- | وَ-
            the+fierce/violent/forceful |-AlEaniyfi | -الْعَنِيفُ))))))
سيستمر في النهج المتشدد والعنيف
sa+yasotamir~+u + fiy + Al+nahoj+i + Al+muta$ad~id+i + wa+Al+Eaniyf+i
will+continue + in + the+procedure + the+fanatic + and+the+violent
He will continue in the fanatic and violent procedure
```

Adjective with NP complement (false idafa):

All adjectives in false idafa must be multi-word, and therefore must be adjoined to the NP they modify, if they are in a modifying position.

```
(S (VP jA'+a
  (NP-SBJ (NP Al+rajul+u)
    (ADJP Al+Tawiyl+u
      (NP Al+qAmap+i))))))
جاء الرجل الطويل القامة
jA'+a + Al+rajul+u + Al+Tawiyl+u + Al+qAmap+i
came + the+man + the+tall + the+stature
The man with the tall stature came
```

```
(NP (NP (NP (ADJP musician/musical_performer |-EAzifi | عازف-
  (NP the+lute/stick |AlEuwdi | العُودُ
    and |wa- | وَ-
      the+violin |-AlkamAni | -الْكَمانُ)))
  (NP Simon |siymuwn | سيمُونُ
    Shaheen/Chahine |$Ahiyn | شاهين))
  (ADJP (ADJP the+New_Yorker |Aln~iyuwyuwrokiy~i | النُّيُورُوكِيِّ
    (NP the+residency/setting_up |Al<iqAmap+i | الإقامة))
    and |wa | وَ
      (ADJP the+Palestinian |AlfilasoTiyaniy~i | الفِلَسْطِينِيِّ
        (NP the+membership/affiliation/commitment |
          AlAinotimA'i | الانْتِمَاءُ))))))
عازف العود والكمّان سيمون شاهين النيويوركي الإقامة والفلسطيني الانتماء
EAzif+u + Al+Euwd+i + wa+Al+kamAn+i + siymuwn + $Ahiyn + Al+nuyuwyuwrokiy~+i +
  Al+<iqAmap+i + wa+Al+filasoTiyaniy~+i + Al+<inotima'+i
musical performer + the+lute + and+the+violin + Simon + Shaheen + the+New Yorker +
  the+residency + and+the+Palestinian + commitment
Simon Shaheen, the lute and violin musical performer who is a New Yorker by residency and
  Palestinian by commitment
```

Adjective with PP complement:

(S (NP-SBJ it/he |huwa | هُوَ)
 (PP-PRD from |min | مِنْ)
 (NP (NP the+marks/signs/points |AlEalAmAti | العَلَامَاتِ
 the+shining/bright |AlmuDiy}api | المُضِيئَةِ)
 (ADJP the+rare/unusual/infrequent |Aln~Adirapi | النَادِرَةِ)
 (PP in |fiy | فِي)
 (NP the+life |AlHayApi | الْحَيَاةِ
 the+cultural/intellectual |Alv~aqAfiy~api | الثَّقَافِيَّةِ
 the+Lebanese |All~ubonAniy~api | اللُّبْنَانِيَّةِ))))))
 هو من العلامات المضيئة في الحيات الثقافية اللبنانية
 huwa + min + Al+EalAmAt+i + Al+muDiy}ap+i + fiy + Al+HayAt+i + Al+vaqAfiy~ap+i +
 Al+lubnAniy~ap+i
 he + from + the+signs + the+shining + in + the+life + the+cultural + the+Lebanese
He is from the shining signs in the Lebanese cultural life

Adjective with internal modification:

(S (VP kAnat كانت
 (NP-SBJ *)
 (NP-PRD (NP fatApF فتاة)
 (ADJP *akiy~pF ذكيّة
 jidAF جداً)))))
 كانت فتاة ذكيّة جداً
 kAnat + fatAp+F + *akiy~ap+F + jid~+AF
 was + girl + smart + very
She was a very smart girl

3.3.2 Nominal modification

The NP adjunct agrees with the head NP in definiteness and case.

3.3.2.1 "Verbal corroboration"

This section is about the construction discussed in Wright, 2005, p.282 (tawkiyd lafoDiy~emphasis concrete).

The NP adjunct in this construction should an emphatic repetition of the NP head. It can be a simple repetition of the same word, or it can be a repetition of the word along with different or additional modification. Either way, the repetition should be adjoined to the NP head as apposition.

(S (VP kAnat كانت
 (NP-SBJ tiloka تِلْكَ)
 (NP-PRD (NP mufAja>apF مُفَاجَأَةٌ
 rahiybapF رَهِيْبَةٌ)
 (NP (NP mufAja>apF مُفَاجَأَةٌ)
 (SBAR (WHNP-1 0)
 (S (VP haz~at هَزَّتْ
 (NP-SBJ-1 *T*)
 (NP-OBJ AljamiyEa الْجَمِيْعُ))))))))
 كانت تلك مفاجأة رهيبة مفاجأة هزّت الجميع
 kAnat + tiloka + mufAja'ap+F + rahiybap+F + mufAja'ap+F + haz~at + Al+jamiyE+a
 was + that + surprise + terrible + surprise + agitate + everybody
That was a terrible surprise. It shook everybody

If a bound/clitic pronoun is being repeated for emphasis, the “repetition” will be the unbound form of the pronoun (rather than a strict repeating of the clitic). The unbound form of the pronoun is annotated as an NP adjoined to the clitic pronoun. (See section ????? for more about pronouns.)

(S (VP >adoEuw أَدْعُو
 (NP-SBJ *)
 (NP-OBJ (NP ka ك)
 (NP >anota أَنْتَ))))
 أَدْعُوكَ أَنْتَ
 >adoEuw+ka + >anota
 invite + you
I invite you, yourself

To emphasize a pronominal subject., the empty pro-drop subject is replaced by the full unbound form of the pronoun:

(S (VP jA'a جاء
 (NP-SBJ huwa هُوَ)
 (PP-MNR bi بِ
 (NP nafosi نَفْسِ
 (NP hi هِ)))))
 جاء هو بنفسه
 jA'+a + huwa + bi+nafos+i+hi
 arrived + by+himself
He arrived by himself

NB: need an example here of how this is different from restarts and EDITEDs in speech (can be same as in BN guidelines)

3.3.2.2 Corroboration in meaning

This section is about the construction discussed in Wright, 2005, p.282 (tawkiyd maEonawiy~/emphasis abstract).

A number of nouns function as emphasizers of other nouns. This is similar to the repetition emphasis above, except that the head noun is not repeated. The “repetition” is in meaning – the *same* or the *totality/all of it* – rather than a strict repetition of the same word. These emphasizing nouns are in apposition to the head noun, and they usually have a pronominal complement that is co-referential with the head noun.

The emphasizing nouns include the universal quantifiers (kul~, kilaA (-ayo), jamiyE), nafos, *At, Eayon, >ajomaE, >ajomaEuna (-iyna).

Universal quantifier (kul~, kilaA (-ayo), jamiyE) (all with POS = NOUN_QUANT), followed by a co-referential pronoun:

```
(S (NP-SBJ (NP the+people | -AlnAsu | الناس-)
      (NP every/all/each_one | kul~u- | كُلُّ-
        (NP (their | -hum | هُمْ)))
      (NP-PRD ears | |*AnN | آذَانُ
        attentive/listening | SAgiyapN | صاغية))
```

الناس كلهم آذان صاغية

Al+nAs+u + kul~+u+hum + |*An+N + SAgiyap+N

the+people + all+their + ears + listening

All the people are listening

Same/self (nafos, *At, Eayon), followed by a co-referential pronoun:

```
(NP (NP the+work/action | -AlEamal+u | العمل-)
      (PP-TMP in | fiy | في
        (NP (NP the+time/moment/period_of_time | Alwaqoti | الوقت)
              (NP same/self/essence | *Ati- | ذات-
                (NP its/his | -hi | هـ)))
        (PP for/to | li- | ل-
          (NP stopping/ceasing | -waqofi | وَقْف-
            (NP the+confrontations/encounters | AlmuwAjahAti |
              المواجهات))))))
```

العمل في الوقت ذاته لوقف المواجهات

Al+Eamal+u + fiy + Al+waqot+i + *At+i+hi + li+waqof+i + Al+muwAjahAt+i

the+work + in + the+time + same+its + to+ceasing + the+confrontations

Working at the same time to cease the confrontations

```
(S (VP be_filtered/be_distilled+it/they/she | -takar~arat | تَكَرَّرَتْ-
      (NP-SBJ (NP the+expressions | Alt~aEAbiyru | التعبيرُ)
        (NP same/self | nafosu- | نَفْسُ-
          (NP its/their/her | -hA | هـ))))))
```

تكررت التعبير نفسها

takar~arat + Al+taEabiyr+u + nafos+u+hA

repeated + the+expressions + same+them

The expressions themselves were repeated

This modification can be made by a single word noun without a pronoun complement: >ajomaE, >ajomaEuna (-iyna). This noun should be adjoined to the head NP.

(S (VP indeed |<in~a | إِنَّ
 (NP-SBJ (NP the+people/nation |Al\$~aEoba | الشَّعْبُ
 the+Libyan |All~iybiy~a | اللَّيْبِيُّ
 and |wa- | وَ-
 (NP (NP the+world |-AlEAlama | العَالَمَ-
 the+Arab/Arabic |AlEArabiy~a | الْعَرَبِيُّ))
 (NP entire/whole |>ajomaEa | أَجْمَعُ)))
 (ADJP-PRD unhappy/displeased/dissatisfied |musotA'N | مُسْتَأْ))
 إِنَّ الشَّعْبَ اللَّيْبِيَّ وَالْعَالَمَ الْعَرَبِيَّ أَجْمَعُ مُسْتَأْ
 <in~a + Al+\$aEob+a + Al+liybiy~+a + wa+Al+EAlam+a + Al+EArabiy~+a + >ajomaE+a +
 musotA'N
 indeed + the+people + the+Libyan + and+the+world + the+Arab + entire + displeased
The Libyan people and the entire Arab world are displeased

A distinction should be made between modification of the noun and modification of the verb.

In this first example, nafosu+hu is an NP adjunct to the noun Alrajulu:

(S (VP jA'a جاء
 (NP-SBJ (NP Alrajulu الرَّجُلُ)
 (NP nafosu نَفْسُ
 (NP hu هُ))))
 جاء الرجل بنفسه
 jA'+a + Al+rajul+u + nafos+a+hu
 came + the+man + same+his
The same man came

In this second example, however, the PP bi+nafosihi is modifying the verb jA'a (Hal), so it is a child of the VP, not adjoined to the NP:

(S (VP jA'a جاء
 (NP-SBJ Alrajulu الرَّجُلُ)
 (PP-MNR bi ب
 (NP nafosi نَفْسِ
 (NP hi هِ)))))
 جاء الرجل بنفسه
 jA'+a + Al+rajul+u + bi+nafos+i+hi
 came + the+man + by+himself
The man came by himself

3.3.2.3 Restrictive apposition (badal permutation)

This section is about the construction discussed as permutative, in Wright, 2005, p.284.

The NP adjunct in restrictive apposition "limits or clarifies the foregoing [NP] in some crucial way" (<http://en.wikipedia.org/wiki/Apposition>, as accessed 7-30-2007).

3.3.2.3.1 Names in apposition

Names in apposition are an exception to the 'all adjuncts on same level' rule. The whole NP prior to the appositive name is annotated as usual, but the appositive name is an adjunct to that full NP, which is to say, there is an extra NP level: for example, (NP (NP (NP head noun) (XP any adjunct)) (NP appositive name))

```
(NP (NP (NP head noun)
        (XP any adjunct))
    (NP appositive name))
```

```
(S-ADV (VP representing |mumav~ilAF | ممثلاً
        (NP-SBJ *)
        (NP-OBJ (NP (NP president/head/chairman |ra}iysa | رئيس
                    (NP the+ministers |AlwuzarA'i | الوزراء
                    (ADJP the+libanease|AllubonAniy~i | اللبناني
                    (NP Rafiq/Rafeek |rafiyq | رفيق
                    the+Hariri |AlHariyriy~ | الحريري))))
        (t-('t-tt''))
        ممثلاً رئيس الوزراء اللبناني رفيق الحريري
```

mumav~il+AF + ra}iys+a + Al+wuzarA'+i + Al+lubnaniy~+i + rafiya + Al+Hariyriy~
 representing + chairman + the+ministers + the+Lebanese + Rafeek + al-Hariri
Representing the Lebanese prime minister Rafeek al-Hariri

```
(NP (NP (NP ra}isu رئيس
        (NP AlwuzarA'i | الوزراء ) )
        (SBAR (WHNP-1 Al~*iy | الذي
        (S (VP mAta مات
            (NP-SBJ-1 *T*)
            (PP fiy في
            (NP ZuruwfK ظروف
            gAmiDapK غامضة )))))
        (NP rafiya رفيق
        AlHariyriy | الحريري ) ) )
```

ra}is+u + Al+wuzarA'+i + Al~*iy + mAt+a + fiy + Zuruwf+K + gAmiDap+K + rafiya +
 Al+Hariyriy
 chairman + the+ministers + who + died + in + circumstances + ambiguous + Rafeek + al-Hariri

The prime minister who died in ambiguous circumstances, Rafeek al-Hariri

```
(S (VP zAra زار
    (NP-OBJ Almadiynapa | المدينة )
    (NP-SBJ (NP (NP AlmuTribu | المطرب
                AleirAqiy~u | العراقي
                (ADJP Alma$huwru | المشهور
                jid~AF | جداً ) )
            (NP kAZim | كاظم
            AlsAhir | الساهر ) ) ) )
```

zAr+a + Al+madiynap+a + Al+muTorib+u + Al+EirAqiy~+u + Al+ma\$ohur+u + jid~+AF +
 kAZim + Al-sAhir
 زار المدينة المطرب العراقي المشهور جداً كاظم الساهر

visit + the+city + the+singer + the+Iraqi + the+famous + very + Kazim + Al-Sahir
The very famous Iraqi singer, Kazim al-Sahir, visited the city

3.3.2.3.2 Demonstratives

Demonstrative pronouns are annotated as the head of an adjunction structure, modified by an appositive NP.

```
(NP (NP h`*ihi:: ذه`ه::this/these )
  (NP Al+tagoyiyr+At+i:: التغيرات::the+change/modification/replacement))

(S (VP it/they/she+be_intertwined/be_interwoven |tata$Abaku | تتشابك
  (NP-SBJ mentioning/citing/remembering |*Akiratu- | ذاكِرَتْ-
    (NP its/their/her |-hA | -ها))
  (PP in |fiy | في)
    (NP (NP this |h`*A | ه`ذا)
      (NP the+space/cosmos |AlfaDA'i | الفضاء
        the+civilized/cultural |AlHaDAriy~i | الحضاري))))
  هذه التغيرات تتشابك ذاكرتها في هذا الفضاء الحضاري
h`*ihi + Al+tagoyiyr+At+i + tata$Abak+u + *Akirat+u+hA + fiy + h`*A + Al+faDA'+i +
Al+HaDAriy~+i
these + the+changes + be interwoven + remembering+its + in + this + the+space +
the+cultural
Remembering these changes is interwoven with this cultural space
```

This analysis follows the categorization of demonstratives as badal/restrictive apposition in Arabic traditional grammar, and the interpretation of the demonstratives in Arabic. Arabic pre-nominal demonstratives are always interpreted as full independent noun phrases, unlike in English where they are often interpreted as determiners.

NB: The ATB I guidelines treated pre-nominal demonstratives as determiners, and as such flat in an NP with the following noun. The annotation of demonstratives as the head NP with an adjoined NP appositive is one of the new policies in these ATB II guidelines.

Post-nominal demonstrative pronouns are adjoined to the head noun phrase.

```
(PP to/towards |-<ilaY | إلى-
  (NP (NP day |yawomi- | يَوْم-
    (NP our |-nA | نَا))
    (NP this |h`*A | ه`ذا)))
<ilaY + yawom+i+nA + h`*A
to + day+our + this
till the present time
```

الى يومنا هذا

3.3.2.3.3 NP modified by quantificational NP (non-universal)

Noun phrases that are modified by a post-nominal non-universal quantificational noun phrase are annotated as adjunction structures. The post-nominal quantificational NP is adjoined to the head NP.

```
(S (VP >akalotu أَكَلْتُ
  (NP-SBJ * )
  (NP-OBJ (NP AldajAjapa الدجاجة
    (NP niSofa نصف
      (NP hA ها ) ) ) ) )
```

أكلت الدجاجة نصفها

>akalot+u + Al+dajAjap+a + niSofa+hA
ate + the+chicken + half +its
I ate half of the chicken

```
(S (VP jA'a جاء
  (NP-SBJ (NP AlTalabapu الطالبة
    (NP baEoDu بعض
      (NP hum هُم ) ) ) ) )
```

جاء الطالبة بعضهم

jA'+a + Al+Talabap+u + baEoD+u+hum
came + the+students + some+their
Some of the students came

3.3.2.3.4 NP modified by numerical NP

A noun phrase modified by a post-nominal numerical NP (including a numerical noun that is itself modified by a prepositional phrase) is annotated as an adjunction structure. The post-modifying numerical NP is adjoined to the head NP.

```
(S (VP qara>otu قَرَأْتُ
  (NP-SBJ * )
  (NP-OBJ (NP AlkitAba الكتاب
    (NP (NP Ei$rina عشرين
      (NP SafoHapF صفحة ) )
      (PP min من
        (NP hu هُ ) ) ) ) ) )
```

قرأت الكتاب عشرين صفحة منه

qara>ot+u + Al+kitAb+a + Ei\$rina + SafoHap+F + min+hu
read + the+book + twenty + page from+it
I read twenty pages of the book

3.3.2.4 SBAR-NOM or S-NOM modified by an adjunct

An SBAR-NOM or S-NOM that is modified by an adjunct is adjoined to that adjunct with an NP adjunction level.

(NP jamiyEa جَمِيعَ
 (NP (SBAR-NOM (WHNP-1 mA مَا
 (S (NP-PRD Hawla حَوْلَ
 (NP hA هَا))
 (NP-SBJ-1 *T*)))
 (PP min مِن
 (NP Albuyuwti البُيُوتِ))))

جَمِيعَ مَا حَوْلَهَا مِنَ الْبُيُوتِ

jamiyEa mA HawlahA min Albuyuwti
 ‘all that (is) around+her of the+houses’

3.4 Flat

3.4.1 Titles

Single word personal titles are annotated flat along with the proper name they belong to, when they are interpreted as part of that name.

(NP Al+Eamiyd+i::الْعَمِيد::the+brigadier_general+[def.gen.]
 riyuwn::رَيْوْن::Raymond
 EAzAr::عَازَر::Azar)

العميد ريمون عازار

Al+Eamiyd+i + riyuwn + EAzAr
 the+dean + Raymond + Azar
Dean Raymond Azar

Preliminary list of titles:

NB: need transliterations and glosses for all of these lists

Social titles

السيد
 السيدة
 الأنسة

Academic titles

الدكتور
 الأستاذ

Religious titles

الشيخ
 الإمام
 الفقيه

Military titles

العرِف

الكولونيل
الجنرال
اللواء
المقدم
القبطان
العميد
المستشار
الفريق

Political titles

الرئيس
الوزير
الملك
الأمير
النائب
الرفيق
الأخ
القائد
الزعيم

Multi-word personal titles are shown as apposition. (This is similar to the policy of single word noun and adjective combinations being flat NPs.)

(NP (NP جلالة
(NP الملك))
(NP (عبد الله

جلالة الملك عبد الله

Highness the king Abdallah

His Highness the King Abdallah

NB: needs the transliteration, and add to the tree also

3.4.2 Proper nouns

3.4.2.1 Case ending/part-of-speech policy for proper nouns

Though proper nouns are traditionally used with their case-endings (as in the example: Daraba muHam~aduN Ealiy~AF, Beat Muhammad [SBJ] Ali [OBJ]), it was decided that Arab personal names would not be assigned case endings in the ATB POS annotation.

So most, if not all ATB personal names are without case endings. However, an exception was made for Arab place names, which do have case endings in our annotation. Note also that according to Tim Buckwalter, we encountered no instances of personal names used in the indefinite since in the ATB as this is highly improbable (e.g. "50 Mohamed's...").

3.4.2.2 Treebanking proper nouns

Proper names are annotated as flat NPs.

```
(NP zayon::زَيْن::Zein
  Al+EAbdyn::العابدين::the+NOT_IN_LEXICON
  bin::بن::bin
  Ealiy~::عَلِيَّ::Ali )

zayn + Al-EAbidiyn + bin + Ali
Zain + al-Abidin + bin + Ali
Zain al-Abidiyn bin Ali
```

زين العابدين بن علي

It is, however, important to distinguish between the words used to reference a proper name and the same words that may be used as a common noun phrase. When the proper noun or the noun of the named entity is complex, the meaning of the phrase should be carefully considered, to determine whether it is a proper name (and therefore annotated flat) or referring to common noun (and therefore annotated with structure).

South/southern Africa جنوب إفريقيا is an example of a noun phrase that may be ambiguous in this respect.

If it means the name of the country South Africa, it should be treated as proper noun (flat):

```
(S (NP-SBJ mAndiylA مانديلا )
  (ADJP-PRD >a$oharu أَشْهُرُ
    (NP ru&asA'i رؤساء
      (NP januwbi جَنُوبِ
        <ifriyqiYA إفريقيا ) ) ) )
```

مانديلا أشهر رؤساء جنوب أفريقيا

mAndiylA + >a\$ohar+u + ru&asA'i + januwb+i + <ifriyqiYA
Mandela + most famous + presidents + South + Africa
Mandela is the most famous president in South Africa

If it refers to a geographical part of Africa (the southern part of Africa), however, it should be represented as complementation structure (the Africa NP is a complement of the south NP):

```
(S (VP zurtu زرت
  (NP-SBJ * )
  (NP-OBJ kul~a كُلُّ
    (NP buldAni بلدان
      (NP januwbi جَنُوبِ
        (NP <iforiqiYA إفريقيا ) ) ) ) ) )
```

زرت كل بلدان جنوب إفريقيا

Zurot+u + kul~+a + bulodAn+i + januwb+i + <iforiqiYA
Visited + all + countries + south + Africa
I visited all the countries of South Africa

3.4.3 Adjective single word

A single word adjective modifying a single word noun makes a flat NP:

```
(NP Al+mudiyar+a::المدير::the+director/manager/chief+[def.acc.]
  l+EAm~+a::العام::the+general/common/public+[def.acc.] )
```

المدير العام

Al+mudiyar+a + Al+EAm~+a
the+director + the+general
The general director

3.4.4 Compound numbers

Compound, multi-word numbers should be annotated flat, as if they were a single token.

```
(NP mA{apN مائة
  wa و
  wAHidN واحد
  wa و
  Ei$oruwna عشرون
  (NP kitAbAF كتاباً ) )
```

مائة واحد وعشرون كتاباً

mA{ap+N + wa+wAHid+N + wa+Ei\$oruwna + kitAb+AF
hundred + and+one + and+twenty + book
One hundred and twenty one books

We should point to the fact that there is complement structure whenever mA'ap, >alof,milyuwn is coming as complement of compound number

It is not flat for compound numbers having hundreds, thousands, millions as a second element in a numeral, i.e. complement structure.

3.4.5 Coordination between single words is flat

Single words of any type that are coordinated to single words of the same type are annotated as flat. See section ????? for more on coordination in general.

A single word noun coordinated with another single word noun will be a flat NP:

```
(PP fiy::في::in
  (NP $u&uwn+i::شؤون::affairs/matters+[def.gen.]
    (NP Al+Harob+i::الحرب::the+war/warfare+[def.gen.]
      wa-::و-::and
      -Al+salAm+i::-السلام::the+peace+[def.gen.] )))
```

في شؤون الحرب والسلام

fiy + \$u&uwn+i + Al+Harob+i + wa+Al+salAm+i
in + affairs + the+war + and+the+peace
In war and peace matters

3.5 Pronoun constructions

3.5.1 Forms and cases of pronouns in Arabic

See also section ????? on pronouns in the POS/morphological guidelines.

Referential overt pronouns are not coindexed with other full noun phrases in the Treebank, as this form of discourse coreference is not shown in the Treebank in general. Pronouns are coindexed only if they control an empty category. See section ????? on empty categories.

Forms of the bound and unbound pronouns vary according the syntactic position occupied by the pronoun. In nominative position (subject), the pronouns are

NB: need Arabic script and glosses for this list

>anA
naHonu
>anota
>anoti
>anotumA
>anotum
>anotun~a
huwa
hiya
humA (mas)
humA (fem)
hum
hun~a

(S (NP-SBJ hiya::هي::it/they/she)
(ADJP-PRD maEoruwf+ap+N::مَعْرُوفَةٌ::known/well-known+[fem.sg.]+[indef.nom.])
(NP-TMP HAliy~+AF::حَالِيًا::presently/currently+[acc.indef.]))

هي معروفة حاليًا

hiya + maEoruwfap+N + HAliy~+AF

it + known + now

It is known now

In accusative or genitive position (object, complement of preposition or noun), the bound/cliticized forms of the pronouns are

NB: need Arabic script and glosses for this list

-iy/niy(for the verb)

-nA
- ka
-ki
-kumA
-kum
-kun~a

-hu
-hA
-humA
-hum
-hun~a

```
(S (PP-PRD fiy::في::in
  (NP (NP kul~+i::كُل::every/all/each_one+[def.gen.]
    (NP wAHid+K::واحد::one+[indef.gen.] ))
    (PP min-::من::from
      (NP -hA::-ها::it/them/her ))))
  (NP-SBJ majomuWE+ap+N::مجموعة::collection/group/bloc+[fem.sg.]+[indef.nom.] ))
```

في كل واحد منها مجموعة

fiy + kul~+i + wAHid+K + min+hA + majomuWEap+N
in + every + one + from+their + group
There is a group in each one of them

When these accusative or genitive bound/clitic pronouns are separated in the text from the word (verb, noun, particles) that they are normally cliticized to, they are suffixed to the dummy element -يَا <iy~A instead. The dummy element and the pronoun are annotated together as a flat NP.

This construction occurs when

- The pronoun object is preposed before its verb.

```
(S (NP-TPC-1 <iy~A يَا
  ka كَ )
  (VP >aEoniy أَغْنِي
    (NP-SBJ * )
    (NP-OBJ-1 *T*) ) )
```

إِيَّاكَ أَغْنِي

<iy~A+ka + >aEoniy
NO GLOSS+you + mean
I mean you, yourself

- Both accusative arguments of a ditransitive verb (-OBJ and -DTV) are bound/clitic pronouns. Because the verb will be too long with two bound pronouns, the second pronoun is detached of the verb and suffixed to <iy~A.

```
(S (VP >aEoTA أَعْطَا
  (NP-SBJ * )
  (NP-DTV hu هُوَ )
  (NP-OBJ <iy~A يَا
    hA هَا ) ) )
```

أَعْطَاهُ إِيَّاهَا

>aEoTA+hu + <iy~A+Ha
gave+him + NO GLOSS +it
He gave it to him

- The active participle has a verbal reading – to show that the bound pronoun/clitic object is an accusative argument of the verb, and not to confuse it with the bound pronoun/clitic complement which is genitive when the participle has the adjectival or nominal reading.

- the verbal reading with <iy~AhA

```
(S (VP jA'a جاء
    (NP-SBJ-1 * )
    (S-MNR (VP HAmilAF حاملا
        (NP-SBJ-1 * )
        (NP-OBJ <iy~A إِيَّا
            hA ها ) ) ) ) )
```

جاء حاملا إِيَّاها

jA'+a + HAmil+AF + <iy~A+hA
 came + carrying+it
He came carrying it

- the nominal reading with -hA

```
(S (VP $akartu شَكَرْتُ
    (NP-SBJ * )
    (NP-OBJ HAmila حَامِلًا
        (NP hA ها ) ) ) )
```

شكرت حاملها

\$akarot+u + HAmil+a+hA
 thank + holding+it
I thanked the one who was holding it

- Following the preposition wa (waw AlmaEiy~ap).

```
(S (VP xarajotu خَرَجْتُ
    (NP-SBJ * )
    (PP wa وَ
        (NP <iy~A إِيَّا
            hu هُوَ ) ) ) ) )
```

خرجت وإِيَّاه

xarajot+u + wa+<iy~A+hu
 went out + and+him
I went out with him

NB: The dummy element <iy~A- was frequently annotated as PRT in previous versions of the ATB. This is not correct – it should be annotated as a flat NP with the bound pronoun, as above.

NB: Mohamed, Sondos, Dalila resolve the flat question here

3.5.2 Clitic pronouns (bound pronouns)

The prevalence of cliticization in Arabic sentences of determiners, prepositions, conjunctions, and pronouns led to a necessary difference in tokenization between the morphological/POS files and the treebanking files. Clitics that play a role in the syntactic structure are split off into separate tokens (e.g., object pronouns cliticized to verbs, subject pronouns cliticized to complementizers, cliticized prepositions, etc.). Clitics that do not affect the structure are not separated (e.g., determiners). Since the word boundaries necessary to separate the clitics are taken from the POS tags, and since it is not possible to show the syntactic structure unless the clitics are separated, correct POS tagging is very important in order to be able to properly separate clitics prior to treebanking.

Clitic/bound pronouns appear in the following contexts:

- **Complement of noun (possessive pronoun)**

```
(S (VP -yu+qad~im+u::-يُقَدِّمُ::he/it+offer/present/introduce+[ind.]
  (NP-SBJ Al+nA$ir+u::الناشر::the+publisher+[def.nom.] )
  (NP-OBJ kitAb+a-::كِتَاب::book+[def.acc.]
    (NP -hu::-هُ::its/his )))
```

يقدم الناشر كتابه

yuqad~im+u + Al+nA\$ir+u + kitAb+a+hu
 introduce + the+publisher + book+his
The publisher introduces his book

- **Complement of preposition**

```
(S (VP -na+taman~aY+[null]::-نَتَمَنَّى[نُلل]::we+desire/wish_for/hope+[ind.]
  (NP-SBJ (-NONE- *))
  (PP la-::-لِ-::to/for
    (NP -hum::-هُمْ::them_[masc.pl.]_[they_have] ))
  (NP-OBJ Haj~+AF::حَجًّا::pilgrimage/Haj+[acc.indef.]
    maboruwr+AF::مَبْرُورًا::blessed+[acc.indef.] )))
```

نتمنى لهم حجاً مبروراً

na+taman~aY+o + la+hum = Haj~+AF + maboruwr+AF
 wish + to+them + pilgrimage + blessed
We wish them a blessed pilgrimage

- **Object of verb**

```
(SBAR (WHNP-1 Al~atiy التي )
  (S (VP (PRT lam لم )
    yamas~ يَمَسُّ
    (NP-OBJ (NP hA ها )
      (NP-1 *T* ) )
    (NP-SBJ suw'N سُوء ) ) ) ) )
```

التي لم يمسها سوء

Al~atiy + lam + yamas~+hA + suw'+N
 which + did not + touch + evil
Which was not effected by evil

- **Subject following inna and its sisters**

```
(S (VP ان
  (NP-SBJ-1 هـ )
  (S (VP قاسى
    (NP-SBJ-1 *)
    (NP-OBJ الأمرين ) ) ) ) )
```

انه قاسى الأمرين

Indeed went he through the worst
He went through the worst

NB: need transliteration, also in the tree

- **Subject following a complementizer**

```
(S (VP أكد
  (NP-SBJ *)
  (SBAR أن
    (S (NP-SBJ هـ )
      (ADJP-PRD (قادم ) ) ) ) ) )
```

أكد أنه قادم.

Insisted that him coming
He confirmed that he's coming

NB: need transliteration, also in the tree

3.5.3 Resumptive pronouns

Resumptive pronouns are overt pronouns that appear in the position of a *T* trace. The trace can be coindexed to either a WH- word or to a topicalized constituent (usually NP-TPC).

The trace is adjoined to the overt resumptive pronoun:

```
(NP (NP ha)
  (NP-1 *T* ) )
```

Some examples of WH- traces with resumptive pronouns follow.

- **Subject in an equational sentence**

```
(SBAR-NOM (WHNP-2 -mA::ما::what )
  (S (NP-SBJ (NP huwa::هُوَ::it/he )
    (NP-2 *T*))
    (PP-PRD li-::لِ-::for/to
      (NP -maSolaH+ap+i::مَصْلَحَة::interest/advantage/
        favor+[fem.sg.]+[def.gen.]
        (NP Al+Tul~Ab+i::الطُّلَّابُ::the+students+[def.gen.] )))))
    ما هو لمصلحة الطلاب
```

mA + huwa + li+maSolaHap+i + Al+Tul~Ab+i
 what + it + for+favor + the+students
What is for the favor of the students

- **Object of a preposition**

```
(PP-LOC <ilaY::إِلَى::to/towards
  (NP (NP sijon+K::سجن::prison+[indef.gen.]
    <isokutolanodiy~+K::إِسْكُوتْلَانْدِيَّ::Scottish+[indef.gen.] )
    (SBAR (WHNP-2 *0*)
      (S (VP sa+yu+naf~i*+u::سَيُنْفِذُ::will+he/it+
        implement /accomplish+[ind.]
        (PP-LOC fiy-::فِي::in
          (NP (NP -hi::-هُوَ::it/him )
            (NP-2 *T*))
            (NP-SBJ Euquwb+at+a-::عُقُوبَتُ-::punishment/sanction
              +[fem.sg.]+[def.acc.]
              (NP -hu::-هُوَ::his/its ))))))))
    إلى سجن إسكتلندي سينفذ فيه عقوبته
```

<ilaY + sijon+K + <isokutlanodiy + sa+yunaf~i*+u + fiy+hi + Euquwb+at+a+hu
 to + prison + Scottish + will+implement + in+it + sanction+his
To a Scottish prison where he will implement his sanction

- **Object of a verb**

```
(SBAR (WHNP-7 mA::ما::what )
  (S (VP sa+ya+quwl+u-::سَيَقُولُ-::will+he/it+say+[ind.]
    (NP-OBJ (NP -hu::-هُوَ::it/him )
      (NP-7 *T*))
      (NP-SBJ Al+maboEuwv+u::الْمَبْعُوثُ::the+envoy/representative+[def.nom.]
        Al+EirAqiy~+u::العِرَاقِيُّ::the+Iraqi+[def.nom.] ))))
    ما سيقوله المبعوث العراقي
```

mA + sa+yaquwl+u+hu + Al+maboEuwv+u + Al+EirAqiy~+u
 What + will+say+it + the+envoy + the+Iraqi
What will the Iraqi envoy say

- **Complement of a noun**

```
(PP-CLR <ilaY::إلى::to/towards
  (NP (NP Haloq+ap+K::خَلْقَة::ring/circle/link+[fem.sg.]+[indef.gen.] )
    (SBAR (WHNP-1 *0*)
      (S (NP-SBJ EunowAn+u-::عُنْوَان-::address+[def.nom]
        (NP (NP -hA::-ها::its/their/her )
          (NP-1 *T*)) )
        (NP-TTL-PRD "
          (NP waDoE+u::وَضْعُ::situation/status/
            condition+[def.nom.]
            (NP Al+maro>+ap+i::الْمَرْأَة::the+woman+
              [fem.sg.]+[def.gen.] ) )
          (PP-LOC fiy::فِي::in
            (NP >afogAnisotAn+a::أَفْغَانِسْتَان::
              Afghanistan+[def.acc.]
              "))))))
            إلى حلقة عنوانها "وضع المرأة في أفغانستان"
```

<ilaY + Haloqap+K + EunowAn+u+hA + waDoE+u + Al+maro>ap+i + fiy +
 >afogAnisotAn+a
 to + episode + title+its + status + the+woman + in + Afghanistan
To an episode titled "The Status of Women in Afghanistan"

Some examples of traces of topicalized constituents with resumptive pronouns follow.

- **Object of a verb**

```
(S (NP-TPC-3 -nAdiy::نادي::club/association
  (NP Al+maEohad+i::المَعْهَد::the+institute+[def.gen.]
    Al+>anoTuwniy~+i::الْأَنْطُونِي::the+Antouni/Antoni+[def.gen.] ))
  (PP fiy::فِي::in
    (NP buTuwl+ap+i::بُطُولَة::championship/starring_role/heroism+
      [fem.sg.]+[def.gen.]
      (NP Al+jabal+i::الْجَبَل::the+mountain+[def.gen.] )))
  (VP ya+liy+[null]-::يَلِي[نُلل]-::he/it+follow/come_after+[ind.]
    (NP-OBJ (NP -hi::هـ::it/him )
      (NP-3 *T*))
    (NP-SBJ nAdiy::نادي::club/association
      (NP Al+$Anfoyl::الشانْفِيل::the+Champville )))
    نادي المَعْهَد الْأَنْطُونِي فِي بطولة الجبل يليه نادي الشنافيل
    -nAdiy Al+maEohad+i Al+>anoTuwniy~+i fiy + buTuwl+ap+i + Al+jabal+i + yaliy+o+hi +
      nAdiy + Al+$anAfoyl
    gloss gloss gloss in + championship + the+mountain + follow+it + club + the+Champville
    trans trans transIn the Mountain championship followed by the Champville club
```

- **Object of a preposition**

(S (PRT >am~A::أَمَّا::as_for/concerning)
 (NP-TPC-1 Al+Tul~Ab+u::الطُّلَّابُ::the+students+[def.nom.])
 (PRT fa-::فَـ::and/so)
 (VP (PRT -qad::قَدْ::[has/have])
 waqaf+a::وَقَفَ::stop/cease/stand+he/it_[verb]
 (NP-SBJ (NP Eadad+N::عَدَدٌ::number/quantity/issue+[indef.nom.])
 (PP min-::مِنْ::from
 (NP (NP -hum::هُم::them_[masc.pl.])
 (NP-1 *T*))))
 (NP-LOC Einoda::عِنْدَ::with/at
 (NP madoxal+i::مَدْخَلُ::entrance/introduction+[def.gen.]
 (NP Al+jAmiE+ap+i::الْجَامِعَةُ::the+university+[fem.sg.]+[def.gen.]))))))
 أمَّا الطلاب فقد وقف عدد منهم عند مدخل الجامعة
 <am~A + Al+Tul~Ab+u + fa+qad + waqaf+a + Eadad+N + Einoda + madoxal+i +
 Al+jAmiEap+i
 concerning + the+students + so+have + quantity + from + them + at + entrance +
 the+university
Concerning the students, some of them have stood at the entrance of the university

- **Complement of a noun**

(S (NP-TPC-1 EqAr+a::عَقَارٌ::real_estate/immovable_property+[def.acc.]
 (NP binAy+ap+i::بِنَايَةٌ::building/structure+[fem.sg.]+[def.gen.]
 (NP Al+AwtwmAtyk::الْأُتُومَاتِيكُ::the+Automatic)))
 (VP (PRT qad::قَدْ::[has/have])
 jaraY+[null]::جَرَى[نُحْل]::occur/happen/take_place/flow+he/it_[verb]
 (NP-SBJ {isotibodAl+u-::{سُتَبَدِّلُ}::substitution/exchange/
 replacement+[def.nom]
 (NP (NP -hu::هُـ::its/his)
 (NP-1 *T*)))))))
 عقار بناية الاوتوماتيك قد جرى استبداله
 EqAr+a + binAyap+i + Al+>uwtuwmAtiyk + qad + jaraY + <isotibdAl+u+hu
 real estate + building + has + happen + substitute+it
The real estate of the Automatic building was substituted

- **When the resumptive pronoun has an appositive**

When the resumptive pronoun has an appositive adjoined to it, the *T* trace is adjoined first, to the lower NP, and the appositive NP is adjoined to the higher NP.

3.5.4 Separating pronouns/ Damiyr Al-faSol

The function of the separating pronouns/Damiyr Al-faSol is to separate the subject from the predicate in equational sentences when both of them are **definite**. These pronouns serve to distinguish the subject-predicate relationship from the apposition/adjunction relationship where there is also an agreement in definiteness.

The Damiyr Al-faSol is adjoined to the subject as an appositive:

```
(S (NP-SBJ (NP الهدف Al+hadaF+u the goal)
            (NP هو huwa it))
  (NP-PRD الحرية Al+Hur~iy~ap+u freedom))
```

الهدف هو الحرية

Al+hadaF+u + huwa + Al+Hur~iy~ap+u
 the+goal + it is + the+freedom
The goal is freedom

With demonstrative subjects, the separating pronoun is almost always used, as these sentences would be indistinguishable from plain NPs without the Damiyr Al-faSol.

```
(S (NP-SBJ (NP هذه h`*ihi this)
            (NP هي hiya it))
  (NP-PRD الحقيقة AlHaqiyqapu truth))
```

هذه هي الحقيقة

h`*ihi + hiya + Al+Haqiyqapu+u
 this + it is + the+truth
This is the truth

When the predicate of the equational sentence is a free relative with ala*y (without mA), the separating pronoun is also necessary to distinguish the equational sentence from a modified NP:

```
(S (NP-SBJ (NP الحرية AlHur~iy~pu freedom
              و wa and
              السلام Alsalamu peace)
  (NP هما humA they))
  (SBAR-NOM-PRD (WHNP-1 اللذان Al~a*Ani what)
    (S (VP نطلب naTolubu we want
      (NP-SBJ *)
      (NP-OBJ-1 *T*)))))
```

الحرية والسلام هما اللذان نطلب

Al+Hur~iy~ap+u + wa+Al+salAm+u + humA + Al~a*Ani + naTolub+u
 the+freedom + and+the+peace + they are + that + want
Freedom and peace are what we want

In Modern Standard Arabic, the use of Damiyr Al-faSol has been extended to include also separating heavy NP subjects from their predicates in equational sentences. In this use, the Damiyr Al-faSol can be used even when the predicate is indefinite.

(S (NP-SBJ (NP (NP Alhadafu الَهْدَفُ
 AlHaqiqiy~u الحَقِيقِي)
 (SBAR (WHNP-1 Al~*iy الّذي
 (S (VP >aradonA أَرَدْنَا
 (NP-SBJ *)
 (NP-OBJ (NP hu هُوَ)
 (NP-1 *T*))
 (PP-TMP mino*u مِنْذُ
 (NP zamanK زَمَنُ
 baEiydK بِعِيدٍ))))))
 (NP huwa هُوَ))
 (NP-PRD AlHur~iy~apu الحُرِّيَّةُ)))
 الهدف الحقيقي الذي أردناه منذ زمن بعيد هو الحرية
 Al+hadaf+u + Al+Haqiqiy~+u + Al~*iy + >aradonA+hu + mino*u + zaman+K baEiyd+K
 + huwa + Al+Hur~iy~apu
 the+goal + the+real + which + want+it + since + time + distant + it is + the+freedom
Freedom is the real goal we wanted for long time ago

This construction is also possible with kAna, and with inna and its sisters.

(S (VP kAna كَانَ
 (NP-SBJ (NP (NP AlsilAHu السلاح)
 (ADJP Almaqbuwlu المَقْبُولُ
 (PP-LOC fiy فِي
 (NP AlminoTaqap المِنطَقَة))))
 (NP huwa هُوَ))
 (NP-PRD silAHu سِلَاحُ
 (NP AlsiyAsapi السِّيَاسَة)))))
 كان السلاح المقبول في المنطقة هو سلاح السياسة
 kAn+a + Al+silAH+u + Al+maqobuwl+u + fiy + Al+minoTaqap+i + huwa + silAH+a +
 Al+siyAsap+i
 was + the+weapon + the+accepted + in + the+area + it is + weapon + the+politics
The accepted weapon in the area was the political one

(PP-PRP li ل
 (SBAR >an~a أَنَّ
 (S (NP-SBJ (NP Alhadafa الَهْدَفُ)
 (NP huwa هُوَ))
 (NP-PRD xidomapu خِدْمَة
 (NP AlmuHAMiyna المُحَامِيْنَ)))))
 لأنّ الهدف هو خدمة المحامين
 li>an~a + Al+hadaf+a + huwa + xidomapu+u + Al+muHAMiyna
 because + the+target + it is + service + the+lawyers
Because the target is to serve the lawyers

NB: This is a change of policy – in previous versions of the ATB, the Damiyr Al-faSol was treated as a resumptive pronoun.

However, if either or both NPs is **not definite** and there is a pronoun, that pronoun is not a separating pronoun but a resumptive pronoun resulting from topicalization.


```
(SBARQ (WHNP-1 mA ما
  (S (NP-SBJ (NP hiy~a هي)
    (NP-1 *T*))
    (NP-PRD SalAHiyap+u صلاحيات
      (NP hu ه)))
  ?)
```

ما هي صلاحياته ؟

mA hiy~a SalAHiyap+u
what she powers his/its
What are his powers?

```
(S (PRT أما >am~A as for)
  (NP-TPC-1 النظام الحالي Al+niDAm+u Al+HAliy~+u the current system)
  (PRT ف fa so)
  (NP-SBJ (NP هو huw~a he/it)
    (NP-1 *T*))
  (ADJP-PRD معادي muEAdiy against
    (PP ل li to
      (NP امريكا >amAriyKA America))))
```

أما النظام الحالي فهو معادي لأمريكا

>am~A Al+niDAm+u Al+HAliy~+u fa huw~a muEAdiy li >amAriyKA
As for the system the current so it against to America
As for the current regime, it is against America

3.5.5 Anticipatory pronouns/ ضمير الشأن / Damiyr Al\$a>on

These pronouns are called “anticipatory” because they anticipate at the beginning of the sentence, the “story” that will come in the sentence. They most commonly follow >in~a or >an~a (and rarely other sisters of >in~a: ka>an~a, laEal~a, lakin~a).

The Arabic Treebank annotates these anticipatory pronouns as topics – that is, they are marked as NP-TPC, but they do not have a co-indexed trace.

```
(SBAR أن >an~a-
  (S (NP-TPC له -hu)
    (VP سيلقي sayuloqiy
      (NP-SBJ الرئيس Alra{iysu)
      (NP-OBJ خطابا xiTAbAF)
      (PP ب -bi
        (NP المناسبة AlmunAsabapi))))))
```

أنه سيلقي الرئيس خطابا بالمناسبة

>an~a+hu + sa+yaloqiy + Al+ra{iys+u + xiTAb+AF + bi+Al+munAsabap+i
that + will+deliver + the+president + oration + in+the+occasion
That the president will deliver an oration for the occasion

Originally, especially in classic Arabic, this pronoun could be used without the >in~a or >an~a:

قل هو الله أحد

Qul huwa All~ahu >aHadN
say it God (is) unique

These will also be annotated as topics in the treebank, if an example like this appears:

```
(S (VP قل say
  (NP-SBJ *)
  (S (NP-TPC هو huwa)
    (NP-SBJ الله All~ahu)
    (NP-PRD أحد >aHadN) ) ) )
```

قل هو الله أحد

Qul + huwa + All~ah+u + >aHad+N
 say + God + unique
Say: He is Allah, the One and Only

In Modern Standard Arabic (MSA), this pronoun is used only after "the sentence modifiers <in~a إن, >an~a أن" (Badawi 2005, 337). The value of this pronoun is to anticipate the sentence that comes after <in~a or >an~a, and this sentence can be

- Most of the time, a verbal sentence with a regular verb or a weak verb (kAna and its sisters)

Ealima >an~ahu lan tatmak~ana Al>um~a min <irDAEi mawoluwdihA

<in~a hu laysa lildumuwei vamanN

- An equational sentence with an inverted order (prepositional phrase predicate)

<in~a hu min AlmusotaHili <inojAzu ha*A Alma\$oruwEi

NB: This is a change of policy – in previous versions of the ATB, the Damiyr Al\$a>on was treated as the subject of the clause.

3.5.5.1 Not to be confused with referential pronouns following in >an~a or >in~a

It is important not to confuse the anticipatory pronouns/Damiyr Al\$a>on with truly referential pronouns, which can also follow >an~a or >in~a.

```
(S (VP <in~a إِنَّ )
  (NP-SBJ hu هُوَ )
  (NP-PRD >axuw أَخُو
    (NP ka كَ ) ) ) )
```

إِنَّهُ أَخُوكَ

<in~a+hu + >axuw+ka
 indeed+he + brother+yours
Indeed he is your brother

How can the referential –hu be distinguished from the anticipatory –hu?

Tests:

1. If there is an overt subject of the same subordinate clause in addition to the –hu, the –hu must be anticipatory and cannot be referential.

(S (VP إِنَّ <in~a
 (NP-TPC هـ hu)
 (PP-PRD من min
 (NP الممكن Almumokin)
 (NP-SBJ تصديق taSodiqu)
 (NP هـ -hu))))

أنه من الممكن تصديقه

>an~a+hu + min + Al+mumokin+i + taSodiyq+u+hu
 that + from + the+possible + believe+him
That it was impossible to believe him

2. If the hu- is followed only by a nominal, adjectival or a prepositional phrase predicate, the –hu must be referential and cannot be anticipatory.

(SBAR إِنَّ >an~a
 (S (NP-SBJ هـ hu)
 (PP-PRD من min
 (NP الجنوب Aljanuwbi)))

أنه من الجنوب

>an~a+hu + min + Al+januwbi+i
 that+he + from + the+south
That he is from the south

3. The anticipatory pronoun is almost always masculine singular (-hu) or rarely feminine singular (-hA) because of some attraction with the overt subject. However, the referential pronoun can be any gender or number, in agreement with the entity that it refers to.

(SBAR >an~a إِنَّ
 (S (NP-SBJ hu هـ)
 (ADJP-PRD majonuwnN مَجْنُونُ)))

أنه مجنون

>an~a+hu + majonuwn+N
 that+he + insane
That he is insane

(SBAR >an~a إِنَّ
 (S (NP-TPC-1 hu هـ)
 (VP لا لا
 يَعْرِفُ yaEorifu
 (NP-SBJ-1 *T*)
 (NP-OBJ AlmustaHiyla المُستحيل))))

أنه لا يعرف المستحيل

>an~a+hu + lA + yaEorif+u + Al+musotaHiyl+a
 that+he + not+know + the+impossible
He does not believe in the impossible

```
(SBAR >an~a أَنْ
  (S (NP-TPC-1 hu هُوَ )
    (VP kAna كَانَ )
      (VP (PRT lA لَا )
        yaEorifu يَعْرِفُ )
        (NP-SBJ-1 *T* )
        (NP-OBJ AlmustaHiyla الْمُسْتَحِيلُ ) ) ) )
```

أنه كان لايعرف المستحيل

>an~a+hu + kAn+a + lA + yaEorif+u + Al+musotaHiyl+a
 that+he + was + not + know + the+impossible
That he did not believe in the impossible

3.6 Relative Clauses

Relative clauses are ALWAYS adjoined to the NP they modify:

```
(NP (NP -Al+mawAd~i:::-المواد:::the+substances/materials+[def.gen.]
  Alhydrwkrbwnyp:::الهيدروكربونية:::nogloss )
  (SBAR (WHNP-2 Al~atiy:::التي:::which/who/whom_[fem.sg.] )
    (S (VP (PRT lam:::لَمْ:::did_not )
      ta+Hotariq+o:::تَحْتَرَقُ:::it/they/she+burn_up/be_burned+[jus.]
      (NP-SBJ-2 *T*)) ) ) )
```

المواد الهيدروكربونية التي لم تحترق

Al+mawAd~i + Al+hiydoruwkarobuwniy~ap+i + Al~atiy + lam + taHotariq+o
 the+materials + the+hydro carbonic + that + did not + burn
The hydro carbonic materials that did not burn

The relative clause SBAR "التي لم تحترق" (Al~atiy lam ta+Hotariq+o/ which she did not burn) is adjoined to the head NP "المواد" (AlmawAd~imaterials).

Relative pronouns:

الذي Al~a*iy
 التي Al~tiy
 اللذان Al~l*Ani (الذين Al~la*ayoni)
 اللتان Al~ltAni (اللتين Al~latayoni)
 الذين Al~a*iyna
 اللاتي Al~lAtiy
 اللواتي Al~lawAtiy
 اللائي Al~lA{iy

Free relative pronouns:

من man
 ما mA

>ay~u can be also a relative pronoun with an obligatory complement in a free relative clause or in a question

Regular relative pronouns can head free relatives if they do not modify a NP.

Relative and interrogative adverbs:

حيث Hayovu (cannot be interrogative)

متى mataY

كيف kayofu

أين >ayona

لماذا lima*A

NB: need glosses for the above lists

For the most part, we do not distinguish among levels or "scope" of modification – all adjuncts are at the same level, sisters of the head NP, with the exception of relative clauses and appositive NPs. **An extra level of NP adjunction is added for the relative clause SBAR and/or for the appositive NP.**

3.6.1 Free relatives

Two relative pronouns are always part of free relatives in Arabic: *man* and *mA*. Free relatives have the internal structure of relative clauses (WH- word with a co-indexed trace, SBAR node label), but they normally function in nominal positions. As such, they are also annotated with the –NOM dashtag. So, most free relatives should have the node label SBAR-NOM.

Note, however, that if the free relative is the complement of a verb that subcategorizes for an SBAR complement (verbs of speech, for example), then the free relative should be simply SBAR (not SBAR-NOM).

See section ????? for more on dashtags, and section ????? for more on relative clauses.

```
(PP-CLR bi-::بـ::by/with
  (SBAR-NOM (WHNP-1 -mA::ما::what )
    (S (VP Talab+a-::طلب-::request+he/it_[verb]
      (NP-OBJ (NP -hu::هو::it/him )
        (NP-1 *T*))
      (PP min-::من::from
        (NP -hu::هو::it/him ))
      (NP-SBJ Al+jinirAl+u::الجنرال::the+general+[def.nom.]
        Al+<isrA}iyliy~+u::الإسرائيلي::the+Israeli+[def.nom.]
        Al+Eajuwz+u::العجوز::the+old_person+[def.nom.] )))))
  بما طلب منه الجنرال الإسرائيلي العجوز
  bi+mA + Talab+a + hu + min+hu + Al+jinirAl+u + Al+<isrA}iyliy~+u + Al+Eajuwz+u
  with+what + requested + from+him + the+general + the+Israeli + the+old
  With what the old Israeli general requested from him
```

>ay~u can be also a relative pronoun with an obligatory complement in a free relative clause or in a question

All of the free relative pronouns can also be used as interrogative pronouns in SBARQ.

A free relative that is coordinated with an NP or that is modified by a post-modifier (such as an NP in apposition) is treated as if it is itself an NP – that is, the higher adjunction level should be NP.

```
(NP jamiyEa جَمِيعَ
  (NP (SBAR-NOM (WHNP-1 mA مَا
    (S (NP-PRD Hawla حَوْلَ
      (NP hA هَا ) )
      (NP-SBJ-1 *T* ) ) )
    (PP min مِن
      (NP Albuyuwti البُيُوتِ ) ) ) ) )
```

جَمِيعَ مَا حَوْلَهَا مِنَ الْبُيُوتِ

jamiyEa mA HawlahA min Albuyuwti
 ‘all that (is) around+her of the+houses’
All the houses that are around it

```
(S (NP-SBJ (SBAR-NOM (WHNP-1 الذي
  (S (VP نطلب
    (NP-SBJ *)
    (NP-OBJ (NP هـ )
      (NP-1 *T* ) ) ) ) )
    (NP هو ) )
  (NP-PRD الحرية ) ) )
```

الذي نطلبه هو الحرية

NB: need transliteration, also in the tree

That asked we I the freedom
All we ask for is freedom

```
(NP (NP قراءاته
  (NP هـ )
  و
  (SBAR-NOM (WHNP-1 مَا
    (S (VP كتب
      (NP-SBJ *)
      (NP-OBJ (NP هـ )
        (NP-1 *T* ) ) ) ) ) ) ) )
```

قراءاته و ما كتبه

NB: need transliteration, also in the tree

Readings his and what wrote he it
His readings and all that he wrote

Note, however, that if the free relative is the complement of a verb that subcategorizes for an SBAR complement (verbs of speech, for example), then the free relative should be simply SBAR (not SBAR-NOM).

(S (VP سأل
 (NP-SBJ *)
 (SBAR (WHNP- 1 ماذا
 (S (VP حصل
 (NP-SBJ-1 *T*)))))))

سأل ماذا حصل

NB: need transliteration, also in the tree

Asked he what happened it

He asked what happened

3.7 Numbers, Quantifiers and Comparatives

3.7.1 Numbers: Cardinal vs. ordinal, noun vs. adjective

Examples	POS	TB
1. One-token numerals: 1.1. from 1 to 10: Cardinals are morphologically distinguishable from ordinal (with the pattern fAEilN or afEal)		
رجل واحد	NOUN + NOUN	(NP (NP رجل rajul man) (NP واحد wAHid one)) One man
واحد من الرجال vs	NOUN+ PREP+NOUN	(NP (NP واحد wAHid one) (PP من min from (NP الرجال Al+rijAl the+men))) One of the men
الرجل الأول	NOUN + ADJECTIVE	(NP الرجل Al+rajul the+man الأول Al+>aw~al the+first) The first man
أول الرجال	ADJECTIVE + NOUN	(ADJP أول >aw~al first (NP الرجال Al+rijAl the+men)) The first of men
أربعة كتب	NOUN + NOUN	(NP أربعة >arobaEap four (NP كتب kutub books)) Four books
كتب أربعة vs	NOUN + NOUN	(NP (NP كتب kutub books) (NP أربعة >arobaEap four)) Four books
الكتاب الرابع	NOUN + ADJECTIVE	(NP الكتاب Al+kitAb the+book الرابع Al+rAbiE the+fourth)

<p>رابع الكتب</p>	<p>ADJECTIVE + NOUN</p>	<p>The fourth book (ADJP رابع rAbiE fourth (NP الكتب Al+kutub the+books)) The fourth one of the books</p>
<p>خمس غرف</p>	<p>NOUN + NOUN</p>	<p>(NP خمس xams five (NP غرف guraf rooms)) Five rooms</p>
<p>غرف خمس</p>	<p>NOUN + NOUN</p>	<p>(NP (NP غرف guraf rooms) (NP خمس xams five)) Five rooms</p>
<p>الغرفة الخامسة</p>	<p>NOUN + ADJECTIVE</p>	<p>(NP الغرفة Al+gurofap the+room الخامسة Al+xAmisap the+fifth) The fifth room</p>
<p>خامس الغرف</p> <p>1.2.20, 30, 40 100, 1000, 1000000 etc.: There is no difference in the form of the word between cardinal and ordinal. We can distinguish them by using these two parameters</p>	<p>ADJECTIVE + NOUN</p>	<p>(ADJP خامس xAmis fifth (NP الغرف Al+guraf the+rooms)) The fifth of the rooms</p>
<p>1.2.1. ordinal is always in these number postnomi nal الصفحة الخمسون</p>	<p>NOUN + ADJECTIVE</p>	<p>(NP الصفحة Al+SafoHap the+page الخمسون Al+xamosuwn the+fiftieth) The fiftieth page</p>
<p>السنة المائة vs cardinal can be prenominal or postnominal</p>	<p>NOUN + ADJECTIVE</p>	<p>(NP السنة Al+sanap the+year المائة Al+mA}ap the+hundredth) The hundredth year</p>

<p>خمسون صفحة</p>	<p>NOUN + NOUN</p>	<p>(NP خمس xamos five (NP غرف guraf rooms)) Five rooms</p>
<p>الصفحات الخمسون الأخيرة</p>	<p>NOUN + NOUN</p>	<p>(NP (NP (NP الصفحات Al+SafaHAt the+pages) (NP الخمسون Al+xamosuwn the+fifty) (ADJP الأخيرة Al+>axiyrap the+last))) The last fifty pages</p>
<p>1.2.2. when cardinal is postnominal the noun modified has to be plural vs for ordinal the noun modified is always singular</p>		
<p>ألف ليلة</p>	<p>NOUN + NOUN</p>	<p>(NP ألف >alof thousand (NP ليلة layolap night)) Thousand nights</p>
<p>الليالي الألف</p>	<p>NOUN + NOUN</p>	<p>(NP (NP الليالي Al+layAliy the+nights) (NP الألف The+thousand)) The thousand nights</p>
<p>الليلة الألف</p>	<p>NOUN + ADJECTIVE</p>	<p>(NP الليلة Al+layolap the+night الألف Al+>alof the+thousandth) The thousandth night</p>
<p>2. Multi-token numerals : The distinction between cardinal and ordinal is based on the form of the first term : if it is cardinal so all the number is cardinal = NOUN if it is ordinal the whole number is ordinal = ADJECTIVE</p>		

واحد وعشرون قصيدة	NOUN + CONJ + NOUN	(NP واحد wAHid one و wa and عشرون Ei\$oruwn twenty (NP قصيدة qaSiydap poem)) Twenty one poems
الواحد القصائد والعشرون vs	NOUN + CONJ + NOUN	(NP (NP القصائد Al+qaSA}id the +poems) (NP الواحد Al+wAHid the+one و wa and العشرون Al+Ei\$oruwn the+twenty)) The twenty one poems
الحادية القصيدة والعشرون	ADJECTIVE + CONJ + ADJECTIVE	(NP (NP القصيدة AL+qaSidap the+poem) (ADJP الحادية Al+HAdiyap the+first و wa and العشرون Al+Ei\$oruwn the+twenty)) The twenty first poem
We can also use the parameter 1.2.2. to distinguish cardinal from ordinal when there is no first term (1 to 9)		
مائة وخمسون صفحة	NOUN+CONJ+NO UN	(NP مائة mA{ap hundred و wa and خمسون xamosuwn fifty (NP صفحة SafoHap page)) One hundred and fifty pages
المائة الصفحات والخمسون vs	NOUN+CONJ+NO UN	(NP (NP الصفحات Al+SafaHAt the+pages) (NP المائة Al+mA}ap the+hundred و wa and الخمسون Al+ xamosuwn the+fifty)) The hundred and fifty pages
المائة الصفحة والخمسون	ADJ + CONJ + ADJ	(NP (NP الصفحة Al+SafoHap the+page) (ADJP المائة Al+mA}ap the+hundred و wa and الخمسون Al+ xamosuwn the+fifty)) The hundred and fiftieth page
3. More examples مرتين اثنتين	NOUN + NOUN	(NP (NP مرتين mar~at+ayoni times+dual) (NP اثنتين <ivnata+ayoni two+ dual)) Two times
القصة الأولى	NOUN + ADJ	(NP القصة Al+qiS~ap the+story الأولى Al+&uwlAY the first) The first story

اليوم الحادي عشر	NOUN + ADJ + ADJ	(NP (NP (NP اليوم Al+yawom the+day) (ADJP الحادي Al+HAdiy the+first عشر Ea\$ar ten)) The eleventh page
اثنا عشر شهرا	NOUN + NOUN + NOUN	(NP (NP (NP اثنا <ivnA two عشر Ea\$ar ten (NP (NP شهرا \$ahor+F month+Acc)) Twelve months
ثلاثة وثلاثون عاما	NOUN+CONJ+NO UN	(NP (NP (NP ثلاثة valAvap three و wa and ثلاثون valAvuwn thirty (NP (NP عاما EAm+F year+Acc)) Thirty three years
الليالي العشر	NOUN + NOUN	(NP (NP (NP الليالي Al+layAliy the+nights) (NP (NP العشر Al+Ea\$r the+ten)) The ten nights
المرّة المليون	NOUN + ADJ	(NP (NP (NP المرّة Al+mar~ap the +time المليون Al+maloyuwn the+million) The millionth time
ست وثلاثون وخمسمائة دينار	NOUN+CONJ+NO UN+CONJ + NOUN	(NP (NP (NP ست sit~ six و wa and ثلاثون valAvuwn و wa and خمسمائة xamosumA }ap five hundred (NP (NP دينار dinAr dinar)) Five hundred and thirty six dinars
مابين خمسة وعشرين وثلاثين جريحا	NOUN+CONJ+NO UN NOUN + NOUN	(SBAR-NOM (WHNP-1 ما mA what) (S (NP-SBJ-1 (*T*) (NP-PRD بين bayon between (NP (NP خمسة xamosap five و wa and عشرين Ei\$oriyn twenty و wa and (NP (NP ثلاثين valAviyn thirty (NP (NP جريحا jariyH+F injured+Acc)))))) between twenty five and thirty injured people
مائة وخمسة وثلاثون أو مائة وأربعة وستون مشاركا	NOUN + CONJ + NOUN + CONJ + NOUN NOUN + CONJ + NOUN + CONJ +	(NP (NP (NP مائة mA }ap hundred و wa and خمسة xamosap five و wa and ثلاثون valAvuwn thirty

	NOUN	(NP-1 *RNR*) أو >aw or (NP مائة mA}ap hundred و wa and أربعة >arobaEap four و wa and ستون sit~uwn sixty (NP-1 *RNR*) (NP-1 مشاركا mu\$Arik+F participant) One hundred and thirty five or one hundred and sixty four participants
الساعة الثانية عشر	ADJ+ADJ	(NP الساعة Al+sAEap the+hour الثانية Al+vAniyap the+second عشر Ea\$ar ten) The twelfth hour
الساعات الأربعة والعشرون الآتية	NOUN+CONJ+NO UN	(NP (NP (NP الساعات Al+sAEAt the+hours) (NP الأربعة Al+>arobaE the+four و wa and العشرون Al+Ei\$ruwn the twenty) (ADJP الآتية Al+ Atiyap the+next))) The next twenty four hours
الحلقة الخامسة والأربعون	ADJ +CONJ+ADJ	(NP (NP الحلقة Al+Halaqap the+episode (ADJP الخامسة Al+xAmisap the+fifth و wa and الأربعون Al+>arobaEuwn the fourtieth)) The forty fifth episode

3.7.2 Numbers

Examples of coordinated numbers with right node raised (*RNR*) complements:

```

(NP (NP 5
    (NP آلاف thousand
      (NP-2 *RNR*)))
  or
  (NP 20
    (NP ألف thousand
      (NP-2 *RNR*)))
  (NP-2 رجل men))

```

5 آلاف أو 20 ألف رجل

NB: need transliteration, also in the tree

5 thousands or 20 thousand men

5 thousand or 20 thousand men

```
(NP (NP اثنان two
    و and
    عشرون twenty
    (NP-3 *RNR*))
    او or
    (NP خمسة five
    و and
    ثلاثون thirty
    (NP-3 *RNR*))
    (NP-3 دولارا dollars))
```

اثنان و عشرون او خمسة و ثلاثون دولارا

NB: need transliteration, also in the tree

Two and twenty or five and thirty dollars

Twenty-two or thirty-five dollars

This is because number terms are considered as one word (both in the treebank annotation and also in Arabic grammar), so the idafa works even though there is more than one token for the number. This is not possible with regular, non-numeric nouns.

```
(NP ثلاثة three
    او or
    اربعة four
    (NP دولارات dollars))
```

ثلاثة او اربعة دولارات

NB: need transliteration, also in the tree

Three or four dollars

Three or four dollars

Below are examples of numbers in many different constructions.

```
(PP bi ب
    (NP 40 40
        (NP milyuwna مليون
            (NP duwlArK دولار ) ) ) )
```

ب 40 مليون دولار

bi 40 milyuwna duwlArK

with 40 million dollar(s)

with 40 million dollars

```
(NP 530
    (NP miloyuwna مئوون
        (NP duwlArK دولار ) ) )
```

530 مئوون دولار

530 miloyuwna duwlArK

530 million dollar(s)

530 million dollars

(PP <ilaY إلى
 (NP zuhA'i زهاء
 (NP 22
 (NP milyuwnAF مليوناً ())))

إلى زهاء 22 مليوناً

<ilaY zuhA'i 22 milyuwnAF
 towards proximity 22 million(s)
to about 22 million

(NP HawAlay حوالي
 (NP 8000
 (NP m م
 2)))

حوالي 8000 م 2

HawAlay 8000 m 2
 about 8000 m2
about 8000 m2

(NP (NP 17
 (NP milyuwna مليون
 (NP mitriK متر
 murab~aEiK مُرَبَّع ()))
 (PP min من
 (NP Al>arADiyi الأراضي
 All~ubnAniy~api اللُّبْنَانِيَّة ()))

17 مليونَ مترَ مُرَبَّعٍ من الأراضي اللُّبْنَانِيَّة

17 milyuwna mitriK murab~aEiK min Al>arADiyi Al~l~ubnAniy~api
 17 million meter square from the+land(s) the+Lebanese

NB: need translation

(ADJP >akvara أَكْثَرُ
 (PP min من
 (NP 25
 (NP EAmAF عاماً ())))

أَكْثَرَ مِنْ 25 عاماً

>akvara min 25 EAmAF
 more than 25 year(s)
more than 25 years

(NP (NP 60
 (NP %))
 (PP min من
 (NP Alrusuwmi الرُّسُومِ ()))

60% مِنَ الرُّسُومِ

60 % min Al+rusuwmi
 60% of the+documents
60% of the documents

(S (NP-SBJ >arbAHu أرباح
 (NP AlS~aydaliy~i الصَّيْدَلِيَّ))
 (NP-PRD (NP 22.5 22.5)
 (PP fiy في
 (NP Almi}api المِئَةِ))))

أرباح الصَّيْدَلِيَّ 22.5 في المِئَةِ

>arbAHu AlS~aydaliy~i 22.5 fiy Almi}api
 (the) gains (of) the+pharmacist (are) 22.5 per cent

NB: need translation

(NP naHwa نَحْوَ
 (NP niSfi نصف
 (NP malyuwni مَلْيُون
 (NP muwATinK مُوَاطِنِ))))

نَحْوَ نِصْفِ مَلْيُونِ مُوَاطِنِ

naHwa niSfi malyuwni muwATinK
 towards half (a)million (of)citizen(s)

NB: need translation

(NP naHwa نَحْوَ
 (NP (NP (NP Tun~iK طُنَّ)
 wa وَ
 (NP niSfi نصف
 (NP Tun~iK طُنَّ)))
 (PP min مِنْ
 (NP Al>agTiy~api الأَعْطِيَّةِ))))

نَحْوَ طُنَّ وَ نِصْفِ طُنَّ مِنْ الأَعْطِيَّةِ

naHwa Tun~K wa+niSfi Tun~K min Al>agTiy~api
 towards (a)ton and+half (of a)ton of the+blankets

NB: need translation

(NP *ikraY ذَكَرَى
 (NP (NP 4)
 (NP-TMP jwiylap جَوِيلِيَّةِ)))

ذَكَرَى 4 جَوِيلِيَّةِ

*ikraY 4 jwiylap
 commemoration (of) 4 July

NB: need translation

ADJUNCTION / see also 24 mArs + <ilaY 17 >uktuwbir 2002

NB: is this an example that needs to be added?

(NP-TMP baEda بَعْدَ
 (NP 3
 >aw أَوْ
 4
 (NP >a\$huriK أَشْهُرِ)))

بَعْدَ 3 أَوْ 4 أَشْهُرِ

baEda 3 >aw 4 >a\$huriK
 after 3 or 4 months

NB: need translation

(SBAR-NOM (WHNP-1 mA ما)
 (S (NP-PRD bayna بَيْنَ
 (NP 3
 wa وَ
 4
 (NP malAyiyna مَلَائِيْن)))
 (NP-SBJ-1 *T*)))

ما بَيْنَ 3 وَ 4 مَلَائِيْن

mA bayna 3 wa 4 malAyiyna
 What (is) between 3 and 4 millions

NB: need translation

(ADJP >akvaru أَكْثَرُ
 (PP min مِنْ
 (NP (NP 50)
 (PP fiy فِي
 (NP Almi}api المِئَةِ)))
 (PP min مِنْ
 (NP Al>a\$xASi الأشخاص)))))

أَكْثَرُ مِنْ 50 فِي المِئَةِ مِنَ الأشخاص

>akvaru min 50 fiy Al+mi}api min Al +>a\$xASi
 more than 50 per the+(one)cent of/from the+people(s)

NB: need translation

(PP fiy فِي
 (NP Huduwdi حُدُودِ
 (NP Al ال
 2600
 (NP mudar~isK مُدَرِّس))))

فِي حُدُودِ ال-2600 مُدَرِّس

fiy Huduwdi Al-2600 mudar~isK
 in (the)limits (of) the-2600 teacher(s)

NB: need translation

(NP-TMP (NP Als~AEapu السَّاعَةُ
 19.07)
 (PP bi بِ
 (NP Altawqiyti التَّوْقِيْتِ
 AlmaHal~iy~i المَحَلِّيَّ)))

السَّاعَةُ 19.07 بِالتَّوْقِيْتِ المَحَلِّيَّ

Als~AEapu 19.07 biAltawqiyti AlmaHal~iy~i
 The+hour (is) 19:07 with+the+time the+local

NB: need translation

(NP naHwa نَحْوَ
 (NP 23
 (NP (NP HizbAF حزباً
 wa وَ
 (NP munaZ~amapaF مُنَظَّمَةً
 AinofisAliy~apF اِنْفِصَالِيَّةً))))

نَحْوَ 23 حزباً وَمُنَظَّمَةً اِنْفِصَالِيَّةً

naHwa 23 HizbAF wamunaZ~amapaF AinfiS Aliy~apF
about 23 parti(es) and+organization(s) separatist

NB: need translation

(NP (NP 1.2
(NP milyuwna مَلْيُونْ
(NP zA}iriK زَائِرِ ()))
(PP fiy فِي
(NP muqAbili مُقَابِلِ
(NP (NP (ADJP >aqal~i أَقْلُ
(PP min مِنْ
(NP 500
(NP >alofiK أَلْفِ ())))
(NP-TMP Alywama الْيَوْمَ ()))))
1.2 مَلْيُونْ زَائِرِ فِي مُقَابِلِ أَقْلُ مِنْ 500 أَلْفِ الْيَوْمَ

1.2 milyuwna zA}iriK fiy muqAbili >aqal~i min 500 >alofiK Alyawma

1.2 million visitor(s) in opposite/contrast (with) less than 500 thousand the+day

NB: need translation

(NP (NP (NP 4
(NP malAyiyna مَلَايِينْ
(NP-1 *RNR* ()))
wa وَ
(NP 470
(NP >alfa أَلْفِ
(NP-1 *RNR* ()))
(NP-1 duwlArK دُولَارِ ()))
(NP (NP (PRT >ay أَيِ
10)
(PP fiy فِي
(NP Almi}api الْمِئَةِ ()))
(PP min مِنْ
(NP qiymapi قِيَمَةِ
(NP Ald~uyuwuni الدُّيُونِ ())))))
4 مَلَايِينْ وَ 470 أَلْفِ دُولَارِ أَيِ 10 فِي الْمِئَةِ مِنْ قِيَمَةِ الدُّيُونِ

4 malAyiyna wa 470 >alfa duwlArK >ay 10 fiy Almi}api min qiymapi Ald~uyuwuni

4 millions and 470 thousand dollar(s) namely 10 in the+cent/hundred from value (of)
the+debts

NB: need translation

(NP naHwa نَحْوَ
(NP (NP 47,4
(PP fiy فِي
(NP Almi}api الْمِئَةِ ())))

نَحْوَ 47,4 فِي الْمِئَةِ

About 47,4 per/in the+cent/hundred

NB: need transliteration and translation

(NP 2150
(NP kitAbAF كِتَاباً ()))

2150 كِتَاباً

2150 kitAbAF

2150 book(s)

2,150 books

(NP >alfAni أَلْفَانِ
 wa وَ
 m}apuN مِئَةُ
 wa وَ
 xamsuwna خَمْسُونَ
 (NP kitaAbAF كِتَاباً))

أَلْفَانِ وَمِئَةُ وَخَمْسُونَ كِتَاباً

>alfAni wa+mi}apuN wa+xamsuwna kitAbAF
 thousand(-two) and+(one-)hundred and+fifty book(s)

NB: need translation

(NP xamsapuN خَمْسَةُ
 wa وَ
 Ei\$ruwna عِشْرُونَ
 (NP kitAbAF كِتَاباً))

خَمْسَةُ وَعِشْرُونَ كِتَاباً

xamsapuN wa+Ei\$ruwna kitAbAF
 five and+twenty book(s)
twenty-five books

(NP (NP Al+kitAbu الْكِتَابُ)
 (ADJP Al+xAmisu الْخَامِيسُ
 wa وَ
 Al+Ei\$ruwna الْعِشْرُونَ))

الْكِتَابُ الْخَامِيسُ وَالْعِشْرُونَ

Al+kitAbu Al+xAmisu wa+Al+Ei\$ruwna
 the+book the+fifth and+the+ twenty
the twenty-fifth book

(NP kul~u كُلُّ
 (NP <ivnayo اِثْنَيْ
 Ea\$ara عَشَرَ
 (NP kitAbAF كِتَاباً)))

كُلُّ اِثْنَيْ عَشَرَ كِتَاباً

kul~u ivnayo Ea\$ara kitAbAF
 every two ten (of) book(s)
NB: need translation

(NP (NP <ivnA اِثْنَا
 Ea\$ara عَشَرَ
 (NP kitAbAF كِتَاباً))
 (NP kul~u كُلُّ
 (NP hum هُمْ)))

اِثْنَى عَشَرَ كِتَاباً كُلُّهُمْ

<ivnA Ea\$ara kitAbAF kul~uhum
 two ten book(s) all+(of)them

NB: need translation

(S (VP zurtu زُرْتُ
 (NP-SBJ *)
 (NP-OBJ hum هُمْ)
 (NP-TMP kul~a كُلَّ
 (NP xamsapi خَمْسَةَ
 (NP >ay~AmiK أَيَّامٍ () () ())

زُرْتُهُمْ كُلَّ خَمْسَةِ أَيَّامٍ

zurtuhum kul~a xamsapi >ay~AmiK
 visited(I)+them every five (of)days

NB: need translation

(S (VP jalastu جَلَسْتُ
 (NP-SBJ *)
 (NP-ADV maEa مَعَ
 (NP hum هُمْ))
 (NP-TMP (NP xamsapa خَمْسَةَ
 (NP >ay~AmiK أَيَّامٍ ())
 (NP kul~a كُلَّ
 (NP hA هَا () () ())

جَلَسْتُ مَعَهُمْ خَمْسَةَ أَيَّامٍ كُلِّهَا

jalastu maEahum xamsapa >ay~AmiK kul~ahA
 spent(I) with+them five days all+(of)them (Fem.Pl)

NB: need translation

(S (VP (PRT lam لَمْ)
 ya>otin يَأْتِ
 (NP-OBJ niy نِي)
 (NP-SBJ kul~u كُلُّ
 (NP AlnaAsi النَّاسِ () () ())

لَمْ يَأْتِي كُلُّ النَّاسِ

lam ya>otiniy kul~u AlnaAsi
 not come+(to)me all (of) the+people
Not all the people came to me

(S (VP (PRT lam لَمْ)
 ya>oti يَأْتِ
 (NP-SBJ (NP AlnAsu النَّاسِ)
 (NP kul~u كُلُّ
 (NP hum هُمْ () () ())

لَمْ يَأْتِ النَّاسُ كُلُّهُمْ

lam ya}oti Al+nAsu kul~u+hum
 not come the+people all+(of)them

NB: need translation

(S (NP-SBJ ha*aA هَذَا)
 (ADJP-PRD >akvaru أَكْثَرُ
 (PP min مِنْ
 (NP kul~i كُلُّ
 (NP Alkutubi الْكُتُبِ () () ())

(NP-ADV EilmAF علماً ()))

هَذَا أَكْثَرُ مِنْ كُلِّ الْكُتُبِ عِلْمًا

ha*aA >akvaru min kul~u Al+kutubi EilmAF
 this more than all (of) the+books (in) knowledge

NB: need translation

(NP HawAlay حَوَالِي
 (NP kul~i كُلِّ
 (NP Alkutubi الْكُتُبِ ()))

حَوَالِي كُلِّ الْكُتُبِ

HawAlay kul~i Al+kutubi
 almost all (of) the+books

NB: need translation

(NP qurAbapa قُرَابَة
 (NP <ivnayo إِثْنَيْ
 Ea\$ara عَشْرَ
 (NP kitAbAF كِتَاباً ()))

قُرَابَة إِثْنَيْ عَشْرَ كِتَابًا

qurAbapa <ivnaY Ea\$ara kitAbAF
 close(to) two ten book(s)

NB: need translation

(NP bayna بَيْنَ
 (NP 12
 wa وَ
 20
 (NP kitAbAF كِتَاباً ()))

بَيْنَ 12 وَ 20 كِتَابًا

bayna 12 wa 20 kitAbAF
 between 12 and 20 book(s)
between 12 and 20 books

(NP 12
 >aw أَوْ
 20
 (NP kitAbAF كِتَاباً ()))

12 أَوْ 20 كِتَابًا

12 >aw 20 kitAbAF
 12 or 20 book(s)
12 or 20 books

(S (VP balagat بَلَّغَتْ
 (NP-SBJ *)
 (NP-OBJ (NP 12)
 (PP <ily إِلَى
 (NP 20
 (NP duwlArAF دُولَارًا ()))))))))

بَلَّغَتْ 12 إِلَى 20 دُولَارًا

balagat 12 <ilY 20 duwlArAF
reached(it/she) 12 to 20 dollar(s)

NB: need translation

(PP min مِن
(NP (NP 12)
(PP <ilY إِلَى
(NP 20
(NP duwlArAF دُولَاراً)))))

مِن 12 إِلَى 20 دُولَاراً

min 12 <ilY 20 duwlArAF
from 12 to 20 dollar(s)
from 12 to 20 dollars

لَمْ اشْتَرِ إِلَّا اثْنَيْ عَشَرَ كِتَابًا

lam Aa\$tariy <il~A <ivnay Ea\$ara kitAbAF
not buy(I) except/only two ten book(s)

NB: need translation – and a tree!

جَاءَ الطَّلَبَةُ إِلَّا اثْنَيْ عَشَرَ مِنْهُمْ

jaA' AlTalabapu <il~A <ivnay Ea\$ara minhum

NB: need translation and gloss – and a tree!

(S (NP-SBJ *)
(PRT la لَيْسَ)
(ADJP-PRD >aqal~a أَقَلُّ
(PP min مِن
(NP 12
(NP kitAbAF كِتَابًا)))))

لَيْسَ أَقَلُّ مِن 12 كِتَابًا

laysa >aqal~a min 12 kitAbAF
'(there are) no fewer than 12 book(s)

NB: need translation

(S (NP-TPC-1 huwa هُوَ)
(VP laysa لَيْسَ
(NP-SBJ-1 *T*)
(ADJP-PRD >akvara أَكْثَرَ
(PP min مِن
(NP >axiy أَخِي
(NP hi هِ)))
(NP-ADV mAlAF مَالاً))))

huwa >akvaru mAliK (??) min >axiyhi

NB: need translation and gloss and Arabic

(S (VP (PRT lA لا)
 yaqil~u يَقِلُّ
 (NP-SBJ *)
 (PP-CLR Ean عَنْ
 (NP 12
 (NP kitAbAF كِتَاباً)))))

لا يَقِلُّ عَنْ 12 كِتَابًا

lA yaqil~u Ean 12 kitAbAF
 not (it/he)make(s)-fewer than 12 book (s)

NB: need translation

(NP (NP 12
 (NP kitAbAF كِتَاباً)
 (PP EalaY عَلَى
 (NP Al>aqal~i الْأَقْلُ)))

12 كِتَابًا عَلَى الْأَقْلُ

12 kitAbAF EalaY Al+aqal~i
 12 book(s) at the+least

NB: need translation

(NP valAvapu ثَلَاثَةُ
 (NP >aDEAfi أضعاف
 (NP AlxasA}iri الخسائر)))

ثَلَاثَةُ أَضْعَافِ الْخَسَائِرِ

valAvapu >aDEAfi AlxasA}iri
 three time(s) (of) the+losses

NB: need translation

(ADJP >akvaru أَكْثَرُ
 (PP min مِنْ
 (NP (NP 12
 (NP-1 *RNR*)
 >aw أَوْ
 (NP (ADVP rub~amA رُبَّمَا) → check ADVP?????
 20
 (NP-1 *RNR*))
 (NP-1 kitAbAF كِتَاباً)))))

أَكْثَرُ مِنْ 12 أَوْ رُبَّمَا 20 كِتَابًا

>akvaru min 12 >aw rub~amA 20 kitAbAF
 more than 12 or maybe 20 books

NB: need translation

(NP EAmAni عامان
>aw أو
valAvapuN ثلاثة
>aw أو
>akvaru أكثر)

عامان أو ثلاثة أو أكثر

EAmAni >aw valAvapuN >aw >akvaru
Years(-two) or three or more

NB: need translation

(NP-TMP (NP sAEapuN ساعة)
(NP-TMP taqriybAF تقريباً)

ساعة تقريباً

sAEatuN taqriybAF
hour(one) approximately

NB: need translation

(NP (NP 12
(NP kitAbAF كتاباً)
>aw أو
(SBAR-NOM (WHNP-1 mA ما)
(S (VP yuqAribu يُقاربُ
(NP-SBJ-1 *T*)
(NP-OBJ *alika ذلك)))))

12 كتاباً أو ما يُقاربُ ذلكَ

12 kitAbAF >aw mA yuqAribu *alika
12 books or what nears that

NB: need translation

(NP (NP 800
(NP kiyluwgrAmAF كيلوغراماً))
wa و
(SBAR-NOM (WHNP-1 mA ما)
(S (NP-PRD fawqa فوق
(NP *alika ذلك))))
(NP-SBJ-1 *T*))))

800 كيلوغراماً وما فوقَ ذلكَ

800 kiyluwgrAmAF wa+mA fawqa *alika
800 kilogram(s) and+what (is) above that

NB: need translation

(S (VP *ahaba ذَهَبَ
(NP-SBJ AlTawiyly الطويلُ)
(PP-DIR >ily إلى
(NP Almadorasapi المدرسة))))

ذَهَبَ الطَّوِيلُ إِلَى المَدْرَسَةِ

*ahaba AlTawiyly >ily Almadorasapi
went(he) the+long(one) to the+school

NB: need translation

```

(NP jamiyEa جَمِيعَ
  (NP (SBAR-NOM (WHNP-1 mA مَا
    (S (NP-PRD Hawla حَوْلَ
      (NP hA هَا ) )
      (NP-SBJ-1 *T* ) ) )
    (PP min مِن
      (NP Albuyuwti البُيُوتِ ) ) ) ) )

```

جَمِيعَ مَا حَوْلَهَا مِنَ الْبُيُوتِ

jamiyEa mA HawlahA min Albuyuwti
all that (is) around+her of the+houses

NB: need translation

3.7.3 Akvr “most”

Akvr can have 2 possible POS core tags: ADJ_COMP, NOUN_QUANT. When it is NOUN_QUANT, it requires only an NP complement (i.e. a muDAf <ilayh). When it is an ADJ_COMP and it is followed by an NP complement, it requires an NP-ADV in addition. ADJ_COMP can also have an NP-ADV and/or a PP without the NP complement.

For example:

ADJP_COMP

الرجل الأكثر جرأة

Aljuro>u + Al>akovar + juro>ap

The man + the most + audacity

The most audacious man

NOUN_QUANT

أكثر الناس

most the people

most of the people

3.7.3.1 Adjective

When it is ADJ_COMP, it should be treebanked as an adjective (single word) or ADJP (multiple words):

- The template for Akvr as a superlative adjective is

```

(ADJP Al>akovar
  (NP-ADV advberial semantic head))

```

Or

(ADJP >akovar
 (NP ...)
 (NP-ADV adverbial semantic head))

The NP and the NP-ADV are both optional.

(NP (NP Alrajulu الرَّجُلُ)
 (ADJP Al>akovar أَكْثَرُ
 (NP-ADV juro>apF (جُرْأَةً)))

الرجل الأكثر جرأة

Aljuro>u + Al>akovar + juro>ap

The man + the most + audacity

The most audacious man

(ADJP أَكْثَرُ
 (NP الرجال)
 (NP-ADV علما))

أكثر الرجال علما

NB: need transliteration, also in the tree

Most the men knowledge

The most knowledgeable man

- The template for Akvr as a comparative adjective is

(ADJP (Al)>akovar
 (NP-ADV ...)
 (PP min
 (NP ...)))

The NP-ADV is optional.

(S (NP-TPC-2 (NP الزيارة::Al+ziyArapu::the+visit)
 (PP ل::li-::for/to
 (NP بيروت::-bayorwta::Beirut)))
 (VP (PRT لن::lan::[will]_not/never)
 تستمر::ta+sotamir~+a::it/they/she+continue/last_[time]+[sub.]
 (NP-SBJ-2 *T*)
 (NP-TMP (ADJP أكثر::>akovar+a::more/most+[def.acc.]
 (PP من::min::from
 (NP ثلاث::valAvi::three
 (NP ساعات::sAEAtK::hours))))))

الزيارة لبيروت لن تستمر أكثر من ثلاث ساعات.

AlzyArapu + li+bayorwta + lan + tasotamir~a + >akovara + min + valAvi + sAEAtK

The visit + to+Beirut + will not + last + more + of + three + hours

The visit to Beirut will not last more than three hours

(NP (NP Alrajulu الرَّجُلُ)
 (ADJP Al>akovaru الأكثرُ
 (NP-ADV juro>apF جُرْأَةً)
 (PP min من
 (NP gayori غَيْرِ
 (NP hi هِ))))))

الرجل الأكثر جرأة من غيره

Aljuro>u + Al>akovaru + juro>ap + min + gayori+hi
 The man + the most + (of) audacity + of + other+his
The most audacious man of all

3.7.3.2 Adjective in nominal position

When an ADJ_COMP is in a nominal position, it should be treebanked as an adjective (single word) or ADJP (multiple words) with a parent NP.

NP-SBJ

(S (NP-ADV تحت::taHota::under
 (NP (NP هذا::h`*A::this_[masc.sg.]
 (NP العنوان::Al+EunowAni::the+address))
 (VP اجتمع::{ijotamaE+a::meet/confer+he/it_[verb]
 (NP-SBJ (NP (ADJP اكثر::>akovaru::more/more
 (PP من::min::from
 (NP (NP 500::500::nogloss)
 (NP طبيب::TabiybK::physician/doctor
 و::wa-::and
 اختصاصي::-{ixotiSASiy~K::specialist)
 (PP-LOC من::min::from
 (NP (NP العالم::Al+EAlami::the+world
 ربي الع::Al+Earaby~i: the+Arab)
 و::wa-::and
 (NP اوربا::->uwrwb~A::Europe)
 و::wa-::and
 (NP اميركا::>amiyrokA::America))))
 (PP-LOC في::fiy::in
 (NP (NP العاصمة::Al+EASimapi::the+capital_city
 الاردنية::Al>urodun~iy~api::the+Jordanian)
 (NP عمان::Eam~An::Amman))))))
 تحت هذا العنوان اجتمع أكثر من 500 طبيب واخصائي من العالم العربي وأوروبا وأمريكا في العاصمة الأردنية عمان
 taHota + ha*A + AlEunwAni + AijotamaEa + >akovaru + min + 500 + TabiybK +
 wa+AixotiSASy~K + min + AlEAlami + AlEaraby~i + wa+Awrb~A + wa+AmorykA
 fy + AlEASimapi + Al>urodun~iy~api + Eam~An
 under + this + title + gathered + more + of + 500 + doctors + specialist + from + the world +
 the Arab + and+Europe + and+America + in + the capital + the Jordanian + Amman
*More than 500 doctors and specialists from the Arabic world, Europe and America conferred
 under this slogan in the Jordanian capital Amman.*

NP-OBJ

```
(S (NP-TPC-3 (NP هذه::h`*ihi::this/these)
              (NP المأساة::Al+ma>osAapu::the+tragedy
                  الانسانية::Al+<inosAniy~apu::the+human/humane/humanitarian))
  (VP تطاول::tu+TAwil+u::it/they/she+contend_with+[ind.]
    (NP-SBJ-3 *T*)
    (NP-OBJ (ADJP اكثر::>akovar+a::more+[def.acc.]
              (PP من::min::from 600::600::nogloss
                (NP (NP بيت::bayotK::house)
                    (ADJP كويتي::kuwayotiy~K::Kuwaiti
                        و::wa-::and
                        عربي::-Earabiy~K::Arab/Arabic))))))
    ha*ihi + Alma>osApu + Al<inosAniy~apu + tuTAWilu + >akovara + min + bayotK
    kuwayty~K + wa+Earaby~K
    this + the tragedy + tragedy + touch + more + of + house + Kuwaiti + and+Arab
    This human tragedy touches more than Kuwaiti and Arab families
```

NB: Is this tree and gloss/translation, etc. correct? It looks like “600” is missing (with “from” in the tree)

This is just like all other multi-word adjectives in nominal position are treated:

```
(NP (ADJP جميل handsome
      (NP الوجه face))
```

جميل الوجه

NB: need transliteration, also in the tree

Beautiful face

The person with a beautiful face

3.7.3.3 Noun quantifier

A token with the part-of-speech tag NOUN_QUANT should be treebanked as an NP. (This is not part of the rule about adjectives in nominal position above – because the POS tag for Akvr will be a type of NOUN in this case.)

>akvar and >aglab are noun quantifiers with the part-of-speech tag NOUN_QUANT when their meaning can be paraphrased with الأَكْثَرِيَّة /Al>akvariyy~ap/the majority and الأَغْلَبِيَّة /Al>aglabiyy~ap/the majority. As such, they take only an NP complement.

(S (NP-TMP بعد::baEoda::after
 (NP 11::11::nogloss
 (NP 9::9::nogloss)))
 (VP حصلت::HaSal+at::obtain/acquire/get+it/they/she_[verb]
 (NP-SBJ اسرائيل::<isorA}iy1+u::Israel+[def.nom.]
 (PP-CLR على::EalaY::on/above
 (NP (NP كامل::kAmil+i::complete/full/integral+[def.gen.]
 (NP الموقف::Al+mawoqifi::the+position
 الاميركي::Al+>amiyrokiy~i::the+American))
 و::wa-::and
 (NP اكثر::->akovari::most of
 (NP الموقف::Al+mawoqifi::the+position
 الاوروبي::Al+>uwruwb~iy~i::the+European))))))
 بعد 9 11 حصلت اسرائيل على كامل الموقف الأمريكي وأكثر الموقف الاوروبي
 baEoda + 11 + 9 + HaSalat + <isovA}iy1 + EalaY + kAmili + Almaoqifi + Al>amoriykiy~i
 + wa+<akovara + Almawoqifi + Al<uwruwbiy~i.
 After 9/11 Israel acquired the full American position and most of the European position
NB: need translation (or else gloss, if the above is a translation)

3.8 Differentiating when an adjective is an ADJP and when it is an NP-ADV

(S (VP Euwqiba عُوَقِبَ
 (NP-SBJ-1 اللاعبُ AllAEibu)
 (NP-OBJ-1 *)
 (PP bi ب
 (NP (NP Al<iyqAfi الإيقاف)
 (NP-ADV (ADJP >akovara أكثر
 (PP min من
 (NP mubArAtK مباراة))))))))
 عوقب اللاعب بالإيقاف أكثر من مباراة
 Euwqiba Al+lAEib+u bi+Al+<iyqAf+i >akovar+a min mubArAt+K
 was punished the+player with+the+suspension more than game
The player was punished by suspension for more than one game

Note in constructions such as the above example that the NP-ADV modifier is labeled as NP-ADV rather than simply as an ADJP because it is modifying the event of the suspension as and adverbial NP rather than modifying the noun itself in an adjectival way. Note that the adjective in this construction does not match “suspension” in case, gender, number or definiteness because it is being used as an adverbial. Because of its adverbial use in this construction,

- the adjective must be in accusative case (which does not match the genitive case of “suspension”)
- the adjective’s default gender must be masculine (which matches “suspension” only by accident in this example – the adjective would remain masculine, even if the modified noun were feminine)
- the adjective must be singular, regardless of the number of the modified noun
- the adjective must be indefinite (which does not match the definite modified noun)

NB: there should be also an ADJP example here to show the difference

```
(S (VP Euwqiba
      (NP-SBJ-2 Al+lAEibu )
      (NP-OBJ-2 * )
      (PP bi
        (NP (NP Al+<iyqaAfi)
          (PP li
            (NP (ADJP >akovar+a
              (PP min
                (NP mubArAt+K))))))))))
```

عوقب اللاعب بالإيقاف لأكثر من مباراة

Euwqiba Al+lAEib+u bi+Al+<iyqAf+i li+>akovar+a min mubArAt+K
 Was punished + the+player + with+the+suspension + for+more + of+game
The player was punished with a suspension of more than one game

```
(S (VP Euwqiba
      (NP-SBJ-2 Al+lAEibu )
      (NP-OBJ-2 * )
      (PP bi
        (NP (NP Al+<iyqaAfi)
          (PP li
            (NP mubArAt+K
              aw
              akovar+a))))))
```

عوقب اللاعب بالإيقاف لمباراة أو أكثر

Euwqiba Al+lAEib+u bi+Al+<iyqAf+i li+mubArAt+K >aw >akovar+a
 was punished + the player + with+suspension + for+game + or + more
The player was punished by suspension one game or more

3.9 Adverbial pre-modifiers

NB: if we want this to be a whole separate section like this, I think it will need some text explaining why it is important and different, etc.

Example of an ADVP premodifying an NP:

```
(S (VP جاء
      (NP-SBJ (NP الجميع)
        (NP (ADVP (PRT ل)
          سيما
          الشبان))))
```

جاء الجميع لسيما الشبان

NB: need transliteration, also in the tree
 Came everybody especially the youth
Everybody especially young people came

Example of an NP-ADV premodifying a PP:

(S (VP (PRT -qad::قد::[has/have])
 >avobat+a::أُثْبِتَ::ascertain/establish+he/it_[verb]
 (NP-SBJ *)
 (NP-ADV sAbiq+AF::سابقاً::formerly/earlier+[acc.indef.])
 (NP-OBJ jadAr+at+a-::جَدَارَتَ::worthiness/aptitude/merit+
 [fem.sg.]+[def.acc.]
 (NP -hu::هُ::his/its))
 ,::,::nogloss
 (NAC wa-::وَ::and
 (PP (NP-ADV -xuSuwS+AF::خُصُوصاً::especially/in_particular+[acc.indef.])
 bi-::بِ-::by/with
 (NP -<izA'+i::إِزاءِ::towards/facing/regarding/
 vis--vis+[def.gen.]
 (NP Al+muEAriD+iyna::المُعَارِضِينَ::the+opponent/adversary/
 opposition+[masc.pl.gen.])))))))
 قد أثبت سابقاً جدارته وخصوصاً بإزاء المعارضين
 qad + >avobat+a + sAbiq+AF + jadArat+a+hu + wa+xuSuwS+AF + bi+<izA'+i +
 Al+muEAriDiyna
 has + establish + earlier + wptitude+his + and+especially + with+facing + the+opponents
Earlier he has established his aptitude especially in facing the opponents

Example of a PP premodifying a PP:

(S (VP نرثلوا
 (NP-SBJ *)
 (NP-OBJ القرآن
 (NAC و
 (PP (PP ب
 (NP خاصة)
 في
 (NP صلات
 (NP لنا))))))))

نرثلوا القرآن وبخاصة في صلاتنا

NB: need transliteration, also in the tree

Chant we the Quran and with specifically in prayer our

We read the Quran especially in our prayer.

NB: also the above tree cannot be correct – it has a PP coordinated with an NP. Should that be UCP, or is one of the nodes incorrect?

Example of an NP-ADV premodifying an ADJP:

(S (VP (PRT لم)
 تكن
 (ADVP-PRD هناك)
 (NP-SBJ (NP أي
 (NP محظورات)
 (ADJP (ADJP (NP-ADV سواء
 (سياسية)
 أو
 (ADJP دينية)
 أو
 (ADJP (جنسية))))))))

لم تكن هناك أي محظورات سواء سياسية أو دينية أو جنسية

NB: need transliteration, also in the tree

Not was there any banned whether political or religious or sexual

There are no taboos whether political, religious or sexual

NB: When sawA' is followed by coordinated Ss, it is annotated as NP-PRD and what comes after is SBAR-SBJ.

(S (VP (PRT س)
 أرحل
 (NP-SBJ *)
 (S-ADV (NP-PRD سواء)
 (SBAR-SBJ *0*
 (S (S (VP قبلت
 (NP-SBJ *)))
 أم
 (S (VP رفضت
 (NP-SBJ *))))))))))

سأرحل سواء قبلت أم رفضت

NB: need transliteration, also in the tree

Will leave I equality accept you or refuse you

I will leave whether you accept or you refuse

3.10 NP and idafa rules, the quick guide

- Possessive pronouns are ALWAYS idafa, always an NP complement to the noun they modify
- Numbers (cardinal) and quantifiers = idafa (or tamiyz), always have an NP complement (not flat, not adjunction)
- Adjunction – an NP adjoined to NP
 - Appositives
 - Dates
- Personal titles are flat – (NP President Bush)
 - See section????? for a partial list of personal titles
 - Coordinated titles need the NP for coordination – for now, that will look the same as adjunction (but is not intended to be read as adjunction)
 (NP (NP President and Commander-in-Chief)
 (NP Bush))
- Proper names are flat – (NP John Jacob Smith)

- False idafa will *always* be an ADJP
- False idafa in a noun position = include both the NP and the ADJP
 - → (NP (ADJP adj (NP noun-gen)))
 - Subject of the sentence, direct object of the verb, object of a preposition
- Modifiers that come at the end of the idafa (ADJP, PP, SBAR) should be adjoined to the NP that they belong with
 - WATCH OUT if the idafa is a quantification or number
 - If the two nouns in the idafa are in a whole-part relationship, the post-modifier will almost always go with the *whole* (the first NP)

4 Verb Phrase Structure

(NB: The VP may be the only child of the S, if nothing precedes the verb.)

As in the Penn English Treebank, the distinction between arguments and adjuncts of the verb or verb phrase is made through the use of functional dashtags rather than with a structural difference. Both arguments and adjuncts are children of the VP node. No structural distinction is made to represent any interpretive differences between VP-level modification and S-level modification. All constituents that appear before the verb are children of S and sisters of VP; all constituents that appear after the verb are children of VP.

ARGUMENTS of the verb are: NP-SBJ, NP-OBJ, NP-DTV, SBAR (no dashtag or -NOM-SBJ/OBJ/DTV), S (no dashtag or -NOM-SBJ/OBJ/DTV), PP-DTV, PP-CLR (closely/clearly related – a PP argument of the verb).

ADJUNCTS are: any XP with any other adverbial dashtag, PP (without dashtag), ADVP (without dashtag).

In this example, the NP-SBJ is the subject, NP-OBJ is the object of the verb, and NP-TMP is an adverbial (temporal) NP:

```
(S wa-::و-::and
  (VP -gAdar+a::غادر::leave/depart+he/it_[verb]
    (NP-SBJ Al+ra}iys+u::الرئيس::president/head/chairman+[def.nom.]
      (NP-OBJ Al+qAhir+ap+a::القاهرة::the+Cairo+[fem.sg.]+[def.acc.] )
      (NP-TMP Eaqiba::عقب::following/subsequent_to
        (NP (NP ziyAr+ap+K::زيارة::visit+[fem.sg.]+[indef.gen.] )
          (PP-LOC li-::ل-::for/to
            (NP -miSor+a::-مصر::Egypt+[indef.gen.] )))))
        وغادر الرئيس القاهرة عقب زيارة لمصر
  wa+gAdar+a + Al+ra}iys+u + Al+qAhir+ap+a + Eaqiba + ziyAr+ap+K + li+miSor+a
  and+left + the+president + the+Cairo + following + visit + to+Egypt
  the president left Cairo after visiting Egypt
```


4.1 Classification of verbs

4.1.1 Inflectional verbs

Inflectional verbs are the verbs that inflect for tense, aspect and mood. These are the most frequent Arabic verbs.

4.1.1.1 Event verbs

These are the verbs that have the meaning of an event related to a certain tense. They have an argument structure requiring agent and/or patient roles to be filled (by their subjects and/or objects).

The event verbs are categorized below according to their valence: intransitive, transitive with preposition, transitive, ditransitive, and clausal complements.

4.1.1.1.1 Intransitive verbs

Intransitive verbs are verbs that have as their only argument a subject.

```
(S (wa-::and
  (VP (PRT قد::-qad::[has/have])
    (tafAqam+at::become_serious/become_critical+it/they/she_[verb]
      (NP-SBJ الازمة::Al+>azom+ap+u::the+crisis+[fem.sg.]+[def.nom.])))
  وقد تفاقت الازمة
```

wa+qad + tafAqamat + Al+>azom+ap+u
and+has + become critical + the+crisis
The crisis has become critical

They can be modified by any modifier that can modify verbs.

```
(S (wa-::and
  (VP (PRT قد::-qad::[has/have])
    (tafAqam+at::become_serious/become_critical+it/they/she_[verb]
      (NP-SBJ الازمة::Al+>azom+ap+u::the+crisis+[fem.sg.]+[def.nom.]
        (NP-ADV تفاقمًا::tafAqum+AF::increase+[mas.sg.]+[indef.acc]
          (xaTiyar+AF::dangerous+[mas.sg.]+[indef.acc])
        (PP-LOC في::fiy::in
          (NP لبنان::lubonAn::libanon))))))
  وقد تفاقت الازمة تفاقمًا خطيراً في لبنان
```

wa+qad + tafAqamat + Al+>azom+ap+u + tafAqum+AF + xaTiyar+AF + fiy + lubonAn
and+has + become critical + the+crisis + increase + dangerous + in + Lebanon
The crisis has increased critically in Lebanon

Some verbs of this category can have a PP-CLR as a second argument after the subject and they are considered transitive with preposition (متعدّ بحرف/mutaEad~K bi Harf) – but in this case they

are considered as different verb with another meaning (see following section ??????????
1.1.1.2.2.)

Example of jA'a intransitive (come):

(S (VP jA'a جاء
(NP-SBJ Alrajulu الرَّجُلُ)))

جاء الرجل

jA'+a + Al+rajul+u
came + the+man
The man came

Example of jA'a transitive with preposition (bring):

(S (VP jA'a جاء
(NP-SBJ Alrajulu الرَّجُلُ)
(PP-CLR bi ب
(NP Aiboni ابْنِ
(NP hi هِ)))))

جاء الرجل بابنه

jA'+a + Al+rajul+u + bi+Aibon+i+hi
brought + the+man + with+son+his
The man brought his son

See section ????? on passives for a note about the two patterns tafaE~ala and tafAEala that are intransitive and active, but with a syntactic subject that is a patient (cf. in English *The window broke* and ??????????????).

4.1.1.1.2 Transitive with preposition

These are verbs that have as second argument a prepositional phrase (PP-CLR).

(S (VP Eavarat عَثَرَتْ
(NP-SBJ AlfatApu الفتاةُ)
(PP-CLR EalaY عَلَى
(NP kanozK كَنْزٍ))))

عَثَرَتْ الفتاة على كنز

Eavarat + Al+fatAp+u + EalaY + kanoz+K
found + the+girl + on + treasure
The girl found a treasure

4.1.1.1.3 Transitive verbs

These are transitive verbs that have a direct object noun phrase.

```
(S (NP-TPC-1 -TarAbulus:::طرابلس::Tripoli_[Libya] )
  (VP ta+notaqid+u:::تنتقد::it/they/she+criticize+[ind.]
    (NP-SBJ-1 *T*)
    (NP-OBJ -TarAbulus:::طرابلس::Tripoli_[Libya] )))
```

طرابلس تنتقد طرابلس

TarAbulus + tanotaqid+u + TarAbulus
 Tripoli + criticize + Tripoli
Tripoli criticizes Tripoli

They can have a PP-CLR argument in addition to the direct object.

```
(S (VP Ait~ahamat اتَّهَمَتْ
  (NP-SBJ AlwizArapu الوزارة )
  (NP-OBJ AlmuwaZ~afa الموظف )
  (PP-CLR bi ب
    (NP AltaqoSiyri التقصير ) ) ) )
```

اتهمت الوزارة الموظف بالتقصير

<it~ahamat + Al+wizArap+u + Al+muwaZ~af+a + bi+AltaqoSiyri+i
 Accuse + the+ministry +the+employee + with+the+negligence
The ministry accused the employee of negligence

The PP-CLR argument should be obligatorily a prepositional phrase.

If it is in alternation with a possible noun phrase argument, the dashtag should be –DTV (see section ?????? below on ditransitive verbs).

Some verbs are transitive because of their morphology. Two patterns, >afoEal and faE~al, add an argument to the basic form of the verb. If the basic verb is intransitive, these two patterns make a verb transitive. (See section ?????? for these patterns with basically transitive verbs.)

4.1.1.1.4 Ditransitive verbs

Ditransitive (**Di**T**ransiti****Ve**) verbs are verbs that have two complements in addition to the subject:

- Direct object (marked –OBJ). The direct object should be the patient role.
- Indirect object (marked –DTV, can be NP-DTV or PP-DTV). In Arabic, the dative shift is not always with the recipient.

For verbs of giving, when the indirect object is a noun phrase, the indirect object normally precedes the direct object. This is true whether the indirect object is a full noun phrase or a clitic pronoun. For these verbs, when the indirect object is a prepositional phrase, the PP-DTV will normally follow the direct object.

Some verbs are ditransitive because of their lexical meaning of giving or taking (list from the ATB3 corpus).

Verbs of giving:

أعطى >aEoTaY give
 منح manaHa donate
 وهب wahaba bestow
 أهدى >ahoday give a present
 دأين dAyana be indebted to
 أودع >awoda3a deposit in
 استودع Aistawoda3a entrust to
 زوّد zaw~ada provide with
 ناشد nA\$ada appeal to
 سأل sa>ala ask from
 أولى >awolaY devote
 أبلغ >abolaga report
 شارك \$Araka share with
 بادل bAdala exchange

Verbs of taking:

سلب salaba rob
 حرم Harama deprive
 جنب Jan~aba help avoid

Some verbs are ditransitive because of their morphology. Two patterns, >afoEal and faE~al, add an argument to the basic form of the verb. If the basic verb is already transitive, these two patterns make a verb ditransitive. (See section ????? for these patterns with basically intransitive verbs.)

nasiya (transitive) → >anosaY (ditransitive)
 qasama (transitive) → qas~ama (ditransitive)

(S (VP -yu+qas~im+u:: -يُقَسِّمُ::he/it+divide/partition/distribute+[ind.]
 (NP-SBJ Al+mu&al~if+u:: الْمُؤَلِّفُ::the+author/composer+[def.nom.])
 (NP-OBJ Al+kitAb+a:: الْكِتَابُ::the+book+[def.acc.])
 (NP-DTV >arobaE+ap+a:: أَرْبَعَةٌ::four+[fem.sg.]+[def.acc.]
 (NP >aqosAm+K:: أَقْسَامُ::portions/sections+[indef.gen.])))
 يقسم المؤلف الكتاب أربعة أقسام

yuqas~im+u + Al+mu&al~if+u + Al+kitAb+a + >arobaEap+a + >aqosAm+K

divide + the+author + the+book + four + sections

The author divides the book into four sections

When both complements of the ditransitive are bound pronouns, the direct object pronoun is cliticized to the dummy particle <iy~A-.

(S (VP Eaw~ad+at-::عَوَّدَتْ-::accustom/condition/habituate+it/they/she_[verb]
 (NP-DTV -nA::نا::us)
 (NP-OBJ <iy~A-::إِيَّا-::to/for/[accus.]
 -hA::ها::it/them/her)
 (NP-SBJ Al+jumohuwriy~+ap+u::الْجُمْهُورِيَّةُ::the+republic+[fem.sg.]+[def.nom.]
 Al+<isolAmiy~+ap+u::الإِسْلَامِيَّةُ::the+Islamic/Islamist
 +[fem.sg.]+[def.nom.])
 Al+<iyrAniy~+ap+u::الْإِيرَانِيَّةُ::the+Iranian
 +[fem.sg.]+[def.nom.])))
 عَوَّدَتْنَا إِيَّاهَا الْجُمْهُورِيَّةُ الْإِسْلَامِيَّةُ الْإِيرَانِيَّةُ
 Eaw~adat+nA + <iy~A+hA + Al+jumohuwriy~ap+u + Al+<isolAmiy~ap+u +
 Al+<iyrAniy~ap+u
 accustomed+us + it+itself + the+republic + the+Islamic + the+Iranian
The Islamic Republic of Iran accustomed us to it

See section ????? on passives for a note about the two patterns tafaE~ala and tafAEala that are transitive and active, but with a syntactic subject that is a patient (cf. in English *The window broke*). These patterns remove one argument role from the argument structure of the basic verb.

See section ??? for verbs with PP-CLR in addition to other complements.

4.1.1.1.5 Event verbs with clausal complements

Clausal complements are S or SBAR arguments that follow the verb.

(S (VP -ta+Eotabir+u-::تَعْتَبِرُ-::it/they/she+consider/regard/believe+[ind.]
 (NP-SBJ -TarAbulus::طَرَابُلُس::the issue)
 (S (NP-SBJ -AlqaDiy~ap+a::القَضِيَّةُ::Tripoli_[Libya])
 (ADJP-PRD siyAsiy~ap+F::سِيَاسِيَّةُ::political+[acc.indef.])))
 تَعْتَبِرُ طَرَابُلُسُ + الْقَضِيَّةُ + سِيَاسِيَّةُ
 ta+Eotabir+u + TarAbulus + Al+qaDiy~ap+a + siyAsiy~ap+F
 consider + Tripoli + the+ issue+ political
Tripoli considers the issue political

See also section ????? for serial verbs, as non-event verbs with clausal complements.

4.1.1.1.5.1 Verbs of speech

Verbs of speech take a clausal complement that can be either an SBAR or and S.

If the speech is indirect, the complement must be an SBAR. Usually this will be with the complementizers >an~a. If there is no overt complementizer heading the complement clause in indirect speech, then it should be annotated with a *0* null complementizer. Most complements of verbs of speech should be treated as SBARs/indirect speech unless there is evidence in the text that the speech is direct.

If the speech is direct (quoted), the complement should be an S (with no complementizers). This occurs especially with the verb *qala*. Often there will be quotation marks that show that the complement is direct speech.

If the complementizer is spelled as **ان**, it could be either <in~a (i.e., S) or >an~a (i.e., SBAR) depending on the clues: punctuation clues, direct speech, etc.

```
(S (VP قلت
  (NP-SBJ *)
  (S (VP إِنَّ
    (NP-SBJ-1 ك )
    (S (VP (PRT س )
      تستعيد
      (NP-SBJ-1 *)
      عافيت
      (NP-OBJ عافيت
        (NP (ك) ) ) ) ) ) ) ) ) ) )
```

قلت : "انك ستستعيد عافيتك"

NB: need transliteration, also in the tree

said I : “ indeed will regain health your”

I said: “Indeed, you’ll regain your health”

```
(S (VP قال
  (NP-SBJ *)
  (SBAR ان
    (S (NP-TPC-1 ني
      (VP (PRT لم
        أستدع
        (NP-SBJ-1 *T*)
        (NP-OBJ له ) ) ) ) ) ) ) ) )
```

قال انني لم أستدعه

NB: need transliteration, also in the tree

Said he that I not invited him

He said that I didn’t invite him.

4.1.1.1.5.2 Verbs of judgment

These verbs in Arabic take an S clause as complement, which consists either of an equational sentence or a verbal predication with the semantic object of main verb becoming the subject of the complement clause and always being the first element of the clausal complement.

The test for distinguishing between verbs of judgment and verbs of labeling is that if an SBAR >an~a clause can be substituted for the complement clause without any change in meaning, then the verb is a verb of judgment. If the SBAR >an~a cannot be substituted, then the verb is not a verb of judgment. If the >an~a is actually present in the sentence with the verb of judgment, then the complement clause is an SBAR (otherwise the complement clause is an S).

Examples of common verbs of judgment:

اعتقد {iEtaqada *believe/think*

اعتبر {iEtabara *consider*

ظن Zan~a *presume*

عرف Earifa *know*

4.1.1.1.5.3 حسب Hasiba think of something as

خال xAla *see as/imagine*

علم Ealima *know as*

دری daraY *know as*

زعم *assume that*

توهم *assume wrongly that*

وجد wajada *find*

رأى ra>Y *see as*

NB: وجد wajada *find* and رأى ra>Y *see* have two different argumentation structures depending on the meaning of the verb in the context:

- when they are verbs of physical perception, they can get an S-ADV or S-MNR rather than an S complement

(S (VP ra>ayotu رَأَيْتُ
 (NP-SBJ *)
 (NP-OBJ-1 AlTifola الطفل)
 (S-ADV (VP yafotaHu يَفْتَحُ
 (NP-SBJ-1 *)
 (NP-OBJ Eayonayo عَيْنَيْ
 (NP hi هـ))
 (PP-MNR bi بـ
 (NP buT'K بَطء ()))))))

رأيت الطفل يفتح عينيه ببطء

ra>ayt+u + Al+Tifol+a + yafotaH+u + Eayon+ayo+hi + bi+buTo'+K

saw + the+baby + open + eye+two+his + with+slowness

I saw the baby opening his eyes slowly

- when they are verbs of mental judgement, they should have a sentential complement

(S (VP wajadat وَجَدَتْ
 (NP-SBJ Almaro>apu الْمَرْأَةُ)
 (S (NP-SBJ AlEayo\$a الْعَيْشُ)
 (NP-TMP baEoda بَعْدَ
 (NP wafApi وَفَاةِ
 (NP zawoji زَوْجِ
 (NP hA هَا))))
 (ADJP-PRD qAsiyAF قَاسِيَةً))))

وجدت المرأة العيش بعد وفاة زوجها قاسية

wajadat + Al+maro>ap+u + Al+Eayo\$a + baEoda + wafAp+i + zawoj+i+hA + qAsiy+AF
 found + the+woman + the+living + after + death + husband+her + hard
The woman found that living is hard after her husband's death

The S complement of these verbs can be an equational sentence –

(S (VP Hasiba حَسِبَ
 (NP-SBJ AlfataY الْفَتَى)
 (S (NP-SBJ AlHayApa الْحَيَاةُ)
 (ADJP-PRD saholapF سَهْلَةً))))

حسب الفتى الحياة سهلة

Hasib+a + Al+fataY + Al+HayAp+a + saholap+F
 thought + the+lad + the+life + easy
The lad thought that life was easy

– or verbal sentence with an accusative subject topicalized.

(S (VP Zananotu ظَنَنْتُ
 (NP-SBJ *)
 (S (NP-TPC-1 AlsamA'a السَّمَاءُ)
 (VP tumoTiru تُفَطِّرُ
 (NP-SBJ-1 *T*))
 (NP-OBJ *ahabA ذَهَبًا
 wa وَ
 fiD~apF فِضَّةً)))))

ظننت السماء تمطر ذهباً وفضة

Zananot+u + Al+samA'+a + tumoTir+u + *ahab+AF + wa+fiD~ap+F
 thought + the+sky + rain + gold + silver
I thought the sky rained gold and silver

4.1.1.1.5.4 Verbs of transformation and labeling verbs

Verbs of transformation and labeling verbs are annotated as having both a direct object and a clausal (S) complement. The empty subject of the S complement is coindexed with the direct object of the main verb.

The test for distinguishing between verbs of judgment and verbs of labeling is that if an SBAR >an~a clause can be substituted for the complement clause without any change in meaning, then

the verb is a verb of judgment. If the SBAR >an~a cannot be substituted, then the verb is not a verb of judgment.

A representative (but not exhaustive) list of verbs to be annotated in this way is below. In general, the interpretation of the verbs involves both a direct relationship between the verb and the direct object (*John appointed Mary*) and also a relationship between the verb and the clause (*John appointed Mary to be president*). Note that the clausal complement is always S in the Arabic Treebank, not S-CLR. (NB: This category of verbs in the Arabic Treebank combines in some sense the “ditransitive verbs” with clausal complements and the “verbs with secondary predicates” as this distinction is made in the English Treebank.)

```
(S (VP verb
  (NP-SBJ subject)
  (NP-OBJ-1 object)
  (S (NP-SBJ-1 *)
    (XP-PRD predicate))))
```

Contrary to verbs of judgment, verbs of transformation allow only an equational sentence complement, with the exception of jaEala.

Examples of common verbs of transformation and labeling:

حول Haw~ala *transform*

صَيَّر Say~ara *make*

اِتَّخَذَ {it~axa*a *take*

اِسْتَعْمَلَ {isotaEomala *use*

جَعَلَ jaEala *make*

سَمَّى sam~aY *call/nominate*

عَيَّن Eay~ana *appoint*

اِخْتَارَ {ixtaAra *choose*

اِنتَخَبَ {inotaxaba *elect*

أَبْقَى <aboqaY *let something be something*

عُدَّ Ead~a (used in the passive yuEad~u) *be counted among/considered as*

```
(S (VP اِتَّخَذَ
  (NP-SBJ*)
  (NP-OBJ-1 مُحَمَّدًا )
  (S (NP-SBJ-1 *)
    (NP-PRD صديقًا ) ) ) )
```

اِتَّخَذَ مُحَمَّدًا صديقًا

NB: need transliteration, also in the tree

Make he Mohamed a friend

He took Mohamed for a friend

This type of structure involves considering the NP-OBJ as being mino+hu with mino being a PRT.

```

(S (VP جعل
  (NP-SBJ*)
  (NP-OBJ-1 (PRT من )
    هـ)
  (S (NP-SBJ-1 *)
    (NP-PRD (مدير ) ) ) ) )

```

جعل منه مديرا

NB: need transliteration, also in the tree

Made he from him manager

He made a manager of him

Example when the object comes before the subject:

```

(S (VP جعل
  (NP-OBJ-1 هـ)
  (NP-SBJ الناس )
  (S (VP تتحوّل
    (NP-SBJ-1 *)
    (PP إلى
      (NP امرأة
        قاسية ) ) ) ) ) )

```

جعلها الناس تتحوّل إلى امرأة قاسية

NB: need transliteration, also in the tree

Made it her people transform into a woman hard

People made her change into a hard woman

4.1.1.1.5.5 Test for distinguishing S vs. PP-CLR with verbs of transformation

If the potential predicate of the complement clause is a PP, it should be tested to determine whether it is a real predicate (and therefore part of an S complement) or not a predicate (and therefore a PP-CLR). The test is to try the NP-OBJ of the verb as the subject of a simple equational sentence with the PP as its predicate. If the equational sentence works (in the same meaning as it would have as the complement of the labeling verb), then it should be annotated as above. If the equational sentence does not work, there is no S complement, and the PP should be annotated as a PP-CLR, as below.

These verbs can also appear in another construction, with a PP-CLR and without the complement clause, since the PP in this case is not a predicate.

Test for distinguishing verbs of judgment and verbs of transformation:

The test for distinguishing between verbs of judgment and verbs of labeling is that if an SBAR >an~a clause can be substituted for the complement clause without any change in meaning, then the verb is a verb of judgment. If the SBAR >an~a cannot be substituted, then the verb is not a verb of judgment.

Test for distinguishing verbs of labeling and Transformation that take NP-OBJ and PP-CLR:

If the potential predicate of the complement clause is a PP, it should be tested to determine whether it is a real predicate (and therefore part of an S complement) or not a predicate (and therefore a PP-CLR). The test is to try the NP-OBJ of the verb as the subject of a simple equational sentence with the PP as its predicate. If the equational sentence works (in the same meaning as it would have as the complement of the labeling verb), then it should be annotated as above. If the equational sentence does not work, there is no S complement, and the PP should be annotated as a PP-CLR, as below.

```
(S (VP وصفت
  (NP-SBJ الصحافة)
  (NP-OBJ الحملة الانتخابية)
  (PP-CLR ب
    (SBAR أن
      (S (NP-TPC-2 ها)
        (VP كانت
          (NP-SBJ-2 *T*)
          (ADJP-PRD هادئة))))))
```

وصفت الصحافة الحملة الانتخابية بأنها كانت هادئة

NB: need transliteration, also in the tree

Described the media the campaign electoral by that it was calm

The media described the electoral campaign as calm

Most of these verbs have the same behavior when they are regular verbs and when they are maSdar. However, وصف (waSafa describe) behaves as a labeling verb when it is a maSdar, but it takes a PP-CLR as a regular verb.

```
(PP ب
  (S-NOM (VP وصف
    (NP-SBJ *)
    (S (NP-SBJ ه)
      (NP-PRD مديرا))))
```

بوصفه مديرا

NB: need transliteration, also in the tree

With describing him a manager

Given that he's the manager

4.1.2 Non-event verbs

4.1.2.1 kAna and layosa

kAna and *layosa* function both as main verbs and as auxiliary verbs, depending on the syntactic context.

kAna and *layosa* are the only auxiliary verbs in the Arabic Treebank. An auxiliary verb in the Treebank has a plain VP complement. Any other verb that is followed by a second verb is analyzed as a verb with a sentential complement.

4.1.2.1.1 kAna

When *kAna* functions as the main verb in the clause, it has a subject and a predicate as arguments.

```
(S (VP kAn+a::كَانَ::be/was/were+he/it_[verb]
  (NP-SBJ jAn::جان::Jean/Jan
    gAnim::غانيم::Ghanim/Ghanem )
  (NP-PRD (NP yad+a-::يَدَ-::hand+[def.acc.]
    (NP -hu::هُ-::its/his ))
    (ADJP Al+yumonaY::الْيُمْنَى::the+right_hand/right_side/lucky ))))
كان جان غانم يده اليمنى
```

kAn+a + jAn + gAnim + yad+a+hu + Al+yumonaY
 was + Jean + Ghanim + hand+his + the+right
Jean Ghanim was his right hand

When *kAna* functions as an auxiliary verb, its only complement is a plain VP.

```
(S (VP -kAn+a::-كَانَ::be/was/were+he/it_[verb]
  (VP ya+taHad~av+u::يَتَحَدَّثُ::he/it+speak+[ind.]
    (NP-SBJ-1 *)
    (NP-OBJ-1 *)
    (PP bi-::بِ-::by/with
      (NP -wuDuwH+K::وُضُوحَ::clarity/plainness+[indef.gen.] ))))
كان يَتَحَدَّثُ بوضوح
```

kAn+a + ya+taHad~av +u + bi+wuDuwH+K
 was + speaking + with+clarity
He was speaking clearly

If there is a topicalized NP before the *kAna*, the trace follows the main verb. That is, the subject trace is in the VP that is the complement of *kAna*; it is not in the *kAna* VP.

(S (NP-TPC-1 milaf~+At+N::ملفًا::dossier/letter_file+[fem.pl.]+[indef.nom.])
 (VP kAn+at::كانت::be/was/were+it/they/she_[verb]
 (VP (PRT qad::قد::[has/have])
 futiH+at::فُتِحَتْ::be_opened/be_conquered+it/they/she_[verb]
 (NP-SBJ-1 *T*)
 (NP-OBJ-1 *)
 (PP-TMP fiy::في::in
 (NP Eahod+i::عَهْد::age/period+[def.gen.]
 (NP Al+ra}iys+i::الرئيس::the+president/head/
 chairman+[def.gen.])))))))
 ملفات كانت قد فتحت في عهد الرئيس
 milaf~At+N + kAnat + qad + futiHat + fiy + Eahod+i + Al+ra}iys+i
 files + were + has + be opened + in + reign + the+president
Files that were opened at the president's reign

If the overt subject NP appears between the kAna and the main verb, it is annotated as an NP child of the kAna VP (sister to both kAna and the complement VP), and it is coindexed with the subject trace that follows the main verb. This raised subject does not have the –TPC dashtag because that dashtag is reserved for NPs that are fronted to the beginning of the sentence.

(S (VP -kAn+a::كان::be/was/were+he/it_[verb]
 (NP-1 Al+HAdiv+u::الحدث::the+incident/event+[def.nom.])
 (VP HaSal+a::حصل::occur/happen/take_place+he/it_[verb]
 (NP-SBJ-1 *T*)
 (PP-TMP fiy::في::in
 (NP (ADJP Al+>aw~al+i::الأول::the+first+[def.gen.]
 (PP min::من::from
 (NP |*Ar+a::آذار::March+[indef.gen.]
 Al+mADiy::الماضي::the+past/bygone)))))))
 كان الحادث حصل في الأول من آذار الماضي
 kAn+a + Al+HAdiv+u + HaSal+a + fiy + Al+>aw~al+i + min + |*Ar+a + Al+mADiy
 was + the+incident + happen + in + the+first + March + the+last
The incident happened in the first of March the last

4.1.2.1.2 layosa

layosa has the same syntax as *kAna*, both as a main verb and as an auxiliary verb.

When *layosa* functions as the main verb in the clause, it has a subject and a predicate as arguments.

(S (VP ليس::laysa (be/was/ +he/it_[verb]
 (NP-SBJ الوضع::Al+waDE+u/ the+situation)
 (ADJP-PRD خطيرا::xaTiyru+AF::serious)))
 ليس الوضع خطيرا
 Layosa + AlwaDoE+u + xaTiyru+AF
 not + the+situation + dangerous
The situation is not serious

(S (VP ليست::layos+at::was not/+ it/they/she_[verb]
 (PP-PRD لي-::ل-::for/to
 (NP -say~id+i-:::-سيد-::Sir/Mr./Mister+[def.gen.]
 (NP -nA:::-نا:::our)))
 (NP-SBJ Aino\$igAl+At+N:::انشغالات::preoccupied/busy/
 occupation+[fem.pl.]+[indef.nom.]
 kaviyr+ap+N:::كثيرة::many/much/numerous+[fem.sg.]+[indef.nom.]
 (PP-TMP في::في::in
 (NP fator+ap+i:::فترة::phase/time_period/interval
 +[fem.sg.]+[def.gen.]
 (NP Al+>aEoyAd+i:::الأعياد::the+holidays/feasts
 +[def.gen.]))))
 ليست لسيدنا انشغالات كثيرة في فترة الأعياد
 layosat + li+say~id+i+nA + <ino\$igAlAt+N + kaviyr+ap+N + fiy + fator+ap+i +
 Al+>aEoyAd+i
 not + for+mister+our + occupation + much + in + interval + the+holidays
Our mister does not have many occupations during holiday intervals

When layosa functions as an auxiliary verb, its only complement is a plain VP.

(S (VP ليس-lays+a:::- /was/ +he/it_[verb]
 (VP يقيم::yu+qim+u he/it+live+[ind.]
 (NP-SBJ *)
 (PP-LOC في::في::in
 (NP القاهرةAlqAhirapi: Cairo))))
 ليس يقيم في القاهرة
 laysa + yuqim+u + fiy + Al+qAhirap+i
 not + live + in + the+Cairo
He does not live in Cairo

Note that layosa also has a non-verbal meaning – it can be used as negative particle when it has the meaning of *IA*.

(S (VP تعود::ta+Euwd+a:::it/they/she+return/go_back+[sub.]
 (NP-SBJ *)
 (NP-OBJ ادراج::>adorAj+u:::steps/stairs+[def.nom]
 (NP ها:::-hA:::its/their/her))
 (PP الى:::ilaY:::to/towards
 (NP (NP الاطفال:::Al+>aTofAl+i:::the+children/infants+[def.gen.]
 و:::wa:::and
 (NP (PRT ليس:::-layos+a:::not+[it])
 الشباب:::Al+\$abAb+u:::the+youth/youths+[def.nom.]))))
 تعود أدرجها الى الأطفال وليس الشباب
 taEuwd+u + >adorAj+u+hA + <ilaY + Al+>aTofAl+i + wa+layosa + Al+\$abAb+i
 return + steps+its + to + the+children + and+not + the+youths
Its steps return to the children rather than to the youths

4.1.2.2 Sisters of kAna

The verbs that are the sisters of kAna take a subject (usually NP-SBJ) and a predicate. The predicate is shown with the -PRD function tag, which is used with all non-verbal predicates: NP-PRD, ADJP-PRD, PP-PRD, etc.

```
(S (VP ظَلَّت Zal~at
    (NP-SBJ الأمطار Al>amoTAru)
    (ADJP-PRD غزيرة gaziyrapF)))
```

ظَلَّت الأمطار غزيرةً

Zal~at + Al>amoTAru + gaziyrap+F
remained + the+rain + heavy
The rain remained heavy

These verbs also can have an S complement instead of the non-verbal predicate. These complements are different from the complements of serial verbs because their subject does not have to be empty, and it is not obligatorily coindexed with the subject of the higher verb.

```
(S (VP ظَلَّت Zal~at
    (NP-SBJ-1 الأمطار Al>amoTAru)
    (S (VP تهطل tahoTulu
        (NP-SBJ-1 *))))))
```

ظَلَّت الأمطار تهطل

Zal~at + Al>amoTAru + tahoTul+u
remained + the+rain + fall
It rained continuously

```
(S (VP ظل Zal~a
    (NP-SBJ زيد ZayodN
    (S (NP-TPC-1 قلب qalobu
        (NP له hu))
        (VP يخفق yaxofuqu
        (NP-SBJ-1 *T*))))))
```

ظلَّ زيد قلبه يخفق

Zal~a + Zayod+N + qalob+u+hu + yaxofuq+u
remained + Zaid + heart+his + beat
Zaid's heart continued beating

4.1.2.2.1 Verbs of becoming أفعال الصيرورة

صار SAra become
أُمسى >amosaY become
أضحى >aDHaY become
غدا gadA become

(S (VP >aSbaHa أصبح
(NP-SBJ zayodN زَيْدُ)
(NP-PRD waziyrAF وَزِيرًا)))

أصبح زيد وزيرا

>aSbaH+a + zayod+N + waziyr+AF
became + Zayd + minister
Zaid became a minister

(S (VP >amosaY أمسى
(NP-SBJ zaydN زَيْدُ)
(ADJP-PRD xA}ifAF خائفاً)))

أمسى زيد خائفا

>amosaY + zayd+N + xA}if+AF
became + Zayd + afraid
Zaid became afraid

(S (VP >aDoHaY أضْحى
(NP-SBJ zaydN زَيْدُ)
(PP-PRD fiy في
(NP DiyqK ضيق))))

أضحى زيد في ضيق

>aDoHaY + zayd+N + fiy + Diyq+K
became + Zayd + in + trouble
Zaid became in trouble

(S (VP >aSbaH+a
(NP-SBJ-1 zayod+N)
(S (VP yugan~iy
(NP-SBJ-1 *)
(PP fiy
(NP Al+HafalAt+i
Al+xAS~ap+i))))))

أصبح زيد يغني في الحفلات الخاصة

>aSbaH+a + zayod+N + yugan~iy + fiy + Al+HafalAt+i + Al+xAS~ap+i
Became + zayd + sing + in + the+parties + the+privet
Zaid began to sing in privet parties

4.1.2.2.2 Verbs of duration or continuity أفعال الاستمرار

ظلّ Zal~a remain

بات bAta carry on

ما زال mA zAla be still

ما فتئ mA fati}a not stop

ما انفكّ mA {inofak~a remain

(S (VP (PRT mA ما)
 zAla زال
 (NP-SBJ zayodN زَيْدُ)
 (NP-PRD (NP >axAF أَخاً)
 (PP la لَ)
 (NP ka كَ)))))

ما زال زيد أخاك

mA+zAl+a + zayod+N + >ax+AF + la+ka
 not+gone + Zaid + brother + for+you
Zaid is still your brother

(S (VP (PRT mA)
 Ainfak~a ما انفك
 (NP-SBJ Alrajulu الرَّجُلُ)
 (ADJP-PRD mariyDAF مريضاً))))

ما انفك الرجل مريضاً

Zal~+a + Al+rajul+u + mariyD+AF
 Remain + the+man + sick
The man remained sick

NB: this transliteration does not match the tree, not sure which one needs to be fixed

(S (VP sayaZal~u سَيَظَلُّ
 (NP-SBJ zayodN زَيْدُ)
 (PP-PRD fiy في
 (NP AlHafalAti الحَفَلَاتِ
 AlxAS~pi الخاصة)))))
 (S (VP sayaZal~u سَيَظَلُّ
 (NP-SBJ-1 zayodN زَيْدُ)
 (S (VP yugan~iy يَغْنِي
 (NP-SBJ-1 *)
 (PP fiy في
 (NP AlHafalAti الحَفَلَاتِ
 AlxAS~pi الخاصة)))))))))))

سيظل زيد يغني في الحفلات الخاصة

sa+yaZal~+u + zayod+N + yugan~aiy + fiy + Al+HafalAt+i + Al+xAS~p+i
 Will+remain + zayd + singing + in + the+parties + the+private
Zaid will remain singing in privet parties

4.1.2.2.3 Verb of simultaneity فعل التوقيت

This verb can only be the verb in an SBAR headed by mA.

طفق Tafaqa
أنشأ >ano\$a>a
قام qAma

(S (VP >axa*a أخذ
(NP-SBJ-1 AlTiflu الطفل)
(S (VP yaDoHaku يَضْحَكُ)
(NP-SBJ-1 *)))))

أخذ الطفل يضحك

>axa*+a + Al+Tifl+u + yaDoHak+u
take + the+child + laugh
The child started laughing

(S (NP-TPC-2 العالم::Al+EAlam+a::the+world+[def.acc.])
(VP بدأ::bada>a::start/begin+he/it_[verb]
(NP-SBJ-2 *T*)
(S (VP يعي::ya+Eiy+-LRB-null-RRB::he/it+be_aware_of/
pay_attention_to+[ind.]
(NP-SBJ-2 *)
(NP-OBJ ذلك::*`lika::that))))))

العالم بدأ يعي ذلك

Al+EAlam+u + bada>a + yaEiy + *`lika
the+world + start + be aware + that
The world started to recognize it

NB: please note that with an overt subject, this subject must always be traced with the first verb and the second verb takes an empty subject . So, the example above is annotated correctly, but it would be wrong to annotate the example above with the *T* trace of the NP-TPC following the verb yaEiy.

4.1.2.3.2 Verbs meaning ‘be on the point of..’ >afoEAl Almuqaraba

كاد kAda (be on the point of being/doing something...)
أوشك >awo\$aka (same gloss)

(S (VP تكاد::ta+kAd+u::it/they/she+almost/hardly/no_sooner+[ind.]
(NP-SBJ-1 إسرائيل : Israel)
(S (VP نتهم::ta+t~ahim+u::it/they/she+accuse+[ind.]
(NP-SBJ-1 *)
(NP-OBJ (NP لنا::-nA::us)
(NP جميعا::jamiyE+AF::all/entirely/together))
(PP-CLR ب::bi-::with/by
(NP الارهاب::-Al+<irohAb+i::the+terrorism/
terrorizing+[def.gen.]))))))
تكاد إسرائيل نتهمنا جميعاً بالإرهاب

takAd+u + <isorA}iyl+a + tat~ahim+u+nA + jamiyE+AF + bi+Al+<irohAb+i
almost + Israel + accuse+us + all + with+the+terrorism
Israel accuses almost all of us of terrorism

When these verbs occur with S complements, the empty subject of the complement clause should be co-indexed if appropriate.

NB: kAda and >awo\$aka are more and more used with a subordinate clause with >an. In this case, they are not considered as serial verbs. Do not co-index the empty subject across and SBAR.

```
(S (VP كادت:::-kAd+at:::almost/hardly/no_sooner+it/they/she_[verb]
  (NP-SBJ المضادات:::Al+muDad~+At+u:::the+anti-/counter-/
    contra-+[fem.pl.]+[def.nom.] )
  (SBAR ان:::>an:::to
    (S (VP تصيب:::tu+Siyb+a-:::it/they/she+strike/afflict+[sub.]
      (NP-SBJ *)
      (NP-OBJ ها:::-hA:::it/them/her))))))
```

كادت المضادات أن تصيبها

kAdat + Al+muDad~+At+u + >an + tuSiyb+u+hA

almost + the+anti-missiles + to + strike+it

The anti-missiles were about to strike it

>awo\$aka can also take a PP-CLR as an argument, which must be EalaY followed by a noun masdar or SBAR with >an.

```
(S (VP أوشك:::->awo$ak+a:::be_on_the_verge_of/be_about_to/almost+he/it_[verb]
  (NP-SBJ *)
  (PP-CLR على:::EalaY:::on/above
    (NP إحكام:::<iHokAm+i:::accuracy/precision+[def.gen.]
      (NP (NP قبضت:::qaboD+at+i-:::grip/seizure/fistful+[fem.sg.]
        +[def.gen.]
        (NP هـ:::-hi:::his/its))
      (PP على:::EalaY:::on/above
        (NP كل:::kul~+i:::every/all/each_one+[def.gen.]
          (NP الملف:::Al+milaf~+i:::the+file/dossier
            +[def.gen.]
            الفلسطين:::Al+filasoTiyniy~+i:::the
              +Palestinian+[def.gen.]))))))))
  أوشك على إحكام قبضته على كل الملف الفلسطيني
```

>aw\$ak+a + EalaY + <iHokAm + qaboD+at+i+hi + EalaY + kul~+i + Al+milaf~+i +

Al+filasoTiyniy~+i

be about to + on + precision + grip+his + on + all +the+file + the+Palestinian

He is about to control the whole Palestinian file

The main difference between these verbs and the sisters of kAna is that these verbs require an S or an SBAR complement, so they cannot have a non-verbal predicate (NP-PRD, ADJP-PRD, PP-PRD, etc.). (Sisters of kAna and pseudo-verbs can have either non-verbal predicates or clausal complements – see section 4.1.2.2. ??????????)

4.1.3 Non-inflectional verbs

The non-inflectional verbs are verbs that don't inflect in order to express a different tense or to agree in gender and number with their subjects. They always have the same form, and they are

classified as verbs because they require a nominative subject. The non-inflectional verbs are usually intransitive, and they are relatively infrequent verbs.

Complete list of non-inflectional verbs:

بئس/bi{osa (how evil)
 نعم/niEoma (how good)
 ينبغي yanobagiy must
 تعال taEAla come
 أف >uf~F fed up
 هيا hay~A let's
 هيهات hayohAta how far
 شتان \$at~Ana how different
 سرعان suroEAna promptly

It is common to have an NP-ADV (tamyiyz) directly after بئس/bi{osa and نعم/niEoma preceding the nominative subject, indicating in what way the subject is “good” or “bad” (“good as a friend” or “bad as a man”, etc.).

(S (VP bi}osa بئس
 (NP-ADV SadiyqAF صديقاً
 (NP-SBJ zaydN زيد)))

بئس صديقاً زيد

bi}osa + Sadiyq+AF + zayd+N
 how evil + friend + Zaid
How evil to have Zaid as a friend

حبذا/Hab~*A (how nice)

The subject of حبذا/Hab~*A is usually (specially in MSA) an SBAR with the subordinating conjunction لو.

(S (VP Hab~a*A حبذا
 (SBAR-SBJ law لو
 (S (VP samiEota سمعت
 (NP-SBJ *)
 (NP-OBJ baEoDa بعض
 (NP (NP ha*ihi هذه)
 (NP AlnatA}iji النتائج))))))))
 حبذا لو سمعت بعض هذه النتائج

Hab~a*A + law + samiEot+a + baEoD+a + h~*ihi + Al+natA}ij+i
 how nice + if + heard + some + these + the+results
It would be really nice to hear some of these results

شتان/\$at~Ana

hayohAta + >an + yaloHaquwA + biy
 how far + to + reach + me
How far would it be for them to reach me!

سرعان/suroEAna

This verb is always followed by a subject SBAR with the subordinating conjunction mA.

```
(S (VP سرعان::suroEAn+a::quick/prompt+[def.acc.]
  (SBAR-SBJ ما::mA::that
    (S (VP تحول::taHaw~al+a:: be_transformed+he/it_[verb]
      (NP-SBJ واحد::>aHad+u-::one+[def.nom]
        (NP هما::-humA::their))
      (NP-ADV ضد::Did~a::against
        (NP الآخر::Al+|xar+i::the+other/another/
          latest+[def.gen.]))))))
سرعان ما تحول أحدهما ضد الآخر
```

suroEAn+a + mA + taHaw~al+a + >aHad+u+humA + Did~a + Al+|xar+i
 prompt + that + be transformed + one+their + against + the+other
Promptly, each one of them changed against the other

أفعل >afoEala (verb of exclamation) is a non-inflectional verb, and it is used in two patterns : mA >afoEala and .../ >afoEil bi. The second pattern is no longer used in Modern Standard Arabic. However, mA >afoEala is still used and should be analyzed as follows:

```
(SBAR (WHNP-1 ما mA )
  (S (VP >aHosana أحسن
    (NP-SBJ-1 *T* )
    (NP-OBJ AlsamA'a السماء ) ) ) )
```

ما أحسن السماء

mA + >aHosan+a + Al+samA'+a
 how + beautiful + the+sky
How beautiful is the sky!

حسبُ Hasobu

is a non inflectional verb and is annotated as follows:

```
(S (VP Hasobu حَسْبُ
  (NP-OBJ nA نا )
  (NP-SBJ Allahu اللّهُ ) ) )
```

حسبنا الله

Hasobu + nA + Allahu
 be enough (for) + us + God
God is enough for us

It can also happen in the imperative form with an empty subject:

(S (VP Hasobu حَسْبُ
(NP-SBJ *)
(NP-OBJ ki كِ)))

حسبك

Hasobu + ki

Be enough (for) he/it + you (fem.sg.)

enough!

Stop!

You can't go any further!

أف >uf~K

(S (VP أف
(NP-SBJ *)
(PP-CLR min من
(NP ka كِ))))

أف منك

NB: need transliteration, also in the tree

Fed I up with you

I'm fed up with you

4.1.4 Pseudo-verbs: <in~a and its sisters

These words are classified in the traditional Arabic grammar as pseudo- or quasi-verbs (الحروف المشبهة بالأفعال / AlHuruwfu Almu\$ab~hap bi Al>afoEAL), although they are at the same time considered particles. But because of their verbal interpretation and because they require an argument structure similar to the argumentation of the weak verbs (see section 1.1.4.), we annotate them in the ATB as heading a VP and requiring an accusative subject and a nominative non-verbal predicate (NP-PRD, ADJP-PRD, PP-PRD, etc.) or a sentential complement. If the S complement includes an empty subject that is co-referent with an NP higher in the sentence, the NP and the empty subject should be co-indexed.

إنَّ <in~a
لَعَلَّ laEal~a
عَلَّ Eal~a
كَأَنَّ ka>an~a
ليت layota
لَكِنَّ lakin~a
لا النافية للجنس lA

(S (VP إن ::-<in~a::indeed
(NP-SBJ (NP هذا ::h`*A::this_[masc.sg.]
(NP التحول ::Al+taHaw~ul+a::the+change/transformation+[def.acc.]))
(ADJP-PRD آني ::|niy~+N::actual/present/simultaneous+[indef.nom.])))

إن هذا التحول آني

<in~a + h`*A + Al+taHaw~ul+a + lniy~+N

indeed + this + the+change + present

This change is present

(S (VP layota لَيْتَ
 (NP-SBJ-1 Al\$abAba الشَّبَابُ)
 (S (VP yaEuwdu يَنْعُودُ)
 (NP-SBJ-1 *)
 (NP-TMP yawomAF يَوْمًا)))))

ليت الشباب يعود يوماً

layota + Al\$abAb+a + yaEuwd+u + yawom+AF
 hope + the+youth + returns one day
I wish youth will return someday

(S (NP-TPC-1 (NP هذه)
 (NP الفرصة))
 (VP لعلَّ
 (NP-SBJ (NP ها)
 (NP-1 *T*))
 (ADJP-PRD (الأخيرة)))))

هذه الفرصة لعلها الأخيرة

NB: need transliteration, also in the tree

This chance perhaps it (is) the last
This chance may be the last.

NB: Most of these pseudo-verbs were treated as SBAR complementizers in previous versions of the ATB.

EassaY, which is originally a serial verb that requires a nominative subject, can also be treated as a pseudo-verb with the meaning of laEal~ (hope) and be followed by an accusative subject.

(S (VP EasA عساه
 (NP-SBJ hu هُ)
 (S (VP ya>atiy يَأْتِي
 (NP-SBJ *)))))

عساه يأتي

EasA+hu + ya>atiy
 Hope+him + came
I hope he will come

4.1.4.1 1A as a pseudo-verb (1A bud~a, etc.), -SBJ only no -PRD

“1A” in this context will now be treated as a sister of <in~a, and as such as a pseudo-verb. However, it is exceptional among pseudo verbs in that it heads a VP and takes a subject, but does not have a predicate complement.

```
(S (VP لا 1A
  (NP-SBJ غالبwinner)))
```

لا غالب

NB: need transliteration, also in the tree

Exist not a winner

A winner doesn't exist

```
(S (VP لا 1A
  (NP-SBJ رجل man)
  (PP-LOC في in
    (NP الدار the house))))
```

لا رجل في الدار

NB: need transliteration, also in the tree

Exist not a man in the house

There is no man in the house

```
(S (VP لا
  (NP-SBJ (NP أحد )
    (SBAR (WHNP-1 *0*)
      (S (VP يستطيع
        (NP-SBJ-1 *T*)
        (NP-OBJ مقاومه
          (NP له ))))))))
```

لا أحد يستطيع مقاومته

NB: need transliteration, also in the tree

Exist not one (that) can resist him

No one can resist him

```
(S (VP لا 1A
  (NP-SBJ (NP بدّ bud~a)
    (PP من min
      (NP المحاولة))))))
```

لا بدّ من المحاولة

NB: need transliteration, also in the tree

No way out from the trying

There is no way out of trying

NB: is the Arabic string correct? It looks like it's missing something

```
(S (VP لا 1A
  (NP-SBJ بدّ bud~a
    (SBAR أن >an
      (S (VP يستجيب
        (NP-SBJ القدر ))))))))
```

لا بدّ أن يستجيب القدر

NB: need transliteration, also in the tree

No way out that respond the fate

There is no way that fate won't respond

```

(S (VP lA لا
  (NP-SBJ (NP bud~a بُدْ )
    (PP min مِنْ
      (SBAR >an أَنْ
        (S (VP taroHala تَرْحَلْ
          (NP-SBJ * ) ) ) ) ) ) ) )

```

لا بد من أن ترحل

lA+bud~a + min + >an + taroHal+a
 no+way out + from + that + leave
You must leave

```

(S (VP lA لا
  (NP-SBJ (NP $ak~a شَكْ )
    (PP-PRD fiy فِي
      (NP Sidoqi صِدْقِ
        (NP hi هِ ) ) ) ) ) )

```

لا شك في صدقه

lA \$ak~a fiy Sidoqihi
 no+doubt + in + truth+his
There is no doubt that he says the truth

Note that “lA” in this context has been annotated in several ways in previous ATB corpora, but should now always be annotated as a pseudo-verb.

4.1.4.2 lakin~a usually heads an S-ADV, with exceptions

NB: Sondos and Dalila say that this is ALWAYS an independent clause, never s-adv. And they have been correcting it this way in pass2. What to do about this as a policy change? What to do about revising the existing data?

NB: is this still an issue?? Mohamed?

The pseudo-verb lakin~a will usually head an S-ADV that modifies a VP. The primary issue with annotating lakin~a is normally determining which VP it properly modifies.

Since the clause headed by lakin~a is usually an adverbial modification of a preceding VP, when the discourse “wa” appears before lakin~a in this position, the structure is annotated as an NAC. (See also ????????? on the use of NAC with adverbials preceded by wa.)

However, lakin~a can also head an independent matrix clause. In this circumstance, the clause is labeled “S” without the –ADV dashtag.

4.1.4.3 >an~a as complementizer, not pseudo-verb

>an~a, which is classified in traditional Arabic grammar as a sister of <in~a but which is in the same time a complementizer, is treated in the ATB always as a subordinating conjunction.

(S (VP >ak~ada أَكَّدَ
 (NP-SBJ Almut~ahamu الْمُتَّهَمُ)
 (SBAR >an~a أَنَّ
 (S (NP-SBJ hu هُوَ)
 (ADJP-PRD bariy'N بِرِيءٍ)))))

أكد المتهم أنه بريء

>ak~ad+a + Al+mut~aham+u + >an~a+hu + bariy'+N
 assure + the+accused + that+he + innocent
The accused assured that he was innocent

4.1.5 New S complement construction in MSA

jaEala (to make, begin, assume, give, grant etc...

ja'ala + object + verb structure

We may be living witnesses to a new MSA syntactic structure coming into being, namely: the 'otiose' causative, i.e., a conjugated form of j-`-l + direct object pronoun + imperfect tense. For example,

/ja`alat-nii afhamu/ "she made me understand," instead of /afhamat-nii/ or /fahhamat-nii/; and

/naj`alu-hum ya`limuuna/ for "we make them understand" instead of /nu`limu-hum/, etc.

Although we have come across this construction more and more [in Al-Majalla], we have not yet seen it described in any of the newly published MSA grammars.

4.1.6 Interaction of active, passive, and topicalization with S complements

Active Small Clause

(S (VP اعتبر considered
 (NP-SBJ الرئيس the president)
 (S (NP-SBJ الوقت the timing)
 (ADJP-PRD ملائما good)))

اعتبر الرئيس الوقت ملائما

NB: need transliteration, also in the tree

Considered the president the timing good

The president considered (that) the timing is good

Passive Small Clause

```
(S (VP أُعتبر was considered
    (NP-SBJ-1 الوقت the time)
    (S (NP-SBJ-1 *
        (ADJP-PRD ملائما good))
        (PP من by
            (NP قبل
                (NP الرئيس the president))))))
```

أُعتبر الوقت ملائما من قبل الرئيس

NB: need transliteration, also in the tree

Was considered the time good by the president

The timing was considered as good by the president.

```
(S (VP أُعتبر was considered
    (NP-SBJ-1 الوقت the time)
    (S (NP-SBJ-1 *
        (ADJP-PRD ملائما good))))
```

أُعتبر الوقت ملائما

NB: need transliteration, also in the tree

Was considered the time good

The timing was considered good.

Passive Small Clause with Topicalized Subject

```
(S (NP-TPC-1 الوقت the delay)
    (VP أُعتبر was considered
        (NP-SBJ-1 *T*)
        (S (NP-SBJ-1 *)
            (ADJP-PRD good))))
```

الوقت أُعتبر ملائما

NB: need transliteration, also in the tree

The timing was considered good

The timing was considered good

The passive trace is the same, even if the subject is topicalized, as above.

4.2 Subjects

The subject (labeled with the dashtag -SBJ) is inside VP after verb.

A simple sentence with NP subject following the verb:

(S (VP شهدت:::šahid+at:::witness/observe+it/they/she_[verb]
 (NP-SBJ أندونيسيا:::<inoduwniysiyA:::Indonesia)
 (NP-OBJ (NP تظاهرات:::taZAhur+At+K::: exhibition+[fem.pl.]+[indef.acc.]
 (ADJP مؤيدة:::mu&ay~id+ap+F:::supporting+[fem.sg.]+[indef.acc.]
 (PP ل::li-:::to/for
 (NP فلسطينيين:::-Al+filasoTiyiny~+iyina:::
 the+Palestinian+[masc.pl])))
 شهدت أندونيسيا تظاهرات مؤيدة للفلسطينيين
 \$ahidat + <inoduwniysiyA + taZAhurAt+K + mu&ay~idap+F + li+Al+filasTiyiny~iyina
 witnessed + Indonesia + demonstrations + supporting + the+Palestinians
Indonesia witnessed demonstrations supporting the Palestinians

S-NOM gerund/participle/masdar subject:

(S (VP ينبغي yanobagiy
 (S-NOM-SBJ (VP تحذير taHo*iyru
 (NP-SBJ *)
 (NP-OBJ hu هـ)))
 (PP-CLR من min
 (SBAR-NOM (WHNP-1 ما mA)
 (S (VP ينتظر yanotaZiru
 (NP-SBJ-1 *T*)
 (NP-OBJ hu هـ))))))
 ينبغي تحذيره من ما ينتظره
 yanobagiy + taHo*iyru+u+hu + min + mA + yanotaZiru+hu
 should + warn+him + from + what + wait+him
He should be warned of what is waiting for him

Note: The object of a preposition can never be the subject of a sentence.

4.2.1 Pre-verbal/Topicalized subjects

If the subject precedes the verb, it is labeled NP-TPC and traced to (NP-SBJ *T*) following the verb. The NP-TPC and the trace are coindexed.

(S (NP-TPC-1 kul~+u:::كُلُّ:::every/all/each_one+[def.nom.]
 (NP >aromaniy~+K:::أَرْمَنِيَّ:::Armenian+[indef.gen.]))
 (VP ya+HofaZ+u:::يَحْفَظُ:::he/it+preserve/protect/maintain+[ind.]
 (NP-SBJ-1 *T*)
 (NP-OBJ (NP هـ`*A:::هـ`:::this_[masc.sg.])
 (NP Al+kalAm+a:::الكلام:::the+speech/statementtalk/
 saying+[def.acc.]))))
 كل أرمني يحفظ هذا الكلام
 kul~+u + >aromaniy~+K + yaHofaZ+u + h`*A + Al+kalAm+a
 every + Armenian + maintain + this + the+speech
Every Armenian memorizes this speech

4.2.2 Clausal subjects

Some verbs require clausal subjects (list below). These subjects can be either SBAR-SBJ or, with gerunds/participles, S-NOM-SBJ. These verbs have a modal meaning like “must” or “can.”

ينبغي yanobagiy must
 أمكن >amokana can
 يستحيل yasotaHiylu is impossible
 لزم lazima was obligatory
 آن Ina was time to
 حق Haq~a come as obligation
 يجب yajibu must
 ...

(S (VP yajibu يَجِبُ
 (SBAR-SBJ >ano أَنْ
 (S (VP tataxal~aY تَتَخَلَّى
 (NP-SBJ *)
 (PP-CLR Ean عَنْ
 (NP >aTomAEi أَطْمَاعِ
 (NP ka كَ)))))))

يجب أن تتخلى عن أطماعك

yajibu >ano tataxal~aY Ean >aTomAEika
 must + to + give up + greediness+your
You must give up your greediness

(S (VP yanobagiy يَنْبَغِي
 (S-NOM-SBJ (VP <iEoTA'u إِعْطَاءُ
 (NP-SBJ *)
 (NP-DTV kul~i كُلُّ
 (NP *iy ذِي
 (NP Haq~K حَقِّ)))
 (NP-OBJ Haq~a حَقِّ
 (NP hu هُ))))))

ينبغي إعطاء كل ذي حقٍّ حقّ

yanobagiy + <iEoTA'+u + kul~+i + *iy + Haq~+K + Haq~+a+hu
 should + giving + every + owner of + right + right+his
Every right should be given to its owner

4.2.3 Empty subjects

If there is no overt lexical subject, an empty, pro-drop subject (NP-SBJ *) is inserted following the verb.

```
(S (VP -ta+Eomal+u:::تَعْمَلُ::it/they/she+work/function/act+[ind.]
  (NP-SBJ *)
  (NP-ADV taHota:::تَحْتَ::under
    (NP <i$orAf+i-::إِشْرَافٍ-::supervision/direction/
      patronage+[def.gen.]
      (NP -hi:::هِ-::its/his )))))
```

تعمل تحت إشرافه

taEomal+u + taHota + <i\$orAf+i+hi
 work + under + supervision+his
She works under his supervision

The subject can be pro-drop even if it is semantically empty:

715-9-7 (1-12) It appears that John is happy

The subjects of imperatives are usually empty also:

```
(S (VP AistamiEuwA
  (NP-SBJ * )
  (PP-CLR <ilaY إلى
    (NP kalAmi كلام
      (NP >um~ahAti أمّهات
        (NP kum كُمْ ) ) ) ) ) )
```

إستمعوا الى كلام أمّهاتكم

<isotamiEuwA + <ilaY + kalAm+i + >um~ahAt+i+kum
 listen + to + speech + mothers+your
Listen to your mothers' speech

Pro-drop empty subjects function as independent pronouns that are phonetically silent. They are not coindexed with other NPs in the Arabic Treebank because the ATB does not mark pronoun co-reference in general.

Only empty subjects that are syntactically controlled (or that are traces) are coindexed. (Pro-drop subjects cannot appear in contexts where they are syntactically controlled.) See section ???? on serial verbs and ???? on complement clauses.

```
(S (VP >axa*a أَخَذَ
  (NP-SBJ-1 AlTiflu الطفل )
  (S (VP yaDoHaku يَضْحَكُ
    (NP-SBJ-1 * ) ) ) ) )
```

أخذ الطفل يضحك

>axa*+a + Al+Tifl+u + yaDoHak+u
 take + the+child + laugh
The child started laughing

The post-verbal subject position in clauses with topicalized subjects or with WH- subjects is also empty, but it is a different type of empty category. The trace of topicalization or WH- movement is *T*, and it is ALWAYS coindexed with the moved constituent.

(S (NP-TPC-1 kul~+u::كُلُّ::every/all/each_one+[def.nom.]
 (NP >aromaniy~+K::أَرْمَنِيَّ::Armenian+[indef.gen.]))
 (VP ya+HofaZ+u::يَحْفَظُ::he/it+preserve/protect/maintain+[ind.]
 (NP-SBJ-1 *T*)
 (NP-OBJ (NP h`*A::هَذَا::this_[masc.sg.])
 (NP Al+kalAm+a::الكَلَامُ::the+speech/statementtalk/
 saying+[def.acc.]))))
 كل أرمني يحفظ هذا الكلام

kul~+u + >aromaniy~+K + yaHofaZ+u + h`*A + Al+kalAm+a
 every + Armenian + maintain + this + the+speech
Every Armenian learns this speech by heart

(NP-OBJ (NP (NP زوجت::zawoj+at+a--:wife+[fem.sg.]+[def.acc.]
 (NP ه::hu::his/its))
 (NP لورا::luwrA::Lora))
 (SBAR (WHNP-2 التي::Al~atiy::which/who/whom_[fem.sg.])
 (S (VP كانت::kAn+at::be/was/were+it/they/she_[verb]
 (VP تتحدث::ta+taHad~av+u::it/they/she+speak/discuss+[ind.]
 (NP-SBJ-2 *T*)
 (PP على::EalaY::on/above
 (NP الهاتف::Al+hAtif+i::
 the+telephone+[def.gen.])))))))
 زوجته لورا التي كانت تتحدث على الهاتف

zawoj+at+a+hu + luwrA + Al~atiy + kAn+at + ta+taHad~av+u + Ealay + Al+hAtif+i
 wife+his Laura + who + was + talk + on + the telephone
His wife Laura who was talking on the telephone

4.3 Direct and indirect objects

NP direct objects of the verb are labeled NP-OBJ.

(S (VP >arosal+a::أُرْسِلَ::send/transmit+he/it_[verb]
 (NP-SBJ jamAl::جَمَال::Jamal/Gamal
 bA\$A::بَاشَا::pasha
 Al+saf~AH+u::السَّفَّاحُ::the+shedder_of_blood/butcher+[def.nom.])
 (NP-OBJ tiligrAf+AF::تِلْغَرَاْفَا::telegraph/telegram+[acc.indef.])
 (PP <ilaY::إِلَى::to/towards
 (NP Al+mutaSar~if+i::الْمُتَصَرِّفُ::the+provincial_governor
 +[def.gen.]))))
 أُرسل جمال باشا السفاح تلغرافاً إلى المتصرف
 >arosal+a + jamAl + bA\$A + Al+saf~AH + tiligorAf+AF + <ilaY + Al+mutaSar~if+i
 send + Jamal + Basha + the+butcher + telegraph + to + the+provincial governor
Jamal Basha sent a telegraph to the butcher provincial governor

(S (VP waZ~afa وَظَّفَ
 (NP-SBJ zayodN زَيْدُ)
 (NP-OBJ-1 EamorAF عَمْرًا)
 (S (NP-SBJ-1 *)
 (NP-PRD mudiyraAF مُدِيرًا))))

وظف زيد عمرا مديرا

waZ~af+a + zayod+N + Eamor+AF + mudiyra+AF
 employ + Zayd + Amr + director
Zaid employed Amr as a director

Ditransitive (indirect) objects are labeled NP-DTV or PP-DTV, as appropriate. An example of a sentence with two objects (one labeled NP-OBJ and the other labeled NP-DTV) is seen in

(S wa-::و-::and
 (VP ->aEoTaY+[null]::أَعْطَى[نُلل]::give/provide+he/it_[verb]
 (NP-SBJ Al+baToriyarok+u::البَطْرِيَرُكُ::the+Patriarch+[def.nom.])
 (NP-DTV Al+mawArin+ap+a::المَوَارِنَةُ::the+Maronites+[fem.sg.]+[def.acc.])
 (NP-OBJ (NP kaniys+ap+F::كَنِيسَةُ::church/temple+[fem.sg.]+[indef.acc.])
 (PP-LOC fiy::فِي::in
 (NP \$AriE+i::شارع::street+[def.gen.]
 (NP >wlm::أُولْمَ::Ulm)))))))

وأعطى البطريرك الموارنة كنيسة في شارع أولم

wa+>aEoTaY + Al+baToriyarok+u + Al+mawArin+ap+a + kaniysap+F + fiy + \$AriE+i +
 >uwlom
 and+provided + the+Patriarch + the+Maronites + church +in + street + Ulm
The Patriarch provided the Maronites with a church in Ulm Street

Gerunds/participles that are functioning as the direct object or indirect object are labeled S-NOM-OBJ or S-NOM-DTV.

(S (VP >aEAd+a::أَعَادَ::[do]_again/re-[do]/return/give_back+he/it_[verb]
 (NP-SBJ *)
 (S-NOM-OBJ (VP takoliyf+a::تَكْلِيْفٌ::charging/commissioning+[def.acc.]
 (NP-SBJ *)
 (NP-DTV -hu::هُ::its/his)
 (NP-OBJ ta>oliyf+a::تَأْلِيْفٌ::composing/constituting
 +[def.acc.]
 (NP Hukuwm+ap+K::حُكُومَةٌ::government/
 administration+[fem.sg.]+[indef.gen.]
 jadiyd+ap+K::جَدِيدَةٌ::new/modern
 +[fem.sg.]+[indef.gen.]))))))
 أعاد تكليفه تأليف حكومة جديدة

>aEAd+a + takoliyf+a+hu + ta>oliyf+a + Hukuwmap+K + jadiydap+K
 do again + charging+it + composing + government + new
Its charge repeated composing a new government

```
(S (VP >aEoTawA أَعْطُوا
  (NP-SBJ * )
  (S-NOM-DTV (VP AijotiyAHa إْجْتِيَاخُ
    (NP-SBJ hu هُوَ )
    (NP-OBJ Alkuwayta الكُوَيْتِ ) ) )
  (NP-OBJ >aboEAdAF أَبْعَاداً
    xaTiyrapF خطيرة)))
```

أعطوا إْجْتِيَاخَهُ الكُوَيْتِ أَبْعَاداً خطيرة

>aEoTawA + <ijotiyAH+a+hu + Al+kuwayt+a + >aboEAd+AF + xaTiyrap+F
 gave + invasion+his + the+Kuwait + dimensions + serious
They thought his invasion for Kuwait has serious dimensions

4.3.1 Clitic object pronouns

Cliticized object pronouns are split from the verb, so that the node label for the object can be annotated.

```
(S (VP qara>a قَرَأَ
  (NP-SBJ * )
  (NP-OBJ hA هَا ) ) )
```

قرأها

qara>+a+hA
 read+it
He read it

When the object pronoun is separated from the verb in the text, it should always be cliticized to the dummy particle >iy~A. (See section ?????????? on pronouns for more on >iy~A.)

```
(S (VP >aEoTA أَعْطَا
  (NP-SBJ * )
  (NP-DTV hu هُوَ )
  (NP-OBJ <iy~A إِيَّاهُ
    hA هَا ) ) )
```

أعطاه إِيَّاهَا

>aEoTAhu <iy~AhA
 gave+him + it+itself
He gave it to him

4.4 Closely related prepositional phrases (PP-CLR)

PPs that are "Closely Related" to the verb are given the -CLR function tag. This is used for all PPs that are complements of the verb, with the exception of ditransitive verbs where PP-DTV is used. PP-CLR should be obligatory for the sense of the verb being used in context, as with all other arguments.

(S (VP kAn+a::كَانَ::be/was/were+he/it_[verb]
 (VP ya+EoTif+u::يَعْطِفُ::he/it+be_sympathetic/incline+[ind.]
 (NP-SBJ *)
 (PP-CLR EalaY::عَلَى::on/above
 (NP Al+ma>omuwr+iyna::المُؤْمَرِينَ::the+officer
 +[masc.pl.gen.])))))
 كان يعطف على المُؤْمَرِينَ

kAn+a + yaEoTif+u + EalaY + Al+ma>muwriyn
 was + be sympathetic + on + the+officers
He was sympathetic with the officers

(S (NP-TPC-1 kul~+u::كُلُّ::every/all/each_one+[def.nom.]
 (NP (NP h`*ihi::هَذِهِ::this/these)
 (NP Al+Turuq+i::الطُرُق::the+roads/ways/means/methods
 +[def.gen.])))
 (VP tu+&ad~iy+[null]::تُؤَدِّي[نُلل]::it/they/she+direct/guide/lead+[ind.]
 (NP-SBJ-1 *T*)
 (PP-CLR <ilaY::إِلَى::to/towards
 (NP TAHuwn+ap+K::طاحُونَةٍ::windmill/grinder+[fem.sg.]+[indef.gen.]
 wAHid+ap+K::وَاحِدَةً::one/single+[fem.sg.]+[indef.gen.]))))
 كل هذه الطرق تؤدي الى طاحونة واحدة

kul~+u + h`*ihi + Al+Turuq+i + tu+&ad~iy + <ilaY + TAHuwnap+K + wAHidap+K
 all + these + the+roads + lead + to + windmill + one
All these roads lead to one windmill

See section ????? on passive verbs with PP-CLR.

4.4.1 With NP-OBJ and PP-CLR

Some verbs have both an NP-OBJ and a PP-CLR.
 For example, نعت naEata:

(S (VP نعت
 (NP-SBJ *)
 (NP-OBJ هـ)
 (PP-CLR بـ
 (NP (الختال)))))

نعتة بالمختال

NB: need transliteration, also in the tree
 Called he him the crook
He called him a crook

4.4.2 With PP-CLR and PP-CLR

Some verbs have two PP-CLR complements.

For example,

```

(S (VP سمح
  (NP-SBJ الحكم
    (PP-CLR لـ
      (NP (اللاعب ))
        (PP-CLR بـ
          (NP مغادرة
            (NP الملعب ))))))

```

سمح الحكم للاعب بمغادرة الملعب

NB: need transliteration, also in the tree

Allowed the judge to the player to leaving the stadium

The judge allowed to the player to leave the stadium.

4.5 Adverbial modification

All adverbial modification of the sentence and the verb phrase appears within the VP if the modifier follows the verb, and within the S if the modifier precedes the verb. PPs (Prepositional Phrases) and ADVPs (Adverb Phrases) are by default adverbial (i.e., they don't need the non-specific -ADV tag). NP, S and SBAR are arguments by default, and so all of these need some kind of adverbial function tag when they are analyzed as having adverbial function (-ADV if no more specific tag is appropriate).

A specific adverbial function tag is used for all adverbials whenever it is appropriate: -TMP temporal, -LOC locative, -DIR directional, -PRP purpose, -MNR manner. If no specific function is appropriate, -ADV must be used for adverbial noun phrases and clauses: NP-ADV, S-ADV and SBAR-ADV. (Since PP and ADVP are assumed to be adverbial by nature, they do not get the general -ADV tag. They do, however, get any more specific adverbial tag as appropriate.)

See section ????? for a definition of each dashtag.

NP-TMP
 NP-ADV
 S-ADV
 S-MNR
 SBAR-LOC
 PP-LOC
 SBAR-ADV
 SBAR-PRP
 PP-DIR
 NP-DIR
 NP-LOC

4.5.1 Adverbial modifiers vs. arguments

Note that temporal expressions such as today, tomorrow, yesterday (need Arabic), for example, can be used in argument or predicate positions, not always modifying positions.

(S (NP-SBJ ha*A هَذَا)
 (NP-PRD yawomN يَوْمٌ
 SaEobN صَعْبٌ))

هذا يوم صعب

ha*A + yawom+ N + saEob+N

this + day + hard

This is a hard day

In this example, “day” is the predicate of the sentence and does not serve and modifying function. As such, it is marked –PRD and not as –TMP.

Locative expressions can also be used in argument positions:

(S (VP gAdara غادرَ
 (NP-SBJ *)
 (NP-OBJ AldAr الدار
 AlbayDA' البيضاء)))

غادر الدار البيضاء

gAdar+a + Al+dAr+a + Al+bayDA'+a

left + the+house + the+white

He left Casablanca

In this example, the name of the city is a direct object of the transitive verb, and as such it should be annotated NP-OBJ. City and place names are marked as –LOC *only if they are functioning as adverbial modifiers of the action*. When they function in an argument role, they are annotated as an argument.

4.5.2 S-ADV vs. S-MNR (HAI)

Both S-ADV and S-MNR are HAI. They can be distinguished by whether the HAI expresses the manner in which the main verb’s action takes place (S-MNR) or if it expresses a more general, less specific adverbial modification (S-ADV).

S-MNR example:

(S (VP انقضَّ
 (NP-SBJ *)
 (PP-CLR على
 (NP له))
 (S-MNR (VP انقضاض
 (NP-SBJ الأسد)
 (PP-CLR على
 (NP فريسته
 (NP له)))))))

انقضَّ عليه انقضاض الأسد على فريسته

Jumped he on him jumping the lion on prey his

He attacked him the way a lion would attack its prey

S-ADV example:

(S-NOM (VP -qiyAm+i-::قيام::undertaking/carrying_out/setting_up+[def.gen.]
 (NP-SBJ-2 -hi::هـ::its/his)
 (PP-CLR bi-::بـ::by/with
 (NP (NP ->aEomAl+i-::أعمال::actions/activities/work
 +[def.gen.]
 (NP salob+K::سَلْب::robbing/dispossession
 +[indef.gen.]))
 (ADJP Eid~+ap+K::عِدَّة::several/numerous/many
 +[fem.sg.]+[indef.gen.])))
 (S-ADV (VP musotaxodim+AF::مُسْتَخْدِمًا::employing/using+[acc.indef.]
 (NP-SBJ-2 *)
 (NP-OBJ say~Ar+ap+a::سَيَّارَةً::car/automobile/vehicle
 +[fem.sg.]+[def.acc.]
 (NP riynuw::ريْنُو::Renault)))))
 قيامه بأعمال سلب عدة مستخدماً سيارة رينو
 qiyAm+i+hi + bi->aEomAl+i + salob+K + Eid~ap+K + musotaxodim+AF + say~Arap+a +
 riynuw
 undertaking+his + with+activities + robbing + several + using + car + Renault
His undertaking several robbing activities using Renault

(S (VP qAla قال
 (NP-SBJ-1 *)
 (NP-OBJ *alika ذلك)
 (S-ADV (VP mu\$irAF مُشِيرًا
 (NP-SBJ-1 *)
 (PP-CLR <ilaY إلى
 (SBAR-NOM (WHNP-1 mA ما)
 (S (VP HaSala حَصَلَ
 (NP-SBJ-1 *T*)
 (PP-LOC fiy في
 (NP dArfuwr دارْفُور))))))))))
 قال ذلك مشيراً إلى ما حصل في دافور
 qAl+a + *alika + mu\$ir+AF + <ilaY + mA + HaSal+a + fiy + dArfuwr
 said + that + pointing + to + what + happened + in + Dafur
He said that referring to what happened to Dafur

The empty subject of the HAl is coindexed with the subject or the object of the higher clause if appropriate (that is, if the interpretation is that the two have the same referent).

4.5.3 –DIR vs. –LOC

–DIR marks adverbial constituents expressing motion to or from a place, while –LOC marks adverbial constituents expressing static location. –DIR is related to verbs expressing physical movement; –LOC is related to verbs expressing static physical location.

(S (VP fight/wage_war+it/they/she |HArabat- | حَارَبَتْ-
 (NP-SBJ Damascus |dima\$oqu | دِمَشْقُ)
 (NP-OBJ (NP the+fundamentalist |-Al>uSuwliiy~apa | الأُصُولِيَّةُ-
 (PP-LOC in |fiy | فِي
 (NP the+areas/zones/territories |AlmanATiqi | المناطق
 the+Syrian |Als~uwriy~api | السُّورِيَّةُ))))))
 حَارَبَتْ دِمَشْقُ الْإِصْوَليَّةُ فِي الْمَنَاطِقِ السُّورِيَّةِ
 HArabat + dima\$oqu+u + Al+>uSuwliiy~ap+a + fiy + Al+manATiq+i + Al+suwriy~ap+i
 fought + Damascus + the+fundamentalist + in + the+territories + the+Syrian
Damascus fought against fundamentalism in the Syrian territories

(S (VP infiltrate+he/it |tasal~ala | تَسَلَّلَ
 (NP-SBJ foreigner |>ajonabiy~N | أَجْنَبِيٌّ
 (PP-DIR from |min | مِنْ
 (NP Afghanistan |>afogAnisotAn | أَفْغَانِسْتَانُ))
 (PP-DIR to/towards |<ilaY | إِلَى
 (NP Iran |<iyrAn | (إِيرانُ))))))
 تَسَلَّلَ أَجْنَبِيٌّ مِنْ أَفْغَانِسْتَانُ إِلَى إِيْرَانُ
 tasal~al+a + >ajonabiy~+N + min + >afogAnisotAn + <ilaY + <iyrAn
 sneaked + foreigner + from + Afghanistan + to + Iran
A foreigner sneaked from Afghanistan to Iran

(See section ????? on distinguishing adverbial NPs from arguments.)

4.5.4 NP-TMP vs. NP-ADV

NPs representing time or temporal expressions should be annotated as NP-TMP when they function as temporal modifiers. This includes expressions such as today, tomorrow, yesterday (need Arabic) when used as temporal modifiers.

(S (NP-TMP Alyawoma |اليَوْمُ
 (VP >anohayotu |أَنْهَيْتُ
 (NP-SBJ *)
 (NP-OBJ Eamal |عَمَلُ
 (NP iy |يِ
 (NP-ADV bAkirAF |بَاكِراً))
 اليوم أنهيته عملي باكراً
 Alyawom+a + >anohayot+u + Eamal+iy + bAkir+AF
 today + finished + work+my + early
Today, I finished my work early

(See section ????? on distinguishing adverbial NPs from arguments.)

Only NPs representing an adverbial expression without a more specific meaning should be annotated as NP-ADV.

4.5.5 Nouns formerly prepositions

Note that the NPs headed by nouns that were previously treated as prepositions (see section ???? for the full listing of these nouns) ALL need to have an adverbial dashtag. Many will be NP-TMP or NP-LOC, but any without the more specific adverbial meaning must be annotated as NP-ADV (these correspond to the previous annotation of these words as PPs that had no dashtag).

```
(S (VP xaraq+at::خَرَقَتْ::tear/violate+it/they/she_[verb]
  (NP-SBJ <isorA}iyl+u::إِسْرَائِيلُ::Israel+[def.nom.] )
  (PP-CLR bi-::بِ-::by/with
    (NP -TA}ir+At+i-::-طَائِرَات-::aircraft/airplane+[fem.pl.]
      +[def.gen.]
      (NP -hA::-هَا::its/their/her )))
  (NP-OBJ jidAr+a::جِدَارٌ::wall+[def.acc.]
    (NP Al+Sawot+i::الصَّوْتُ::the+voice/sound+[def.gen.] ))
  (NP-LOC fawoqa::فَوْقَ::above/over
    (NP bayoruwt+a::بَيْرُوتُ::Beirut+[indef.gen.] ))))
خرقت إسرائيل بطائراتها جدار الصوت فوق بيروت
xaraqat + <isorA}iyl + bi+TA}irAt+i+hA + jidAr+a + Al+Sawot+i + fawoqa + bayoruwt+a
tear + Israel + with+aircraft+its + wall + the+sound + above + Beirut
With its aircraft, Israel tore the sound wall above Beirut
```

4.6 Passive Verbs

Verbs in the passive form always have a passive object trace which is co-indexed to the subject: (NP-OBJ-1 *). The passive trace is * rather than *T* because of the type of movement involved.

```
(S (VP ذهل *uhila
  (NP-SBJ-1 الجمهور Aljumhuwru)
  (NP-OBJ-1 *)))
```

ذهل الجمهور

*uhil+a + Al+jumohuwr+u
shocked + the+audience
The audiences were shocked

The passive trace is the same, even if the subject is topicalized.

All passive traces are NP-OBJ * or NP-DTV * as appropriate (even if the subject of the passive verb is an S or an SBAR, for example).

NB: Note to discuss in the future whether passive traces should always be invariable NP-OBJ/DTV * or if the passive trace node label category should match the category of the moved constituent (surface subject).

4.6.1 Passive of “monotransitive” verbs

With simple transitive verbs, the passivization process moves the object of the active verb into the subject position of the passive verb, leaving a trace in the object position. This “passive” trace is coindexed with the subject.

```
(S wa-::و-::and
  (VP -qutil+a::-قُتِلَ::be_killed+he/it_[verb]
    (NP-SBJ-1 qA}id+u::قائد::leader/commander+[def.nom.]
      (NP Al+TA}ir+ap+i::الطائرة::the+aircraft/airplane
        +[fem.sg.]+[def.gen.] )))
    (NP-OBJ-1 *)
    (PP-TMP EalaY::على::on/above
      (NP Al+fawor+i::الفور::the+immediately/at_once+[def.gen.] ))))
  و قتل قائد الطائرة على الفور
  wa+qutil+a + qA}id+u + Al+TA}ir+ap+i + EalaY + Al+fawor+i
  and+be killed + commander + the+airplane + on + the+immediately
  The pilot of the airplane was immediately killed
```

4.6.2 Passive of ditransitive verbs

With ditransitive verbs, the passivization process moves either the direct object (NP-OBJ) or the indirect object (NP-DTV) of the active verb into the subject position of the passive verb, leaving a trace in the respective object position. This “passive” trace is coindexed with the subject.

4.6.2.1 Passivized direct object

The subject of this passive sentence is coindexed with a passive trace in the direct object position.

Note that when the direct object of a ditransitive verb is passivized, the indirect object must be a prepositional phrase (PP-DTV) rather than a noun phrase.

```
(S (VP suliba سلب
  (NP-SBJ-1 AlHaq~ الحق )
  (NP-OBJ-1 * )
  (PP-DTV min من
    (NP >aSoHAbi أصحاب
      (NP hi هـ ) ) ) ) )
  سلب الحق من أصحابه
  sulib+a + Al+Haq~+u + min + >aSoHAb+i+hi
  was snatched + the+right + from + owners+its
  The right was snatched from its owners
```

4.6.2.2 Passivized indirect object

The subject of this passive sentence is coindexed with a passive trace in the indirect object position. The direct object is marked NP-OBJ as usual.

```
(S (VP منحت:::muniH+at:::be_granted/be_awarded +it/they/she_[verb]
  (NP-SBJ-3 بعض:::baEoD+a:::some/several+[def.acc.]
    (NP المؤسسات:::Al+mu&as~as+At+i:::the+institutions/
      organizations+[fem.pl.]+[def.gen.])))
  (NP-DTV-3 *)
  (NP-OBJ تراخيص:::tarAxiyS+a:::granting_of_permission/
    licensing+[def.acc.])))
منحت بعض المؤسسات تراخيص
muniHat + baEoD+u + Al+mu&as~as+At+i + tarAxiyS+a
was granted + some + institutions + licensing
Some institutions were granted licensing
```

4.6.3 Passive of verbs with a clausal complement

4.6.3.1 SBAR complement

SBAR becomes the subject of the passive sentence, coindexed with an NP-OBJ passive trace.

4.6.3.2 S complement

4.6.3.2.1 Verbs of judgement

The subject of complement clause in the active version becomes subject of the passive sentence, and the passive trace is in complement clause.

```
(S (NP-TMP yawom+a-:::يَوْم-:::today+[def.acc.]
  (NP -hA:::-ها:::its/their/her ))
  (VP }uEotubir+a:::اُعْتَبِرَ:::be_considered/be_regarded/be_believed+he/it_[verb]
    (NP-SBJ-1 nawAz:::نواز:::NOT_IN_LEXICON
      $ariyf:::شريف:::Sharif )
    (S (NP-SBJ-1 *)
      (ADJP-PRD mutaxA*il+AF:::مُتَخَذِلًا:::weak/exhausted+[acc.indef.] ))))
يومها إعتبر نواز شريف متخاذلاً
yawom+a+hA + >uEotubir+a + nawAz + $ariyf + mutaxA*il+AF
day+its + was considered + Nawaz + Sharif + exhausted
At that time, Nawaz Sharif was considered as exhausted
```

4.6.3.2.2 Verbs of transformation and labelling

Object of verb becomes subject of sentence, trace in object position, no passive trace in complement clause (although the empty subject of the lower clause may be coindexed with the object of the higher verb as appropriate).

4.6.4 Passive of verbs that are transitive with a PP-CLR:

See section ???? for a description of active sentences with these verbs.

These verbs have only two arguments: the subject and the PP-CLR complement. They do not ever have an NP direct object. In the passive sentence, the subject argument disappears, as expected, but the PP-CLR complement remains as a PP (and it is not possible to front the object of the preposition as in English). The passive verb in these constructions does not inflect for gender or number.

The annotation of this kind of passive sentence is to insert an empty subject, and to coindex that subject with the object of the preposition. Note that this particular coindexing is specific to this construction, and it is unlike all other coindexing in the ATB (it does not represent syntactic coindexing or structural coindexing). There is no passive trace inserted (because these verbs do not allow for any NP object).

```
(S (VP passive-verb
  (NP-SBJ-1 *)
  (PP-CLR preposition
    (NP-1 noun))))
```

Active:

```
(S (VP عثر stumbled/discovered
  (NP-SBJ زيد Zayd)
  (PP-CLR على on
    (NP كنز treasure))))
```

عثر زيد على كنز

NB: need transliteration, and in the tree

discovered Zayd on treasure
Zayd found a treasure

Passive:

```
(S (VP عُثر was-discovered/was-stumbled
  (NP-SBJ-1 *)
  (PP-CLR على on
    (NP-1 كنز treasure))))
```

عُثر على كنز

NB: need transliteration, and in the tree

was-discovered on treasure
A treasure was found

```
(S (VP عثر::Euvir+a::be_discovered/be_found+he/it_[verb]
  (NP-SBJ-1 *)
  (PP-TMP بعد::baEoda::after
    (NP ظهر::Zuhor+i::noon/afternoon+[def.gen.]
      (NP امس::>amosi::yesterday)))
  (PP-CLR على::EalaY::on/above
    (NP-1 محمد::muHam~ad::Muhammad/Mohamed ))))
```

عثر بعد ظهر أمس على محمد

Euvir+a + baEoda + Zuhor+i + >amos+i + EalaY + muham~ad
 was found + after + noon + yesterday + on + Muhammad
Yesterday afternoon, Muhammad was found

```
(S (VP قبض::qubiD+a::be_arrested/be_apprehended/be_seized+he/it_[verb]
  (NP-SBJ-2 *)
  (PP-CLR على::EalaY::on/above
    (NP-2 (NP سائق::sA}iq+i-::chauffeur/driver+[def.gen.]
      (NP ها::-hA::its/their/her))
      (NP علي::Ealiy~::Ali
        محمد::muHam~ad::Muhammad/Mohamed )))))
```

قبض على سائقها علي محمد

qubiD+a + EalaY + sA}iq+i+hA + Ealiy~ + muHam~ad
 was seized + on + driver+its + Ali + Muhammad
Its driver, Ali Muhammad, was seized

This passive form can occur with a topicalized NP also, where the topicalized NP is co-referent with the resumptive pronoun object of the preposition. For example,

```
(S (NP-TPC-3 Al+tahodiyd+At+i)
  (VP (PRT لا)
    yu+sotahAn+u
    (NP-SBJ-2 *)
    (PP-CLR bi-
      (NP-2 (NP -hA)
        (NP-3 *T*)))))
```

the+threats/dangers not he+be disdained with it

4.6.5 Passive of verbs that are ditransitive with a PP-CLR:

Note that verbs that take both a direct object and also a PP-CLR complement follow passivization rules similar to ditransitive verbs. That is to say, when the direct object is passivized, the normal passive structure applies and the PP-CLR is unaffected.

```

(SBAR (WHNP-2 -Al~atiy||which/who/whom_[fem.sg.])
  (S (VP tuw~ij+at||be_crowned+it/they/she_[verb]
    (NP-SBJ-2 *T*)
    (NP-OBJ-2 *)
    (PP-CLR bi-||by/with
      (NP -tawoqiye+i||signing+[def.gen.]
        (NP (NP {it~ifAq+i||agreement/accord/treaty+[def.gen.]
          (NP Al+tijAr+ap+i||the+commerce/business+
            [fem.sg.]+[def.gen.]
            Al+Hur~+ap+i||the+free/independent+
              [fem.sg.]+[def.gen.])))
          (NP-ADV bayon+a-||between/among+[def.acc.]
            (NP -humA||their))))))
      (NP-TMP Al+$ahor+a||the+month+[def.acc.]
        Al+mADiy+a||the+past/bygone+[def.acc.])))

```

**NB: Need Arabic string,
transliteration,
gloss,
and translation**

And when the object of the PP-CLR is passivized, the “PP-CLR passive” rules apply – the subject is co-indexed with the object of the preposition, and there is no passive trace.

These verbs do not follow the simple “transitive with PP-CLR” passive rules! (see section ?????? above for the passive structure of verbs that are simply “transitive with PP-CLR” without the direct object.)

4.6.6 Passives in the passive participle form

See section ?????? for a full description of the annotation of participles, gerunds and masdar.

In sentences where the passive participle is annotated as a VP, the subject is coindexed with a passive trace just as it is with a regular verbal form.

(S (VP >uj~ila أَجَّلَ
 (NP-SBJ-2 (NP AlAijotimAEu الاجتماع)
 (SBAR (WHNP-3 0)
 (S (VP Almuqar~ru المقرّر)
 (NP-SBJ Eaqoduhu عَقْدُهُ)
 (NP-OBJ-3 *T*)
 (PP-LOC fiy في
 (NP bayorwt بيروت))))))
 (NP-OBJ-2 *)
 (NP-TMP <ilaY إلى
 (NP (NP >ajalK أَجَل)
 (NP gayri غَيْر
 (NP musam~Y مُسَمًى))))))
 أَجَل الاجتماع المقرّر عقده في بيروت إلى أجل غير مسمّى
 >uj~il+a + Al+<ijotimAE+u + Al+muqar~ar+u + Eaqod+u+hu + fiy + bayorut + <ilaY +
 >ajal+K + gayr+i + musam~Y
 was deferred + the+meeting + the+decided + holding+its + in + Beirut + to + time + not +
 named
The meeting that was decided to be held in Beirut was deferred to an indefinite date

4.7 Active verbs that are semantically similar to passives

With some verbal morphological patterns like tafaE~l, tatAEala and AifoEala, the syntactic subject is a patient of the action. However, these verbs should not be treated as passive because they are in the active form. They cannot have a passive form because they are intransitive.

Care must be taken to annotate verbs in these forms as active, not passive. They should not have passive trace objects (in fact, because they are intransitive, they should not have direct objects of any kind).

(cf. The door opened, the window broke, etc.)

Form 5 p. 24 bottom table in Fischer

taC1aC2aC3~a (tafaEal~a)

4.8 Floating Quantifiers

Floating quantifiers always occur in Arabic with a resumptive pronoun. The structure is a typical topicalization structure with resumptive pronoun.

```

(S (NP-TPC-1)
  (VP جاءوا
    (NP-SBJ كل
      (NP (NP هم)
        (NP-1 *T*))
      (NP (NP إلى
        (NP الاعتصام))
      (NP (NP إلى
        (NP الاعتصام))

```

الطلبة جاءوا كلهم إلى الاعتصام

NB: need transliteration, and in the tree

Students came all them to the sit-in

All students came to the sit-in.

5 Gerunds (maSdar/gerund) and participles (>ism AlfaEil and >ism AlmafoEuwl)

See related sections in the POS guidelines on definition and morphology of maSdar/gerund and active and passive participles.

All gerunds (maSdar), active participle (>ism Al-fAEil) and passive participle (>ism Al-mafEuwl) constructions are analyzed as NPs, ADJPs or VPs, depending on the context. Active participle (>ism Al-fAEil) and passive participle (>ism Al-mafEuwl) forms are often analyzed as adjectives, unless they are lexicalized as nouns. See section ????? of the POS guidelines for more details on how to distinguish the nouns from the adjectives.

5.1 Non-verbal reading of gerunds and participles

The following rules force a nominal or adjectival reading of gerunds and participles. These rules are not necessarily ordered but if one applies all the other rules are pre-empted.

5.1.1 Plural form

The plural form of a gerund or participle always forces a nominal reading, since the word is lexicalized as a noun.

```

(S (NP-SBJ اجتماعات meetings
  (NP اللجنة committee))
  (ADJP-PRD مهمة important))

```

اجتماعات اللجنة مهمّة

NB: need transliteration, and in the tree

Meetings the committee important

The committee's meeting are important

5.1.2 Single word, not in HAl position

See also section ????? 5.2.2.2.2 on the HAl construction.

If the maSdar/gerund, active participle or passive participle is a single word without any argument related to it and the active participle and passive participle are NOT in a HAl position, the maSdar/gerund is treated as an NP and the active or passive participle is treated as an adjective or noun, depending on its position in the sentence.

Single-word maSdar/gerund, always NP:

(S (VP >alogay ألغى
(NP-SBJ Al+mudiyru+u المدير
(NP-OBJ Al+<ijotimAE+a الاجتماع)))

الغى المدير الاجتماع

>alogay + Al+mudiyru+u + Al+<ijotimAE+a
cancel + the+director + the+meeting
The director canceled the meeting

Single-word active participle adjective and ADJP:

(S (NP-SBJ h`*A هذا)
(NP-PRD manotuwj+N منتوج
rA}ij+N رائج))

هذا منتوج رائج

h`*A + manotuwj+N + rA}ij+N
this + product + widespread
This product is wide spread

(S (VP kAna
(NP-SBJ zayodN)
(ADJP-PRD jAlisAF)))

كان زيد جالسا

kAna zayodN jAlisAF
was + Zayd + sitting
Zayd was sitting

Single-word active participle NP:

(S (NP-SBJ najiyb maHofuZ)
(NP-PRD kAtibN))

نجيب محفوظ كاتب

najiyb maHofuZ kAtibN
Najeeb MAhfouz + writer
Najeeb MAhfouz is a writer

Single-word passive participle adjective and ADJP:

(S (NP-TTL-SBJ Bayna
 (NP Al+qasrayoni))
 (NP-PRD riwAyatN
maEoruwfapN)

بين القصيرين رواية معروفة

Bayna Al qasrayoni riwAyatN maEoruwfapN
 Bayna Al qasrayni + novel + famous
BAyna Al Qasrayni is a famous novel

(SQ (PRT hal)
 (NP-SBJ zayodN)
(ADJP-PRD mut~ahamN))

هل زيد متهم

Hal zayodN mut~ahamN
 PART + Zayd + accused
Is Zayd accused

Single-word passive participle NP:

(S (VP >axoraj+at أَخْرَجَتْ
 (NP-SBJ Al+wizArap+u + الوزارَة)
(NP-OBJ (NP mano\$uwr+AF مَنشوراً)
 (SBAR (WHNP-1 *0*)
 (S (VP Dam~+a ضَمَّ
 (NP-SBJ-1 *T*)
 (NP-OBJ Al+taEodiylAt+i + التَّعْدِيلَاتُ
 Al+jadiydap+a + الجَدِيدَة))))))
 أَخْرَجَتِ الْوَزَارَةُ مَنشوراً ضَمَّ التَّعْدِيلَاتُ الْجَدِيدَة
 >axoraj+at + Al+wizArap+u + mano\$uwr+AF + Dam~+a + Al+taEodiylAt+i +
 Al+jadiydap+a
 make + the+ministry + flier + include + the+modifications + the+new
The ministry released a flier which included the recent modifications

5.1.3 Definite maSdar/gerund

If the maSdar/gerund is definite (carries the determiner Al-), it must be an NP. These forms with Al- are treated as nominal regardless of any other modification or argumentation.

(S (VP >alogaY أُلغِيَ
 (NP-SBJ Al+mudiyar+u المَدِيرُ)
(NP-OBJ Al+<ijotimAE+a الإِجْتِمَاعُ)))

ألغى المدير الإِجْتِمَاعُ

>alogaY + Al+mudiyar+u + Al+<ijotimAE+a
 cancel + the+director + the+meeting
The director canceled the meeting

No matter how strong the adverbial modifier is, a maSdar/gerund with Al- must be an NP:

(S (VP Euwqiba عُوقِبَ
 (NP-SBJ-1 AllAEibu اللاعب)
 (NP-OBJ-1 *)
 (PP bi ب
 (NP (NP Al<iyqAfi الإيقاف)
 (NP-ADV (ADJP >akovara أكثر
 (PP min من
 (NP mubArAtK مباراة))))))))
 عوقب اللاعب بالإيقاف أكثر من مباراة
 Euwqiba Al+lAEib+u bi+Al+<iyqAf+i >akovar+a min mubArAt+K
 was punished the+player with+the+suspension more than game
The player was punished by suspension for more than one game

5.1.4 Single word modified by an adjective

If the maSdar/gerund or the participle is modified by an adjective, it is annotated as an NP.

maSdar:

(S (VP >alogaY ألغى
 (NP-SBJ Al+mudiyru المدير)
 (NP-OBJ <ijotimAEAF إجتماعاً
 muhim~AF مهماً)))
 ألغى المدير إجتماعاً مهماً
 >alogaY Al+mudiyru+u <ijotimAE+AF muhim~+AF
 cancelled the+manager meeting important
The manager cancelled an important meeting

Active participle:

(S (VP kAna كان
 (NP-SBJ *)
 (NP-PRD kAtibAF كاتباً
 EaZymAN عظيماً)))
 كان كاتباً عظيماً
 kAn+a kAtib+AF EaZym+AN
 was + writer + great
He was a great writer

Passive participle:

(S (VP yaniS~u ينص
 (NP-SBJ Alman\$uwru المنشور
 Aljadiydu الجديد)
 (PP-CLR EalaY على
 (NP (NP majmuwEapK مجموعة)
 (PP min من
 (NP AltawSiyAti التوصيات
 Almuhim~api المهمة)))))
 ينص المنشور الجديد على مجموعة من التوصيات المهمة
 yaniS~+u Al+man\$uwr+u Al+jadiyd+u EalaY majmuwEap+K min Al+tawSiyAt+i
 Al+muhim~api
 text the+flier the+new on group of the+recommendations the+important
 The new flier enlists some important recommendations

5.1.5 Single word followed by an adverbial of degree

If an active participle or passive participle is a single word modified by an adverbial of degree (such as jid~AF or Haq~AF), it must be an ADJP. (maSdar/gerund is not considered here because maSdar/gerund(s) cannot be modified by degree adverbs.)

Active participle:

(S (VP kAna كان
 (NP-SBJ xiTAbu خطاب
 (NP Alra}iysi الرئيس))
 (ADJP-PRD mu&av~irAF مؤثراً
 Haq~AF حقاً)))
 كان خطاب الرئيس مؤثراً حقاً
 kAn+a xiTAb+u Al+ra}iys+i mu&av~ir+AF Haq~+AF
 was+speech the+president+effective+really
 The speech of the president was really impressive.

Passive participle:

(S (VP kAna كان
 (NP-SBJ Eam~Aru عمار)
 (NP-PRD (NP zamiylAF زميلاً)
 (ADJP maHobuwbaF محبوباً
 jid~AF جداً))))
 كان عمار زميلاً محبوباً جداً
 kAna Eam~Ar zamiyl+AF maHobuwbaF jid~+AF
 was Ammar colleague admirable very
 Ammar was a very admirable colleague

5.1.6 Single word followed by a complement in genitive

A single word maSdar/gerund or participle followed by a complement in genitive must be an NP.

maSdar/gerund:

(S (VP Tuliba طُلِبَ
 (PP min مِن
 (NP hu هُوَ))
 (NP-SBJ-1 <ixlA'u إخلاءُ
 (NP Almanozili المنزِلِ))
 (NP-OBJ-1 *)))

طُلِبَ مِنْهُ إِخْلَاءُ الْمَنْزِلِ

Tuliba min+hu <ixlA'+u Al+manozil+i
 was required from+him evacuation the+house
He was required to evacuate the house

(S (VP kAna كَانَ
 (VP yuSir~u يُصِرُّ
 (NP-SBJ *)
 (PP-CLR EalaY عَلَى
 (NP (NP qatoli قَتْلِ
 (NP him هِمَّ))
 >aw أَوْ
 (NP tahojiyri تَهْجِيرِ
 (NP him هِمَّ)))))))

كَانَ يَصِرُّ عَلَى قَتْلِهِمْ أَوْ تَهْجِيرِهِمْ

kAna+a + yu+Sir~+u + EalaY + qatol+i+him + >aw + tahojiyr+i+him
 was + insist + on + killing+them + or + exiling+them
He used to insist on killing them or forcing them to exile.

Active participle:

(S (VP kAna كَانَ
 (NP-SBJ *)
 (NP-PRD (NP rajulAF رَجُلًا
 (ADJP EAliya عَالِيَا
 (NP Alhim~api الْهَمَّةِ))))))

كَانَ رَجُلًا عَالِيَا الْهَمَّةِ

kAna + rajul+AF + EAliy+a + Al+him~ap+i
 was + man + high + will
He was a man with a strong will

Passive Participle:

(SBARQ (WHADVP-1 kayfa كَيْفَ)
 (S (VP yumokinu يُفَكِّنُ
 (SBAR-SBJ an أَنْ
 (S (VP yasotayqiZa يَسْتَيْقِظُ
 (NP-SBJ (NP \$aEobN شَعْبُ)
 (ADJP masoluwbu مَسْلُوبُ
 (NP Al<irAdap+i الإرَادَة))))))
 (ADVP-MNR-1 *T*))))
 ?)

كيف يمكن أن يستيقظ شعبٌ مَسْلُوبُ الإرادة؟

kayfa + yumokin+u + >an + yasotayqiZ+a + \$aEob+N + masoluwb+u + Al+<irAdap+i
 how + be possible + to + rise + people + deprived + the+will
How can people rise when they have no free will?

5.1.7 maSdar/gerund followed by an SBAR complement

If the maSdar/gerund is definite (carries Al-) or indefinite without tanween (in other words, in an iDafa position) and is followed by an SBAR complement, it is an NP (this rule does not apply when the maSdar/gerund is indefinite with tanween and does not apply for active participle and passive participles, which are verbal with SBAR complements, see section ????? 5.2.1.5 below).

Indefinite maSdar:

(S (VP rafaDat رَفَضَتْ
 (NP-SBJ AljamoEiy~apu الْجَمْعِيَّةُ)
 (NP-OBJ ta>okiyda تَأْكِيدَ
 (SBAR >an~a أَنْ
 (S (NP-TPC-1 AlmubArApa الْمُبَارَاةُ)
 (VP >uloEy+at أُلْعِيَتْ
 (NP-SBJ-1 *T*)
 (NP-OBJ-1 *)))))))

رفضت الجمعية تأكيد أن المباراة ألغيت

rafaD+at + Al+jamoEiy~ap+u + ta>okiyd+a + >an~a + Al+mubArAp+a + >uloGiy+at
 refused + the association + confirmation + that + the match + cancelled
The association refused to confirm that the match was cancelled

Definite maSdar:

(S (VP rafaDat رَفَضَتْ
 (NP-SBJ AljamoEiy~apu الْجَمْعِيَّةُ)
 (NP-OBJ ta>okiyda التَّأْكِيدَ
 (SBAR >an~a أَنْ
 (S (NP-TPC-1 AlmubArApa المَبَارَاةُ)
 (VP >uloEy+at أُلْعِيَتْ
 (NP-SBJ-1 *T*)
 (NP-OBJ-1 *)))))))
 رفضت الجمعية التأكيد أن المباراة ألغيت
 rafaD+at + Al+jamoEiy~ap+u + Al+ta>okiyd+a + >an~a + Al+mubArAp+a + >uloGiy+at
 refused + the association + the confirmation + that + the match + cancelled
The association refused to confirm that the match was cancelled

Note that in the case of definiteness, the maSdar is read as a noun not only because of the SBAR complement but mainly because of the Al- (rule ????? 5.1.2.).

Note also that in both cases, where the maSdar is definite *and* indefinite the SBAR following it is always in a complementation position with the NP heading the maSdar. (see section 3.2.1 in the guidelines about complements of nouns).

5.1.8 Single word followed by special 'li' PP

The active participle or maSodar/gerund can have a PP complement headed by the preposition 'li' that fills the argument role of a direct object (which would be in the accusative case if it were an NP). In most cases, this 'li' PP cannot occur as an argument with the same verb in a non-participle/maSodar form – for these verbs, the maSodar or participle followed by this 'li' PP will be annotated as an NP or ADJP. The presence of the 'li' PP forces the nominal reading for verbs of this kind.

However, there are some verbs that take a PP-CLR headed by the preposition 'li' in all verbal forms. If the 'li' PP is a PP-CLR in combination with the verb, the maSodar or participle of that verb in combination with the PP-CLR will be annotated as a VP. See section ????? for more information and examples of PP-CLR with maSodar/participles.

Finally, the 'li' PP could be a simple adverbial modifier of the verb, neither a PP-CLR nor a special 'li' that occurs only with maSodar or participle forms. In this case, the maSodar/participle will be annotated as an NP or ADJP. See section ????? for more information and examples of adverbial modifiers of maSodar/participles.

The various types of 'li' in combination with maSodar/participles are summarized below:

3 types of 'li' PP	Annotation of the maSodar or participle	Test to determine which type of 'li' PP	Guidelines section with more examples
PP-CLR	VP	<p>The PP-CLR is a possible complement of the verb, regardless of the morphological form of the verb (that is, the PP-CLR is possible with the maSodar/participle form of the verb, and also possible with the regular forms of the verb).</p> <p>Example with maSodar و لن يستطيعوا الدفاع عن أنفسهم عند تعرضهم لهجوم</p> <p>Example with regular verb و لن يستطيعوا الدفاع عن أنفسهم إن تعرضوا لهجوم</p>	???
PP adverbial	<p>NP or ADJP with single-word maSodar/participle;</p> <p>VP if there is other evidence for a verbal reading</p>	<p>The PP is optional or not obligatory with any form of the verb – it is a simple adverbial modifier.</p> <p>Example with maSodar كان خروجه كل صباح للبحث عن عمل إهانة له ولعائلته</p> <p>Example with regular verb لا يحب أن يخرج كل صباح للبحث عن عمل</p>	???
Special 'li' PP with maSodar/participle only (replacing direct object with regular verb)	NP or ADJP	<p>The 'li' PP is obligatory with the maSodar or participle form of the verb, but impossible with any other regular form of the verb.</p> <p>Example with maSodar قراءته لنتائج الانتخابات الرئاسية كانت في محلها</p> <p>Example with regular verb قرأ نتائج الانتخابات الرئاسية على وجه صحيح</p> <p>This complement MUST be an NP with any regular form of the verb.</p>	This section

Examples of PP-CLR 'li':

maSdar with PP-CLR:

(S wa و
 (VP (PRT lan لَنْ)
 yasotaTyEwA يَسْتَطِيعُوا
 (NP-SBJ *)
 (NP-OBJ (NP Al+difAEa الدِّفَاعُ)
 (PP Ean عَنْ
 (NP >anofusi أَنْفُسِ
 (NP him هُمْ)))))
 (NP-TMP Einda عِنْدَ
 (S-NOM (VP taEar~uDi تَعْرِضُ
 (NP-SBJ him هُمْ)
 (PP-CLR li لِي
 (NP hujuwM مُجُومُ))))))))
 و لن يستطيعوا الدفاع عن أنفسهم عند تعرضهم لهجوم

NB: need transliteration

and + not + they be able + the defense + on + selves their + at + exposure their + to + attack
And they will not be able to protect themselves if they get attacked.

Cf. regular verb with PP-CLR:

(S wa و
 (VP (PRT lan لَنْ)
 yasotaTyEwA يَسْتَطِيعُوا
 (NP-SBJ *)
 (NP-OBJ (NP Al+difAEa الدِّفَاعُ)
 (PP Ean عَنْ
 (NP <anofusi أَنْفُسِ
 (NP him هُمْ)))
 (SBAR-ADV >in إِنْ
 (S (VP taEar~aDwA تَعْرِضُوا
 (NP-SBJ *)
 (PP-CLR li لِي
 (NP hujuwM مُجُومُ))))))))
 و لن يستطيعوا الدفاع عن أنفسهم إن تعرضوا لهجوم

NB: need transliteration

and + not + they be able + the defense + on + selves their + if + (they) face + to + attack
And they will not be able to protect themselves if they get attacked.

Examples of simple adverbial PP 'li':

With maSdar = nominal:

(S (VP kAna كَانَ
 (NP-SBJ (NP xurwju خُرُوجُ
 (NP hu هُوَ))
 (NP-TMP kul~a كَلَّ
 (NP SabAHK صَبَاحِ))
 (PP-PRP li لِي
 (NP (NP lbaHovi اللَّبْحِ)
 (PP Ean عَنْ
 (NP Eamalk عَمَلِ)))))))
 (NP-PRD (NP <ihAnapF إِهَانَةً)
 (PP (PP la لِي
 (NP hu هُوَ))
 wa وَ
 (PP li لِي
 (NP EA}ilati عَائِلَتِي
 (NP hi هِيَ)))))))))))

كان خروجه كل صباح للبحث عن عمل إهانة له ولعائلته

NB: need transliteration

was + going out his + every + morning + for + looking + for + job + insult + for + him + and
+ for + family his

Going out every morning to look for a job is an insult for him and his family

With regular verb = VP:

```

(S (VP (PRT lA لا )
      yuHib~u يُحِبُّ
      (NP-SBJ * )
      (SBAR >an أن
        (S (VP yaxoruja يَخْرُجُ
              (NP-SBJ * )
              (NP-TMP kul~a كُلَّ
                (NP SabAHK ( صباح ) )
                (PP-PRP li لـ
                  (NP (NP lbaHovi لَبْحَثِ )
                    (PP Ean عن
                      (NP EamalK عَمَلِ )))))))))))

```

لا يحب أن يخرج كل صباح للبحث عن عمل

NB: need transliteration

not + like (he) + that + go out + every + morning + to + searching + for + job

He doesn't like to go out every morning to look for a job

Examples of the special 'li' PP that appears ONLY with maSdar/participle forms:

maSdar/gerund:

(NP (NP Hub~u حُبُّ
(NP nA نَا))
(PP li لِـ
(NP lTabiyEapi لَطَبِيْعَة)))

حُبُّنا للطبيعة

Hub~u+nA + li+l+TabiyEap+i
Love+our+ for+the+nature
Our love of nature

(NP (NP HadrN هَدْرُ
(PP li لِـ
(NP lwaqoti لَوْقَتِ)))

هدرٌ للوقت

Hadr+N + li+l+waqot+i
wasting + for+the+time
Waste of time

(S (NP-TPC-1 (NP qirA'atu قِرَاءَتُهُ
(NP hu هُوَ))
(PP li لِـ
(NP natA}ijzi نَتَائِجِ
(NP Al+<inotixAbAti الإِنْتِخَابَاتِ
Al+ri}Asy~api الرِّئَاسِيَّةِ)))))
(VP kAnat كَانَتْ
(NP-SBJ-1 *T*)
(PP-PRD fy فِي
(NP maHal~i مَحَلُّ
(NP hA هَا)))))

قراءته لنتائج الإنتخابات الرئاسية كانت في محلها

NB: need transliteration

Reading his + for + the + results + the election + the presidential + was + in + place its
His reading of the results of the presidential elections was right

Active participle:

(NP (NP Alduwalu الدُّوَلُ)
(ADJP Almusad~irapu المُصَدِّرَةُ
(PP li لِـ
(NP lnafti لَنَفْتِ))))

الدول المصدرة للنفط

Al+duwal+u + Al+muSad~irap+u + li+l+naftT+i
Countries + the+exporting + for+the+oil
The oil exporting countries

(S (NP-SBJ ha*A هَذَا)
 (NP-PRD (NP mawoqifN مَوْقِفٌ)
 (ADJP muviyrN مُثِيرٌ
 (PP li لـ
 (NP ljadali لِّجَدَلٍ)))))

هذا موقف مثير للجدل

ha*A + mawoqif+N + muviyr+N + li+l+jadal+i
 this + opinion + raising + for+the+questioning
This is a controversial opinion

Note that in all of the above examples, this special 'li' PP cannot occur with the regular form of the verb (i.e., any verbal form that is not a maSdar or participle). For this reason, the maSdar/participles that occur with this special 'li' PP are treebanked as nominal (either NP or ADJP), as above.

Verb cannot occur with special 'li' PP:

(S (VP qara>a قرأ
 (NP-SBJ *)
 (NP-OBJ natA}ija نتائج
 (NP Al+<inotixAbAti الانتخابات
 Al+ri}Asy~api ((الرئاسية))
 (PP-MNR EalY على
 (NP wajohK وجهه
 SaHyHK صحيح)))))

قرأ نتائج الانتخابات الرئاسية على وجه صحيح

NB: need transliteration

read he + results + (of) elections + presidential + on + manner + right
He read the results of the presidential elections in the right way

5.1.9 Adverbial modification

Any maSdar/gerund or participle not involved in one or more rules that force a verbal reading is annotated as nominal, even if it is modified by an adverbial.

maSdar/gerund:

```
(S (VP (PRT mA)
      kAna
      (NP-1 (NP juluwsu
              (NP hu))
            (PP-LOC fi
              (NP Al-bayoti))
            (NP-TMP SabAHAF
              wa
              masA'F))
      (VP yuroDiy
        (NP-SBJ-1 *T*)
        (NP-OBJ hi))))
```

ما كان جلوسه في البيت صباحا ومساء يرضيه

mA kAna juluwsu+hu fiy Al+bayoti SabAHAF wa masA'F yuroDiyhi
 not + was + his staying + at + home + morning + and + vevening + he+satisfy+him
He didn't like staying at home morning and evening

Active participle:

```
(NP (NP AlryiAHu الرياح )
     (ADJP AlqAdimapu القادمة
      (NP-TMP (NP ha*ihi هذه)
        (NP Al>ay~ama الأيام))
      (PP-DIR min من
        (NP Al$amAli الشمال))))
```

الرياح القادمة هذه الأيام من الشمال

AlriyAHu + AlqAdimapu + ha*ihi + Al>ay~Ama + min + Al\$amAli
 The winds + the coming + these + days + from + the North
The winds coming lately from North

Passive participle:

```
(NP (NP luEabu لُعِبَ
      (NP AlaTofAli الأطفال))
     (ADJP AlmaSonwEapu المصنوعة
      (PP-LOC fiy في
        (NP Al$yn الصين))))
```

لعب الأطفال المصنوعة في الصين

luEabu + Al>aTfAli + AlmasonwEapu + fy + Al\$yn
 toys + children + the manufactured + in + China
Children toys made in China

5.2 Verbal reading of gerunds and participles

The following rules force a verbal reading of gerunds and participles. These rules are not necessarily ordered, but if one applies all of the other rules are pre-empted.

5.2.1 Related argument in accusative

5.2.1.1 For maSdar/gerund

The maSdar/gerund acquires a verbal reading if it is (1) derived from a transitive or ditransitive verb and (2) is followed by a noun phrase that fills the position of a direct object, (3) when this noun phrase does not appear in a complementation position with the maSdar/gerunds. The maSdar has already a complement in genitive and the next noun phrase reads itself as the direct object of the maSdar in the regular verb form.

If the maSdar/gerund or participle is followed by an argument in accusative, that should force a verbal reading.

maSdar/gerund:

(S-NOM (VP rafoDu رَفَضَ
(NP-SBJ hu هُوَ)
(NP-OBJ AlmusAwamapa المُساوَمَة)))

رفضه المساومة

rafoDu+hu Al+musAwamapa
rejecting his + the bargaining
His rejecting for a bargaining.

(S-NOM (VP manHu مَنَحَ
(NP-SBJ *)
(NP-DTV hu هُوَ)
(NP-OBJ valAvapa ثَلَاثَة
(NP >ay~AmK أَيَّامَ)
(PP-PRP li لِ
(NP <ixlA' إِخْلَاءِ
(NP Almanzili الْمَنْزِلِ))))

منحه ثلاثة أيام لإخلاء المنزل

manoHu+hu + valAvapa + <ay~AmiK + li+<ikhlA'I + Almanozili
giving + him + three + days + for + deserting + the + house
Giving him three days to move out of the house

When annotated as verb, the direct genitive complement of the maSdar/gerund becomes either the subject (first example above) or the indirect object (second example above) of the verb, according to its distribution and according to context. It will always stand in the genitive case. The direct object is in the accusative case.

Note that in the Arabic Treebank all maSdar/gerunds with verbal reading head an S-NOM. On the distribution of S-NOM in this context, see section xxxx ?????.

See also section ????? 5.1.7 for maSdar/participles followed by a PP headed by 'li.'

5.2.1.2 For participles

Like maSdars, the active and passive participles derived from transitive and ditransitive verbs can be followed by an entity in accusative that fills the direct object position of the regular verb form. The presence of an accusative object after a participle forces a verbal reading.

Active participle:

```
(NP (NP Almudunu المَدُنْ)
  (SBAR (WHNP-1 0)
    (S (VP AlnATiqapu الناطقة
      (NP-SBJ-1 *T*)
      (NP-ADV kul~yAF كُلياً
        >aw أو
        juzo}y~AF جُزئياً)
      (NP-OBJ Allugapa اللُغة
        Alfaransy~iapa الفَرَنسِيَّة))))))
المدن الناطقة كلياً أو جزئياً اللغة الفرنسية
Al+mudunu + Al+nATiqapu + kul~yAF + aw + juzo{y~Af +Al+lugapa + Al+faransy~apa
The cities + the speaking + totally + or + partially + the+language + the French
The cities speaking totally or partially French
```

Passive participle:

```
(NP (NP Aldawolapu الدَوْلَةُ )
  (SBAR (WHNP-1 0 )
    (S (VP AlmamnuwHapu المَمْنُوخَةُ
      (NP-SBJ-1 *T* )
      (NP-DTV-1 * )
      (NP-OBJ <iEAnapF إِعَانَةٌ ) ) ) ) )
```

If the participle is to be treated as a VP as in the examples above and it is not in HAl position, it heads a relative or a free relative clause (SBAR or SBAR-NOM depending distribution rules. See section xxx ?????).

5.2.1.3 Accusative complement overrides non-verbal rules

Note that a number of examples in our corpora have combined typically non-verbal evidence with typically verbal evidence, in particular having an accusative complement. This is not frequent, but when it occurs, the general decision has been to follow the accusative case marking and therefore treat the structure as verbal. Individual instances of such examples should be examined and considered by the annotation group.

Accusative case overriding adjectival modification:

(S-NOM (VP القرار
(NP-SBJ الأمريكي)
(NP-OBJ (NP حاب الانس
(PP من
(NP العراق))))))

القرار الأمريكي الانسحاب من العراق

NB: need transliteration

The decision the American the withdrawing from Iraq

The American decision to withdraw from Iraq

Accusative case overriding plural form:

(S-NOM (VP مُحاولات
(NP-SBJ ها)
(NP-OBJ إجراء
(NP (NP مناقشات
(NP-ADV مع
(NP بعض
(NP الدول))))))

محاولاتها اجراء مناقشات مع بعض الدول

NB: need transliteration

Attempts her doing discussions with some the countries

Her attempts to have some discussions with some countries

Accusative overriding definite article

(S (VP نَرْجُو
(NP-SBJ *)
(S-NOM-OBJ (VP الأخذ
(NP-SBJ *)
(PP-CLR ب
(NP عين
(NP الاعتبار))))
(NP-OBJ زيادة
(NP الأجور))))))

نرجو الأخذ بعين الاعتبار زيادة الأجور

NB: need transliteration

Wish we the taking with eye the consideration increasing the salaries.

We wish that you take into consideration an increase in salary

5.2.2 Argument PPs (PP-CLR and PP-DTV and PP-DIR)

The maSdar/gerund, active participle or passive participle followed by a argument PP forces a verbal reading. PPs that are considered arguments are PP-CLR, PP-DTV, and PP-DIR when it occurs with a verb of motion that requires a directional complement.

Examples of each argument PP type are included below, but the structure is identical for all.

maSdar/gerund:

(S (VP <iHotafala إحتفل
(NP-SBJ Alfariyqu الفريق)
(PP-CLR bi ب
(S-NOM (VP fawzi فوز
(NP-SBJ hi ه)
(PP-CLR bi ب
(NP ka>osi كأس
(NP Al>aboTAl i الأبطال))))))))
إحتفل الفريق بفوزه بكأس الأبطال

<iHotafal+a + Al+fariyq+u + bi+fawz+i+hi + bi+ka>os+i + Al+>aboTAl+i
celebrate + the+team + with+winning+its + with+cup + the+champions
The team celebrated winning their champions cup

(S (VP هاجم
(NP-SBJ *)
(NP-OBJ العدو)
(S-PRP (VP دفاعا
(NP-SBJ*)
(PP-CLR عن
(NP النفس))))))

هاجم العدو دفاعا عن النفس

NB: need transliteration

Attacked the enemy protecting from the self
He attacked the enemy to protect himself

(S-NOM (VP توجه need transliteration and gloss
(NP-SBJ ه need transliteration and gloss)
(PP-DIR إلى need transliteration and gloss
(NP شمال need transliteration and gloss
(NP العراق need transliteration and gloss))))))
توجهه إلى شمال العراق

NB: Need transliteration here, and in the tree

Going his towards north the Iraq
His going towards north of Iraq

Active participle:

(NP (NP Almawoqifu الموقِف)
(SBAR (WHNP-1 0)
(S (VP AlmuEab~iru المعبر
(NP-SBJ-1 *T*)
(PP-CLR Ean عن
(NP ra}iyi رأي
(NP Al>aglabiy~api الأغلبية))))))))

الموقف المعبر عن رأي الأغلبية

Al+mawoqif+u Al+muEab~ir+u Ean ra}iy+i Al+>aglabiy~ap+i
the+attitude the+expressive on opinion the+majority
The attitude which shows the opinion of the majority

```

(S (NP-SBJ الجيش need transliteration and gloss)
  (SBAR-NOM-PRD (WHNP-1 0)
    (S (VP متوجه need transliteration and gloss)
      (NP-SBJ-1 *T*)
      (PP-DIR إلى need transliteration and gloss)
      (NP شمال need transliteration and gloss)
      (NP العراق need transliteration and
gloss))))))

```

الجيش متوجه إلى شمال العراق

NB: Need transliteration here, and in the tree

The army going to north the Iraq

The army is going to the north of Iraq

Passive participle:

```

(NP (NP (NP القانونُ AlqAnuwnu )
  (NP رَقْمُ raqomu )
  (NP 116 ) ) )
(SBAR (WHNP-1 0 )
  (S (VP المنصوص AlmanSuwSu
    (NP-SBJ-2 * )
    (PP-CLR عَليَ Ealay
      (NP-2 (NP هِ hi )
        (NP-1 *T* ))))))))

```

القانون رقم 116 المنصوص عليه

Al+qAnuwn+u raqom+u 116 Al+manSuwS+u Ealay+hi
 the+law number 116 the+mentioned on+it
Law number 116 which was mentioned

NB: Note that a passive participle with a verbal reading is always annotated exactly like its corresponding passive verb. See sections 4.6.4 and 4.6.5 for the treatment of passive verbs with only PP-CLR complements.

```

(NP (NP القروض need transliteration and gloss)
  (SBAR (WHNP-1 0)
    (S (VP الممنوحة need transliteration and gloss)
      (NP-SBJ-1 *T*)
      (NP-OBJ-1 *)
      (PP-DTV ل need transliteration and gloss)
      (NP البلدان الفقيرة need transliteration and
gloss))))))

```

القروض الممنوحة للبلدان الفقيرة

NB: Need transliteration here, and in the tree

The loans the given to the countries the poor

The loans given to the poor countries

```
(S (NP-SBJ انتقادات عديدة {inotiqaAd+At Eadiydap+N many criticisms)
  (SBAR-NOM-PRD (WHNP-1 0)
    (S (VP موجهة muwaj~ahap+N directed
      (NP-SBJ-1 *T*)
      (NP-OBJ-1 *)
      (PP-DIR إلى <ilaY to
        (NP o hi him ))))))))
```

انتقادات عديدة موجهة إليه

{inotiqaAd+At Eadiydap+N muwaj~ahap+N <ilaY hi
 Criticisms many directed to him
Many criticisms are directed to him

5.2.3 In HAl position

The participle in a HAl position, **even if it is a single word**, forces the verbal reading.

```
(S (VP xaraja خَرَجَ
  (NP-SBJ-1 zayodN زَيْدٌ )
  (S-MNR (VP DAHikAF ضاحكاً
    (NP-SBJ-1 * ) ) ) ) )
```

خرج زيد ضاحكاً

xaraj+a zayoduN + DAHik+AF
 left Zayd laughing
Zayd left, laughing

NB: whether the higher treebank node is S-ADV or S-MNR will depend on the interpretation of the relationship between the HAl clause and the main verb.

```
(S (VP xaraja خَرَجَ
  (NP-SBJ-1 * )
  (S-MNR (VP muqay~adAF مُقَيِّدٌ
    (NP-SBJ-1 * )
    (NP-OBJ-1 * ) ) ) ) )
```

خرج مقيداً

xaraja + muqay~adAF
 came out (he) + handcuffed
He came out handcuffed.

5.2.4 Argument in nominative

This rule applies mainly to passive and active participles, since the use of nominative after a maSdar is very rare in Arabic and restricted to very specific conditions (see below).

Passive participle:

(NP (NP AlmuqotaraHAtu المَقْتَرَحَاتُ)
 (SBAR (WHNP-1 0)
 (S (VP AlmaToluwbū المطلوبُ
 (NP-SBJ-2 <i>logA&u إِنْغَاؤُ
 (NP (NP hA هَا)
 (NP-1 *T*)))
 (NP-OBJ-2 *)))))

المقترحات المطلوبُ إِنْغَاؤُهَا

Al+muqotaraHAt+u Al+maToluwb+u <i>logA&+u+hA
 the+suggestions the+required refusal+its
The suggestions whose refusal is required

Active participle:

(S (NP-SBJ ha*A)
 (NP-PRD (NP rajulun)
 (SBAR (WHNP-1 0)
 (S (VP mutaqaal~ibatuN
 (NP-SBJ >aHwAlu
 (NP (NP hu)
 (NP-1 *T*)))))))))

هذا رجل متقلبة أحواله

ha*A rajulun mutaqaal~ibatuN >aHwAluhu
 this + man + changing + temper
This is a man with changing temper

NB: When the passive or active participle is followed by an argument in the nominative, the relative clause already has a subject on its own, and the trace of the WHNP in a participial relative clause that already has a nominative subject is usually adjoined with a resumptive pronoun.

The only time where the nominative is allowed after a maSdar is when the maSdar itself is indefinite with tanween. [Fischer, 2002, #203, p.112-113.]

(S (VP Eajibtu
 (NP-SBJ *)
 (PP-CLR min
 (S-NOM (VP DarbiK
 (NP-SBJ zayduN)
 (NP-OBJ EamrAF)))))))

عجبت من ضرب زيدٌ عمرا

Eajibtu + min + DarobK + zayoduN + EamorAN
 Surprised + from + hitting + zayd + Amr
 It surprises me that Zayd hit Amr

Other than that, every noun following a maSdar is in an >iDafa relationship with it and carries the genitive (independently from the verbal or nominal reading of the maSdar).

```

(S (VP Eajibtu
  (NP-SBJ * )
  (PP-CLR min
    (S-NOM (VP Darbi
      (NP-SBJ zaydiK)
      (NP-OBJ EamrAF))))))

```

عجبت من ضرب زيد عمرا

Eajibtu + min + DarobK + zayodiK + EamorAN
 Surprised + from + hitting + zayd + Amr
 It surprises me that Zayd hit Amr

All the following participles are nominal and NOT verbal except for the last ones ,where the SBAR or the Nominative arguments are complements rather than subjects.

<pre> (S (NP-TPC-1 القرار) (ADJP-PRD ممكن) (NP-SBJ إلغاء (NP (NP هـ) (NP-1 *T*))))) </pre>	<pre> (S (ADJP-PRD ممكن) (NP-SBJ إلغاء (NP القرار))) </pre>	<pre> (S (ADJP-PRD ممكن) (SBAR-SBJ أن (S (VP يلغى (NP-SBJ-1 القرار) (NP-OBJ-1 *)))))) </pre>
القرار ممكن إلغاؤه	ممكن إلغاء القرار	ممكن أن يلغى القرار
NB: Need transliteration here, and in the tree The decision possible repeal its <i>It is possible to repeal the decision</i>	NB: Need transliteration here, and in the tree Possible repeal the decision <i>It is possible to repeal the decision</i>	NB: Need transliteration here, and in the tree Possible that be repealed the decision <i>It is possible for the decision to be repealed</i>

<pre> (NP (NP القرار) (SBAR (WHNP-1 0) (S (ADJP-PRD الممكن) (NP-SBJ إلغاء (NP (NP هـ) (NP-1 *T*))))) </pre>	<pre> (NP (NP القرار) (SBAR (WHNP-1 0) (S (ADJP-PRD الممكن) (SBAR-SBJ أن (S (VP يلغى (NP-SBJ-1 *T*) (NP-OBJ-1 *)))))) </pre>
القرار الممكن إلغاؤه	القرار الممكن أن يلغى
NB: Need transliteration here, and in the tree The decision the possible repeal its <i>The decision which it is possible to repeal</i>	NB: Need transliteration here, and in the tree The decision the possible that is repealed <i>The decision that it is possible to repeal</i>

(S (NP-SBJ القرار) (SBAR-NOM-PRD (WHNP-2 0) (S (VP مطلوب (NP-SBJ-1 إلغاء (NP ه) (NP-2 *T*)) (NP-OBJ-1 *))))))	(S (NP-SBJ القرار) (SBAR-NOM-PRD (WHNP-1 0) (S (VP مطلوب (SBAR-SBJ-2 أن (S (VP يلغى (NP-SBJ-1 *T*)) (NP-OBJ-1 *) (NP-OBJ-2 *))))))
القرار مطلوب إلغاؤه NB: Need transliteration here, and in the tree The decision intended repeal its <i>The decision is intended to be repealed</i>	القرار مطلوب أن يلغى NB: Need transliteration here, and in the tree The decision intended that be repealed <i>The decision is intended to be repealed</i>

5.2.5 maSdar/gerund or Participle followed by an S or SBAR complement

An S or SBAR complement to a participle forces a verbal reading. This is true also for maSdars/gerunds that are indefinite and carry the tanween (indefinite not in an >iDafa structure). Note that this is different from definite maSdars that are either definite or indefinite without tanween, which are necessarily nominal with SBAR complements – see section ?????? 5.1.6 above.

Active participle:

(S (NP-SBJ (NP Alra'yu الزَّائِي)
 (SBAR (WHNP-1 0)
 (S (VP Alqa}ilu القائل
 (NP-SBJ-1 *T*)
 (SBAR >an~a أَنَّ
 (S (NP-TPC-2 Almar>apa المرأة)
 (VP (PRT لا)
 yumkinu يُمكن
 (SBAR-SBJ >an أن
 (S (VP tu\$Arika تُشارك
 (NP-SBJ-2 *T*)
 (NP-OBJ Alrajula الرَّجُل)
 (PP-CLR fiy في
 (NP AlHuquwqi الحُقُوقِ
 wa و
 AlwAjibAti الواجِبَاتِ
 (NP-PRD ra>oyN زَائِي
 xATi}N خَاطِيُ))))))))))))
 الرأي القائل أن المرأة لا يمكن أن تشارك الرجل فل الحقوق والواجبات رأي خاطئ
 Al+ra}iy+u Al+qA}il+u >An~a Al+mar>ap+a lA+yumokin+u >an tu\$Arik+a Al+rajul+a fiy
 + Al+Huquwqi wa+Al+wAjibAt+i ra>y+N xATi}+N
 the+opinion the+saying that the+woman not+be possible to share the+man in the+rights
 and+the+duties opinion false
The opinion which says that women can not share rights and duties with men is false.

Passive participle:

(S (NP-SBJ (NP ha*A هَذَا)
 (NP Alrajulu الرَّجُلُ))
 (SBAR-NOM-PRD (WHNP-1 0)
 (S (VP mut~ahamN مُتَّهَمٌ
 (NP-SBJ-1 *T*)
 (NP-OBJ-1 *)
 (SBAR >an~a أَنَّ
 (S (NP-TPC-2 hu هُوَ)
 (VP qatala قَتَلَ
 (NP-SBJ-2 *T*)
 (NP-OBJ zawojapa زَوْجَةَ
 (NP hu هُوَ))
 (NP-TMP yawoma يَوْمَ
 (NP Eydi عيد
 (NP mylAdi ميلاد
 (NP ha هَا)))))))))))
 هذا الرجل متهم أنه قتل زوجته يوم عيد ميلادها

ha*A + Al+rajulu + mut~ahamuN + >an~a+hu + qatala + zawojapa hu + yawoma + Eydi + milAfi+hA

this + man + accused + that + killed + wife his + day (of) + celebration + birth

This man is accused of having killed his wife on her birthday.

5.2.6 Auxiliary kAwon as a maSdar

When kAwon appears as a maSdar with a VP complement, it is annotated as an S-NOM with the same structure as usual below:

(S-NOM (VP كَوْنُ
(NP-1 (NP مَدِينَةُ
(NP الْقُدْسُ)
(ADJP الْمُحْتَلَّةُ)
(VP أُحْتُلتْ
(NP-SBJ-1 *T*)
(NP-OBJ-1 *)
(NP-TMP عام
(NP (سَبْعَةَ وَسِتِّينَ))))))))

NB: Need transliteration here, and in the tree

Being city the Jerusalem the colonized was colonized year seven and sixty

Given that the colonolized Jerusalem was colonized in 1967

See section ??? for the treatment of kAna as an auxiliary.

See section ???? for the treatment of kAna as an auxiliary.

5.2.7 Annotation of verbal readings

5.2.7.1 maSdar/gerund is S-NOM

With any verbal maSdar/gerund, the top node should be S-NOM followed by a VP node.

The S-NOM node should have additional function tags added as needed – so, S-NOM-SBJ if the maSdar/gerund is the subject of the sentence, or S-NOM-OBJ if the maSdar/gerund is the direct object of the main verb, etc.

```
(S (VP kAna
  (S-NOM-SBJ (VP <iqotiHAMu
    (NP-SBJ hu)
    (NP-OBJ Albalada)))
  (NP-PRD xaTa'AF
    siyAsiy~AF
    qAtilAF)))
كان اقتحامه البلد خطأ سياسيا قاتلا
kAna <iqotiHAMuhu Albalada xaTa'AF siyAsiy~AF qAtilAF
was + invasion+his + the+country + mistake + political + deadly
His invasion of another country was a deadly political mistake
```

NB: Don't forget to put in or mark all of the internal complements of the verbal maSdar/gerund – subject, direct object, PP-CLR, etc.

```
(S (VP kAna
  (S-NOM-SBJ (VP Ai qotiHAMu
    (NP-SBJ hu)
    (NP-OBJ Albalada)))
  (NP-PRD xaTa'AF
    siyAsiy~AF
    qAtilAF)))
كان اقتحامه البلد خطأ سياسيا قاتلا
kAna Ai qotiHAMuhu Albalada xaTa'AF siyAsiy~AF qAtilAF
was + invasion+his + the+country + mistake + political + deadly
His invasion of another country was a deadly political mistake
```

Although the subject of the maSdar/gerund will usually be in the genitive case, it is still marked –SBJ.

5.2.7.2 Participles are SBAR or adverbial S

5.2.7.2.1 Participial relative clauses

If the participle is to be treated as a VP according to the sections above, the participle should be annotated as heading a relative clause, and an empty relative pronoun needs to be created and traced accordingly in the relative clause. The relative clause is annotated with an empty relative pronoun whether the participle is definite with Al- or indefinite.

Participle with Al:

(NP (NP Almawoqifu الموقِفُ)
 (SBAR (WHNP-1 0)
 (S (VP AlmuEab~iru المعبر
 (NP-SBJ-1 *T*)
 (PP-CLR Ean عن
 (NP ra}iyi رأي
 (NP Al>aglabiy~api الأغلبية))))))))
 الموقف المعبر عن رأي الأغلبية

Almawoqifu + AlmuEab~iru + Ean ra}iyi + Al>aglabiy~api
 the position + the expressing + on + opinion + the majority
The position which shows the opinion of the majority

Participle without Al:

(S (NP-SBJ (NP ha*A هذا)
 (NP Alrajulu الرَّجُلُ))
 (SBAR-NOM-PRD (WHNP-1 0)
 (S (VP mut~ahamN مُتَّهَمُ
 (NP-SBJ-1 *T*)
 (NP-OBJ-1 *)
 (SBAR >an~a أَنَّ
 (S (NP-TPC-2 hu هُوَ)
 (VP qatala قَتَلَ
 (NP-SBJ-2 *T*)
 (NP-OBJ zawojapa زَوْجَتَهُ
 (NP hu هُوَ))
 (NP-TMP yawoma يَوْمَ
 (NP Eydi عيد
 (NP mylAdi ميلاد
 (NP ha هُوَ))))))))))
 هذا الرجل متهم أنه قتل زوجته يوم عيد ميلادها
 ha*A + Al+rajulu + mut~ahamN + >an~a+hu + qatala + zawojapa hu + yawoma + Eydi +
 milAfi+hA
 this + man + accused + that + killed + wife his + day (of) + celebration + birth
This man is accused of having killed his wife on her birthday.

5.2.7.2.2 S-ADV vs. S-MNR (HA1 and adverbial participial clauses)

If the active or passive participles are in a HA1 position and are treated as verbal, the adverbial HA1 S modifying the main verb takes a dash tag -ADV if the modification is general(e.g. when it describes the state of the subject or the object when the action happened) and -MNR if the adverbial modification gives us information on the manner in which the verb of the main clause happened.

```
(S (VP xaraja خَرَجَ
  (NP-SBJ-1 zayodN زَيْدُ )
  (S-ADV (VP DAHikAF ضاحكاً
    (NP-SBJ-1 *) ) ) ) ) )
```

خرج زيد ضاحكاً

xaraja + zayoduN + DAHikAF
 left Zayd laughing
Zayd left, laughing

```
(S (VP خرج
  (NP-SBJ-1 *)
  (S-MNR (VP مهرولا
    (NP-SBJ-1 *) ) ) ) ) )
```

خرج مهرولا

NB: Need transliteration here, and in the tree

Went out he running
He went out running

5.2.7.2.3 Passive participles with passive trace

Don't forget to put the passive trace in! Passive participles are annotated in the same way as any other passive verb.

Regular passive:

```
(S (VP xaraja خَرَجَ
  (NP-SBJ-1 * )
  (S-MNR (VP muqay~adAF مُقَيِّدُ
    (NP-SBJ-1 *)
    (NP-OBJ-1 *) ) ) ) ) )
```

خرج مقيداً

xaraja + muqay~adAF
 came out (he) + handcuffed
He came out handcuffed.

- Regular passive with the relative pronoun traced as subject of the relative clause:

```
(S (VP Yajibu يجب
  (NP-SBJ (NP AlqyAmu القيام )
    (PP bi ب
      (NP kul~i كُلّ
        (NP (NP AltaHoDirApi التحضيرات )
          (SBAR (WHNP-1 0 )
            (S (VP Almuqar~arapi المقررة
              (NP-SBJ-1 *T* )
              (NP-OBJ-1 * )
              (PP-PRP li ل
                (NP lqim~api لقيمة
                  )))))))))))
```

يجب القيام بكل التحضيرات المقررة للقيمة

Yajibu Al qyAm+u bi+kul~i Al+taHoDirAt+I Al+muqar~arap+i li+l+qim~ap+i

Is required undertaking of/with+the+preparation the+decided for+the+summit

Preparations decided for the summit need to be made (immediately)

- Regular passive with a different subject and the relative pronoun traced with the matching resumptive pronoun:

```
(NP (NP AlmuqotaraHAtu المقترحات )
  (SBAR (WHNP-1 0)
    (S (VP AlmaToluwbu المطلوب
      (NP-SBJ-2 <ilogA&u إلغاءُ
        (NP (NP hA ها )
          (NP-1 *T* ) ) )
        (NP-OBJ-2 * ) ) ) ) )
```

المقترحات المطلوب إلغاؤها

Al+muqotaraHAAt+u Al+maToluwb+u <ilogA&+u+hA

the+suggestions the+required refusal+its

The suggestions whose refusal is required

Ditransitive passive:

(S wa وَ
 (VP tara>~asa تَرَأس
 (NP-SBJ fArisu فَارِسُ)
 (NP-OBJ (NP <ijotimAEAF إجْتِمَاعاً)
 (PP li لـ
 (NP (NP Allajonapi اللّجْنة
 AlwizAriyapi الوِزارِيةِ)
 (SBAR (WHNP-1 0)
 (S (VP Almukal~afapi المِكلَفَة
 (NP-SBJ-1 *T*)
 (NP-DTV-1 *)
 (NP-OBJ darosa دَرَسَ
 (NP ma\$oruwEi مَشْرُوع
 (NP qAnuwni قَانُون
 (NP AlSafaqAti الصّفَقَاتِ
)))))))

وترأس فارس إجتماعاً للجنة الوزارية المكلفة درسَ مشروع قانون الصفقات
 wa+tara>~as+a fAris+u <ijotimAE+AF li+Al+lajonap+i Al+wizAriyap+i
 Al+mukal~afap+i daros+a ma\$oruwE+i qAnuwn+a Al+SafaqAt+i
*Faris led a meeting for the community of the ministry which is charged to study the
 deals law project.*

Passive with CLR:

(NP (NP (NP AlqAnuwnu القانونُ)
 (NP raqomu رَقْمُ
 (NP 116)))
 (SBAR (WHNP-1 0)
 (S (VP AlmanSuwSu المنصُوصُ
 (NP-SBJ-2 *)
 (PP-CLR Ealay عَلي
 (NP-2 (NP hi هـ)
 (NP-1 *T*)))))))

القانون رقم 116 المنصوص عليه
 Al+qAnuwn+u raqom+u 116 Al+manSuwS+u Ealay+hi
 the+law number 116 the+mentioned on+it
Law number 116 which was mentioned

See also section ????? on passive verbs with PP-CLR complements in general.

5.3 Distribution of S, S-NOM, S-ADV, NP, ADJP

The use of S, S-NOM, S-ADV, S-MNR, SBAR-NOM, NP and ADJP for gerunds and participles is purely distributional. This distribution assumes that you already know whether the word is a verb or a noun/adjective.

- **NP or ADJP** with the appropriate function tags whenever the word is not a verb. Once you know that the word is a noun or an adjective, all of the usual rules about nouns and adjectives apply.

If the word is a verb, use one of the following:

- **S-NOM** when the verbal gerund is in the following positions

1. the subject of a sentence (S-NOM-SBJ)

Example:

```
(S (S-NOM-TPC-1 (VP manoHu مُنَحُ
      (NP-SBJ * )
      (NP-DTV Al>amyri الأمير
      AlhAribi الهارب)
      (NP-OBJ (NP jawAza جواز
      (NP safarK سفر))
      (ADJP dyblwmAsy~AF (ديبلوماسياً))))))
(VP >avAra أثار
  (S-NOM-SBJ-1 *T*)
  (NP-OBJ (NP Al+Eadyda العديد)
    (PP mina مِنْ
      (NP Al<inotiqaAdAti الانتقادات))))))
منحُ الأمير الهارب جواز سفر دبلوماسياً أثار العديد من الانتقادات
giving + the prince + the escaping + permit + (of) travel + diplomatic + raised + the
numerous + of + criticisms
Giving the escaping prince a diplomatic passport has raised a great deal of criticism
```

2. the direct object of a verb (S-NOM-OBJ)

(N.B. This is different from the English Treebank, where all gerund complements of the verb were done as S.)

Example:

(S (VP rafaDat رَفَضَتْ
 (NP-SBJ Al+suluTAtu السُّلْطَاتُ
 (S-NOM-OBJ (VP manoHa مَنَحَ
 (NP-SBJ *)
 (NP-DTV Al>amiyri الأمير
 AlhAribi الهَارِبِ)
 (NP-OBJ (NP jawAza جَوَازَ
 (NP safarK سَفَرِ))
 (ADJP dyblwmAsy~AF (ديبلوماسيَا
 رفضت السلطات منح الأمير الهارب جواز سفر دبلوماسياً
 rejected + the authorities + giving + the escaping prince + permit + (of) travel +
 diplomatic
The authorities refused to give the escaping prince a diplomatic passport

3. the object of a preposition

(S wa وَ
 (VP Eal~aqa عُلِقَ
 (NP-SBJ *)
 (PP-CLR EalY على
 (S-NOM (VP liqotirAHi اقْتَرَحَ
 (NP-SBJ ramsfyld رَمْسْفِيلْدُ
 (NP-OBJ (NP na\$ora نَشْرَ
 (NP nizAmi نظام
 (NP murAqabapK (مُرَاقِبَة
 (PP-LOC EalY على
 (NP xaT~i خَطَّ
 Alhudonapi (الْهُدْنَة
 وعلق على اقتراح رمسفيلد نشر نظام مراقبة على خط الهدنة
 wa + Eal~aqa + EalaY + liqotirAHi + Rumsfeld + na\$ora + niZAmi + murAqabapK +
 EalaY + xaT~I + Alhudnapi
 and + commented (he) + on proposing + Rumsfeld + spreading + system + surveillance +
 at + line + the armistice
And he commented on Rumsfeld's proposal to set a surveillance system on the armistice line

4. when necessary, for coordination with other NPs

(S (VP kAnwA كانوا
 (VP yasotamid~wna يَسْتَمِدُّونَ
 (NP-SBJ *)
 (NP-OBJ qw~ata قُوَّةُ
 (NP hum هُمْ))
 (PP min مِنْ
 (NP (NP Eazymapi عَزِيمَةُ
 (NP raswli رَسُولِ
 (NP Al+lahi اللَّهِ)))
 wa وَ
 (NP Sabri صَبْرِ
 (NP hi هِ))
 wa وَ
 (NP jihAdi
 (NP hi هِ))
 wa وَ
 (S-NOM (VP rafoDi رَفُوضِ
 (NP-SBJ hi هِ)
 (NP-OBJ Al+musAwamapa الْمُسَاوَمَةِ
 wa وَ
 Al+tanAzula التَّنَازُلِ))))))
 وكانوا يستمدون قوتهم من عزيمة رسول الله وصبره وجهاده ورفضه المساومة والتنازل
 and + were + obtain + strength their + from + determination + (of) messenger + (of)
 God + and patience his+ and fighting his + and refusal his + the bargaining + and
 the resignation
*And they used to strive for strength from the prophet's determination, his patience,
 his fighting and his refusal to bargain and to give up*

5. complement of a noun

(S wa وَ
 (VP Tamo>ana طَمَأَنَّ
 (NP-SBJ bw\$ بوش)
 (NP-OBJ Al+SaHafyyna (الصحفيين)
 (NP-TMP qabola قَبْلَ
 (S-NOM (VP mugAdarati مُغَادَرَتِهِ
 (NP-SBJ hi هِ)
 (NP-OBJ Al+bayota الْبَيْتِ
 Al+>aboyaDa الْأَبْيَضِ))))))
 وطمأن بوش الصحفيين قبل مغادرته البيت الأبيض
 and + reassured + Bush + the journalists + prior to + leaving his + the house + the
 white
And Bush reassured the journalists before he left the White House

- **S-ADV or -MNR, -TMP, -LOC, -PRP**, etc. when the gerund or the participle is in an adverbial position, modifying the VP or the predicate. Null subjects of S verbal gerunds can be coindexed to another NP in the sentence if they have a coreferenced interpretation.
- **SBAR-NOM** when the verbal participle is in the following positions

following the nullrelative pronoun and when filling a nominal position

```
(S (NP-TPC-1 (NP Alqay~imwna الْقَائِمُونَ)
  (PP EalY عَلَى
    (NP Alma$orwEi (المشروع))))
  (VP kAnuWA كانوا
    (NP-SBJ-1 *T* )
    (PP-PRD mina مِنْ
      (SBAR-NOM (WHNP-4 0)
        (S (VP Alma$huwdi المشهود
          (NP-SBJ-3 *)
          (PP-CLR la لَ
            (NP-3 (NP hum هُمْ)
              (NP-4 *T*))
            (PP bi بِ
              (NP waTany~ati وَطَنِيَّةِ
                (NP him هِم))))
              القيمون على المشروع هم من المشهود لهم بوطنيتهم
            Alqay~ymuwna + EalY + Alma$oruwEi + kAnwA + min + Alma$ohuwdi + lahum +
            biwaTany~atihim
            The supervisors + on + the project + were + from + the + witnessed + for + them + for
            nationalism their
            The people who are supervising the project are considered among those who are known for
            their nationalism
```

- **SBAR** when the verbal participle is in the following positions

following a (WHNP 0) null relative pronoun and when modifying (and adjoined to) another noun phrase

```
(NP (NP Almawoqifu الْمُؤَقِفُ )
  (SBAR (WHNP-1 0 )
    (S (VP AlmuEab~iru المعبر
      (NP-SBJ-1 *T* )
      (PP-CLR Ean عَنْ
        (NP ra}iyi رأي
          (NP Al>aglabiy~api (الأغلبية))))))
    الموقف المعبر عن رأي الأغلبية
    Al+mawoqif+u Al+muEab~ir+u Ean ra}iy+i Al+>aglabiy~ap+i
    the+attitude the+expressive on opinion the+majority
    The attitude which shows the opinion of the majority
```

S-NOM (like SBAR-NOM free relatives, see section ?????) is treated as an NP if it is in coordination or apposition with another NP. That is to say, the higher adjunction level is NP, as in example 4. above.

6 Coordination

Coordination has the same structure at all phrase levels, regardless of node label.

6.1 Flat

Single words tokens that are coordinated only with other single words or tokens are annotated as flat structures.

6.2 Coordination as adjunction

The presence of a multi-word or multi-token conjunct in a coordination structure results in the annotation of that coordination as Chomsky adjunction.

Coordination done as adjunction (Z (Z) and (Z)) has the same structure at all phrase levels.

6.3 UCP

When constituents of different types are coordinated, the outer coordination-level node label is UCP (Unlike Coordinated Phrase). Any shared function tags are put on the UCP label, and not on the lower labels.

See section ????? on UCP for examples, etc. (2.2.1.3.3 UCP (Unlike Coordinated Phrase)).

6.4 Right node raising (*RNR*)

See section ????? (2.6.1) on Discontinuous Constituents/Rightward Movement.

- RNR is not used when the coordinated elements are modified by the same element in an adjunction structure, e.g.,

```
(NP (NP (NP خروج
          (NP هم )
        و
        (NP دخول
          (NP هم ) )
        (PP-TMP في
          (NP المساء ) ) ) ) )
```

خروجهم ودخولهم في المساء

NB: Need transliteration here, and in the tree
 Going out their and coming in their in the evening
Their going out and coming in in the evening

- RNR is not used when one word nouns are coordinated and have the same complement, e.g.,

(NP بين
 (NP عشرة
 و
 عشرين
 (NP رجلا)))

بين عشرة وعشرين رجلا

NB: Need transliteration here, and in the tree

Between ten and twenty men

Between ten and twenty men

- RNR is used when multi word conjuncts are coordinated and have the same complement, e.g.,

(NP (NP وزير
 (NP خارجية
 (NP-1 *RNR*)))
 و
 (NP وزير
 (NP داخلية
 (NP-1 *RNR*)))
 (NP-1 لبنان)))

وزير خارجية ووزير داخلية لبنان

NB: Need transliteration here, and in the tree

Minister foreign affairs and minister interior Lebanon

Lebanese Ministers of Foreign Affairs and the Interior

(NP بين
 (NP (NP واحد
 و
 عشرين
 (NP-1 *RNR*))
 و
 (NP خمسة
 و
 عشرين
 (NP-1 *RNR*))
 (NP-1 رجلا)))

بين واحد و عشرين و خمسة و- عشرين رجلا

NB: Need transliteration here, and in the tree

Between one and twenty and five and twenty men

Between twenty-one and twenty-five men

6.5 Gapping/ VP template gapping

Gapping is the only construction in the Arabic Treebank that allows VP coordination.

The technicalities of gapping coreference are different in the Arabic Treebank from the original Penn Treebank, although gapping coreference is done this way in the current English Treebanks as well.

All indices are on the node label itself, and gapping co-reference is shown with '=' on both the template and the filler node labels.

```
(S (VP (VP -fAz+a::فاز::win/be_victorious+he/it_[verb]
  (PP fiy-::في::in
    (NP -hi::ه::it/him ))
  (NP-SBJ=1 (NP Al+biriyTAniy~+u::البريطاني::the+British+[def.nom.])
    (NP Ady::ادي::Ed
      <iyrfAyn::إيرفاين::Irvine ))
  (NP-TMP=2 EAm+a::عام::year+[def.acc.]
    ,::,::nogloss
    (NP 1999::1999::nogloss )))
  wa-::و::and
  (VP (NP-SBJ=1 -$uwmAxar::شوماخر::Schumacher )
    (NP-TMP=2 EAm+ayo::عامي::year+two_[acc.]
      (NP 2000::2000::nogloss
        wa-::و::and
        .2001::.2001::nogloss ))))
```

وفاز فيه البريطاني ادي إيرفاين عام وشوماخر عامي 21999

fAz+a + fiy+hi + Al+biriyTAniy~+u + >adiy + <iyrfAyn + EAm+a + 1999 + wa+\$uwmAxar
+ EAmay+o + 2000 + wa+2001

won + in+it + the+British + Eddie + Irvine + year + 1999 + and+Schumacher + two years +
2000 + and+2001

In which the Brit, Eddie Irvine, won in 1999 and Schumacher in 2000 and 2001

See section ????? (2.5.3) on Gapping (VP Template Gapping).

Gapping is the only construction in the Arabic Treebank that allows VP coordination.

```
(S (VP (VP وضع
  (NP-SBJ=1 *)
  (NP-OBJ=2 العصا
  (PP-LOC=3 الى
    (NP جانب
      (NP ه ) ) ) )
  و
  (VP (NP-SBJ=1 *)
    (NP-OBJ=2 البندقية
    (PP-LOC=3 على
      (NP ركبي
        (NP ه ) ) ) ) ) ) ) ) )
```

وضع العصا إلى جانبه والبندقية على ركبتيه

NB: Need transliteration here, and in the tree

Put he the stick on side his and the rifle on knees his

He put the stick on his side and the rifle on his knees

See section ????? (2.5.3) on Gapping (VP Template Gapping).

6.6 Multi-word conjunctions/CONJP

See section ????? (2.2.1.2.10) on CONJP.

6.7 Arabic coordination without conjunction

Arabic allows coordination without an overt conjunction (or separating punctuation), particularly in formal or classical contexts.

The Arabic Treebank, however, requires an overt conjunction for S-level coordination in nearly all cases, but does allow for Ss to be coordinated without a conjunction if they are in exactly the right context, and if it is necessary for the construction of the tree.

The Arabic Treebank allows the coordination of other phrases without a conjunction if they occur. However, since they are essentially indistinguishable from apposition, both in structure and often in interpretation, they will be annotated as adjunction (as both apposition and coordination are annotated as adjunction).

6.8 NAC

See section ????? on NAC (2.2.1.3.2 NAC (Not a Constituent)).

6.9 Initial wa, other initial conjunction

The “wa-“ or other conjunction that frequently begins a sentence in Arabic is treated as a discourse marker, and as such it is not treated as coordination when it is the first word in a sentence. This sentence-initial conjunction is placed inside the S (or inside the S of the first conjunct, if the sentence is later actually coordinated).

See section ?????? on initial wa.

7 Tricky Constructions

7.1 Tokens with more than one possible part-of-speech value

7.1.1 >ay/>ay~ (أَي/أَي)

NB: the table doesn't actually fit in this document – too tall...

Arabic 317 of 375	POS Tag	Syntactic /Semantic Context	TB (NOT AVAILABLE)
1. أَيُّ	VOC_PART Hey	It is used in vocative constructions	(NP (PRT >ayo أَيُّ oh) muHam~ad مُحَمَّد Mohamed) !أَيُّ مُحَمَّد >ay muHamad! Hey Mohamed!
2. أَيُّ	PART That is to say	It provides further explanation or reformulation of a previous statement	(SBAR (PRT >ayo أَيُّ that is to say) >an~a أَنْ that (S (NP(NP hA*ih+hi هَذِهِ this) (NP Al+duwal+i الدُول the countries))...أَيُّ أَنْ هَذِهِ الدُول >ay >an~a h*ihhi Al+duwal That is to say that these countries...
3. أَيُّ	INTERR OG_PRON What Which	In direct questions	(SBARQ (WHPP-1 mino مِنْ from (WHNP >ay~i أَيُّ which (NP balad+K بَلَد country))) (S (NP-SBJ >anota أَنْتَ you) (PP-PRD-1 *T*)) ?) min >ay~i baladK >anta From which country you Which country do you come from?
4. أَيُّ	NOUN_QUANT Any	It is used with one dependent complement.	(S (VP layosa لَيْسَ (PP-PRD la لَ to (NP humA هُمَا them)) (NP-SBJ >ay~u أَيُّ any (NP AisotixodamAt+K استخدامات uses Easokariy~ap+K عَسْكَرِيَّة military)))) لَيْسَ لَهُمَا أَيُّ اسْتِخْدَامَاتٍ عَسْكَرِيَّةٍ laysa la+humA >ay~u <istixodAmAtK Easokariy~apK Not be to them any uses military They don't have any military uses.
5. أَيُّ	REL_PRON Whoever	It introduces indirect questions	(S (SBAR-NOM-TPC-1 (WHNP-2 >ay~u أَيُّ whichever (NP hum هُمْ their)) (S(VP qAbalot+a قَابَلْتُ met (NP-SBJ *) (NP-OBJ-2 *T*) (NP-ADV >aw~al+F أَوَّلًا first)))) (VP >ukal~im+u أَكَلِمُ speak I (NP-SBJ *) (NP-OBJ (NP hu هُوَ him) (NP-1 *T*)) (PP fiy فِي in (NP(NP h*A هَذَا this) (NP Al+>amor+i الأمر the matter)))) أَيُّهُم قَابَلْتُ أَوَّلًا أَكَلِمُهُ فِي هَذَا الْأَمْرِ >ay~u+hum qAbaltu >aw~alAF >ukal~imu+hu fiy h*A Al>amr Whoever their met first speak I him in this the matter Whoever I meet first, I will talk to him about this matter.
6. أَيَّا >ay~AF	REL_PRON Whatever	Only in indefinite accusative where it can be replaced by nouns of conditions such as mahmA	(S (VP <in~a إِنَّ indeed (NP-SBJ Al+>azamAta الْأَزْمَات the crisis Al+duwaliy~apa الدَّوَلِيَّة the international) (SBAR- ADV (WHNP-1 >ay~F أَيَّا whatever) (S(VP takuno تَكُنْ was (NP-SBJ darajapu دَرَجَة degree (NP taEqiydi تَعْقِيد complexity (NP hA هَا its))) (NP-PRD-1 *T*))))) (S(VP yumokinu يُمَكِّنْ was possible (SBAR-SBJ >ano أَنْ that (S(VP tuHal~a تُحَلْ be solved it (NP-SBJ-1 *) (NP-OBJ-1 *T*))))) إِنَّ الْأَزْمَات الدَّوَلِيَّة أَيَّا تَكُن دَرَجَة تَعْقِيدِهَا يُمَكِّنُ أَنْ تُحَلَّ <in~a Al+>azamAti Al+duwaliy~apa >ay~F takun darajapu taEqiydiH yumkinu >an tuHal~a Indeed the crises the international whichever was degree complexity its was possible that was solved it Whatever is the complexity of some international crises , it can be solved.

7.1.2 xalA, EadA, Ha\$A

1. EadA, xalA, and HA\$A are prepositions and annotated as PREP in the following cases:

- When they assign the genitive case to their complement

(S (VP jA'a جاء came he
(NP-SBJ AljamiyEu الجميع the all)
(PP EadA عدا except
(NP zayodK زيد Zayd))))

جاء الجميع عدا زيد

NB: Need transliteration

Came he the all except Zayd

Everybody came except Zayd

- When the case ending is not shown, EadA, xalA, and HA\$A are set as prepositions by default.

(S (VP jA'a جاء came he
(NP-SBJ AljamiyEu الجميع the all)
(PP EadA عدا except
(NP hu هُ it))))

جاء الجميع عدا ه

jA'a + AljamiyEu + EadA + hu

came + the all + except + him/it

Everybody came except him

(S (VP jA'a جاء
(NP-SBJ AljamiyEu الجميع
(PP EadA عدا
(NP AlmuEal~imAti المُعلِّمات))))

جاء الجميع عدا المُعلِّمات

jA'a + AljamiyEu + EadA + Al+muEal~imAti

came + the all + except+the + teachers

Everybody came except the teachers.

2. EadA, xalA, and HA\$A are non-inflectional verbs and annotated as VERB in the following cases:

- When they assign the accusative case to their complement. The accusative has to be marked graphemically via the use of the Alif of the indefinite accusative form.

(S (VP jA'a جاء came it
 (NP-SBJ AljamyEu الجميع the all)
 (S-ADV (VP EadA عدا except
 (NP-SBJ *)
 (NP-OBJ zayodAF زَيْدًا Zayd)))))

جاء الجميع عدا زيدًا

jA'a + AljamyEu + EadA + zayodF
 came + the all + except + Zayd
Everybody came except Zayd

- When they are preceded by mA, which is considered as a relative pronoun. The relative clause modifies any NP constituent of the matrix sentence and not its verb.

(S (VP jA'a جاء came
 (NP-SBJ (NP AljamyEu الجميع the all)
 (SBAR (WHNP-1 mA ما who)
 (S (VP EadA عدا except
 (NP-SBJ-1 *T*)
 (NP-OBJ zayodAF زَيْدًا Zayd))))))

جاء الجميع ما عدا زيدًا

jA'a + AljamyEu + mA + EadA + zayodAF
 came + the all + what + except + Zayd
Everybody came except Zayd

- In cases where the noun and the relative clause introduced by mA are separated by other constituents, the relative clause should be ICHed with the noun it modifies as in the example below.

(S (NP-TPC-1 (NP kul~u كُلُّ all
 (NP Al+madoEw~yna المدعوين the guests))
 (SBAR-3 *ICH*))
 (VP HaDarwA حضروا came
 (NP-SBJ-1 *T*)
 (SBAR-3 (WHNP-2 mA ما who)
 (S (VP EadA عدا except
 (NP-SBJ-2 *T*)
 (NP-OBJ (NP Al+muEoTaSimyna المعتصمين the seeking refuge)
 (PP min من from
 (NP hum مُم them))))))

كل المدعوين حضروا ما عدا المعتصمين منهم

kul~u + AlmadEuw~iyna + HadarwA + mA + EadA + AlmuEotaSimyna + min hum
 every + the invited + attended + what + excepts + the striking + from them
All the invited persons were present except the ones who are on strike

7.1.3 Fa

Arabic	POS Tag+Gloss	Syntactic /Semantic Context	TB
<p>فَ-</p> <p>فاء الجزاء fa-</p> <p>[The fa- of reward]</p>	<p>RC_PART</p> <p>So</p> <p>Then</p>	<p>It is used in conditional sentences introducing the apodosis sentence/main clause which can be imperative, interrogative, equational, or verbal</p> <p>fa- is a response conditional particle when it introduces a main clause that is the response of a condition stated in a subordinate clause introduced by <i*o, <i*A, or <in.</p>	<pre>(S (SBAR-ADV <in ان if/whether (S (S (VP (PRT lam لم did_not) yujomiEo يجمع agree (NP-SBJ Al+Earabu: العرب :the+Arabs))) wa و and (S (VP (PRT lam لم did_not) yuwaH~iduwA يوحدوا unite (NP-SBJ *) (NP-OBJ kalim+ata كلمت word (NP humهم their)))))) (PRT fa ف and/so) (NP-TPC-2 Al+qaDiy~apu القضية the+problem/issue Al+filasoTiyiniy~apu الفلسطينية the Palestinian) (PP-PRD la ل to/for (NP (NP hA ها it) (NP-2 *T*))) (NP-SBJ Eid~apu عدة several/numerous/many (NP >awojuh اوجه nogloss))) <in + lam + yujomiEo + Al+Earabu + wa+lam + yuwaH~iduwA + kalimatahum fa+AlqaDiy~apu + AlfilisTyniy~apu + la+hA + Eid~apu + >awojuhK If did not gather the Arabs and did not unite word their then the issue Palestinian for it several facets If Arabs do not agree and unite their voices then the Palestinian problem will have several facets.</pre>
<p>[الربط] فاء -</p> <p>fa-</p> <p>[The fa- of connection]</p>	<p>CONNEC_</p> <p>PART</p> <p>{ Discourse connective }</p>	<p>It connects two clauses. It is most commonly used to introduce a comment clause after a clause starting with >am~A.</p> <p>It also connects a fronted modifier to the rest of the clause it modifies.</p>	<ol style="list-style-type: none"> 1. A constituent introduced by the focus particle أَمَا (>am~A). Usually, the constituent will be a topicalized NP. <pre>(S (PRT >am~A أما as_for/concerning) (NP-TPC-1 Al+suloTapu السلطة the+power/authority)) (PRT fa ف and/so) (VP layos+at ليست not_be+it/they/she_[verb] (NP-SBJ-1 *T*) (ADJP-PRD msAlmp مسألة nogloss))) أما السلطة فليست مسألة >am~A + AlsulTapu + fa+layosat + musAlimapF. As for the authority the not be pacifist As for the government, then it was not pacifist.</pre> 2. A fronted constituent, whether a PP or an NP-ADV that functions as a modifier of the verb. <pre>(S (PP-PRP li ل for (NP *`lika ذلك that))) (PRT fa- ف and/so) (VP <in~a ان indeed (NP-SBJ (NP Hal~a حل solution (NP Al+qaDiy~api القضية the issue Al+falastiyniy~api الفلسطينية the Palestinian) (NP huwa هو it))))</pre>

			<p>(NP-PRD (NP Al+madoxalu المدخل the gateway (PP ل:li-::for/to (NP Hal~i حل solution (NP Al+qaDAyA القضايا the issues Al+>uxoraY الأخرى the other))))))) لذلك فإن حل القضية الفلسطينية هو المدخل لحل القضايا الأخرى li+*alika + fa+<in~a + Hal~a + AlqaDiy~api + AlfilisTiyniy~api + huwa + Almadxalu li+Hal~i + AlqaDAyA + Al>uxraY <i>For that then solving the Palestenian problem is the introduction for solving other problems.</i> That's why solving the Palestinian issue is the beginning to solving other issues.</p>
<p>- فـ [فاء العطف] fa- [The fa- of coordination]</p>	<p>CONJ And then</p>	<p>It marks a temporal sequence between two events</p>	<p>1.Starting a sentence where fa- is in an initial position just like wa- (S fa then (NP-SBJ tiloka تلك that/those) (NP-PRD ZAhirapN ظاهرة phenomenon+[fem.sg.] TabiyEiy~apN طبيعية natural/normal) .) فتلك ظاهرة طبيعية Fa tilka + ZAhirapN + TabyEiy~apN Then that phenomenon natural Then that is a natural phenomenon. 2.Coordinating 2 constituents with the same syntactic function and category (S (S (VP kAn+a كان be/was/were+he/it (NP-1 Al+nAxibuwna الناخبون the+voter/elector) (VP ya+noZur+uwna ينظرون look they (NP-SBJ-1 *T*) (PP-CLR <ilaY إلى to/towards (NP Suwari صور pictures (NP Al+mura\$a~aHiyn المرشحين the candidates)))))) fa-فـ and/so (S (VP yufaD~iluwna يفضلون they_[people]+prefer (NP-SBJ *) (NP-OBJ Al+muEam~ama المعمم the+wearing_a_turban) (PP EalaY على on/above (NP Al+>afanodiy الافندي the official)))))) كان الناخبون ينظرون إلى صور المرشحين فيفضلون المعمم على الافندي kAna + AlnAxibuwna + yanZuruwna + <ilaY + Suwari + Almura\$a~aHiyna + fa+yufaD~iluwna + AlmuEam~ama + EalY + Al>afandiy Was the electors look they to pictures the candidates then prefer they the wearing a turban on the gentleman The voters used to look at the pictures of the candidates and they preferred the man wearing a turban to the gentleman. 3. fa- can also coordinate two constituents that are of two different syntactic categories but having the same function under a UCP top node. (S (NP-TMP fawora فَوْز after (NP(NP AlsamAHi السماح the allowing) (PP la ل to (NP hum هُم them)) (PP bi ب (NP Eubuwwri عبور crossing (NP AlHudwdi الحدود the borders))))) (VP bada>a بدأ starts it (NP-SBJ-1 All~Aji}wna اللاجئين the refugees) (S(VP yaEoburuwna يغزؤون cross they (NP-SBJ-1 *)</p>

			<p>(NP-OBJ AlDif~apa الضفة the bank) (UCP-MNR (NP Hu\$wdAF حشوداً crowds kabiyrapF كبيرة big) fa ف so (PP bi ب with (NP >aEodAdK أعداد numbers >aqal~a أقل less))))))) فور السماح لهم بعبور بدأ الاجنوت يعبرون الضفة حشودا كبيرة فباعداد أقل Fawora + AlsamAhi + la+hum + bi+Eubuwri + AlHudwdi + bada>a + All~Aji}wna + yaEoburwuna + AlDif~apa + Hu\$wdAF + kabiyrapF + fa + bi+>aEodAdK >aqal~a <i>Immediately after they were allowed to pass the borders, the refugees started crossing the bank (in) big crowds then in smaller numbers.</i></p>
<p>فـ [فاء السببية] fa- [The fa- of causality]</p>	<p>SUB_CONJ J So that</p>	<p>It expresses the result or the cause of the main clause which must be imperative, interrogative, wish, or negation</p> <p>1. The verb of the clause introduced by ‘fa’ is in the subjunctive. 2. This ‘fa’ could be replaced by likayo, kayo, or li>an~a.</p>	<p>(S (VP baHav+nA بحثنا discuss/search_[for]+we_[verb] (NP-SBJ *) (PP-CLR fiy في in (NP <ixtiSArK اختصار brief (NP (NP Almammar~Ati الممرات the passages Aljawi~apa الجوية the air) (NP-LOC بين bayona::between/among (NP (NP bulodAni بلدان countries (NPnA نا our)) (ADJP AlvalAvapi الثلاثة the three)))))) (PP-PRP ل::li-::for (SBAR 0 (S (VP ta\$omala تشمل includes (NP-SBJ *) (NP-OBJ (NP Alxaliyja الخليج the gulf AlEarabi~a العربي the Arab) (SBAR-PRP fa ف and/so (S (VP takuwna تكون was (NP-SBJ >jawA&u أجواء atmospheres (NP nA نا our)) (ADJP-PRD jA*ibapF جاذبة attracting (PP ل li to (NP \$arikAti شركات companies (NP AlTayarAni الطيران The airline)))))))))) بحثنا في اختصار الممرات الجوية بين بلداننا الثلاثة لئشمل الخليج العربي فتكون أجوائنا جاذبة لمختلف شركات الطيران baHavna + fiy + <ixtiSArK + Almammar~Ati + Aljawi~apa + bayona + buldAni+nA + AlvalAvap + li+ta\$mala + Alxaliyja + AlEarabi~a fa+takuwna + >jawA&unA + jA*ibapF + li + \$arikAti +AlTayarAni. <i>We studied in brief the air passages between our three countries for them to comprise the Arabic Gulf and so our location will attract airline companies.</i></p>

7.1.4 mA(ما): WHNP, NP, SUB_CONJ, PRT, Verb, etc.

7.1.4.1 mA WHNP

7.1.4.1.1 Relative Pronoun mA (with trace)

mA can be a relative pronoun with the part-of-speech tag REL_PRON. It usually occurs as the head of a free relative clause (SBAR-NOM) which occupies nominal positions such as subject or direct object of the verb, complement of a noun, and complement of a preposition. This mA needs to be traced to its proper position in the lower clause, as subject, object, predicate, or with a resumptive pronoun, etc. The relative pronoun mA can be substituted by ‘Al~a*iy.’

```
(S (NP-SBJ هذا -h`*A this)
  (SBAR-NOM-PRD (WHNP-3 ما :mA: :what)
    (S (VP كان kAna was
      (NP-SBJ-3 *T*)
      (ADJP-PRD حصل HASil+F happening))))))
```

هَذَا مَا كَانَ حَاصِلًا

NB: Need transliteration here, and in the tree

This what was it happening

This is what was happening

NOTE: mA as a relative pronoun can head **adverbial** SBARs, in which case the SBAR takes only an ADV dashtag without the –NOM tag. This mA operates in the same sense as mim~A, as shown in the following example:

```
(S ثمّ vum~a then
  (VP بدأ bada>a started
    (NP-SBJ الجمهور Aljumohuwru the+supporters)
    (PP-CLR بي bi with
      (NP (NP الهتاف AlhutAfi booing)
        و wa and
        (NP (NP إلقاء <ilqA'i throwing
          (NP القارورات AlqAruwrAti bottles
            الفارغة AlfArigapi empty ) )
          (NP-DIR وسط wasaTa middle
            (NP الملعب AlmaloEabi the+stadium) ) ) ) )
      (SBAR-ADV (WHNP-1 ما mA what)
        (S (VP اضطرّ ADTar~a obliged
          (NP-SBJ-1 *T*)
          (NP-OBJ الحكم AlHakama the+referee)
          (PP-CLR إلى <ilaY to
            (NP إيقاف <iyqAfi stopping
              (NP المباراة AlmubarApi the game))))))))
      ثم بدأ الجمهور بالهتاف وإلقاء القارورات الفارغة وسط الملعب ما اضطر الحكم إلى إيقاف المباراة
```

NB: Need transliteration here, and in the tree

Then started the supporters with the booing and throwing the bottles the emptymiddle the stadium
what obliged the referee to stopping the game

*Then the supporters started booing and throwing empty bottles in the middle of the field, which
forced the referee to stop the match*

7.1.4.1.2 mA interrogative pronoun

mA can be an interrogative pronoun with the part-of-speech tag INTERROG_PRON when it heads an question (SBARQ). This mA needs to be traced in the clause it heads depending on the syntactic category and function of the constituent it substitutes for.

```
(SBARQ (WHNP-1 ما::mA::what)
  (S (VP دفع dafaEa push
    (NP-SBJ-1 *T*)
    (NP-OBJ ك ka you)
    (PP-CLR الى::<ilaY::to/towards
      (NP (NP المشاركة Almu$Arakapi participating)
        (PP في::fiy::in
          (NP (NP هذه::h`*ihi::this)
            (NP المغامرة AlmugAmrapi the adventure))))
      ?)
```

ما دفعك إلى المشاركة في هذه المغامرة؟

mA + dafaEa+ka + <ilaY + Almu\$Arakapi + fiy + ha*ihi + AlmugAmarapi

What pushed it you to the participation in this the adventure

What made you take part in this adventure?

7.1.4.1.3 mA exclamative pronoun

mA is an exclamative pronoun and is annotated as EXCLAM_PRON when it is used in exclamative propositions involving what is called in Arabic traditional grammar verbs of exclamation (أفعل/أفعال التعجب >afoEAl AltaEaj~ub) having the morphological pattern >afoEala (أفعل).

The exclamative mA (mA AltaEajubiy~ap) has a nominal value. It heads an SBAR and is traced in the lower clause as the subject of a verb of exclamation. It is followed by an NP-OBJ in the accusative.

(See section 4.1.3 of the syntactic annotation guidelines for “How beautiful...mA >ajmala...”)

```
(SBAR و::wa-::and
  (WHNP-1 ما::mA::what)
  (S (VP اخطر::>axoTar+a: make dangerous
    (NP-SBJ-1 *T*)
    (NP-OBJ التطرف AltaTar~ufa the extremism
      اليمنى Alyaminiy~a the right))))
```

ما اخطر التطرف اليمني

mA + <axoTara + AltaTar~ufa + Alyaniyny~a

how dangerous the extremism the right

How dangerous is right extremism!

7.1.4.2 Quantifier/Indefinite mA "some" = NP

In this case, mA is a noun (ما المُبَهَمَة/ mA Almubohamap ‘mA of indefiniteness’) and has the part-of-speech tag NOUN when it modifies an indefinite noun by adding (more) indefiniteness to

it. mA here conveys the meaning of “some.” At the Treebank level, this mA is annotated as being in apposition with the noun it modifies.

```
(S (VP كان::kAn+a::be/was/were+he/it_[verb]
  (VP ينتظر::ya+notaZir+u::he/it+expect/wait_for+[ind.]
    (NP-SBJ-3 *T*)
    (NP-OBJ (NP شيئاً::$ayo}AF::something/thing )
      (NP ما::mA::some))))))
```

كان ينتظر شيئاً ما

kAna + yantaZiru + \$ayo+ {AF mA
was he waits thing some
He was waiting something

More examples: yawmin mA "some day," >amrN mA "some question," fiy HAlatK mA "in any state."

7.1.4.3 mA SUB_CONJ

7.1.4.3.1 Subordinating Complementizer mA (ما المصدرية mA >al-maSdariy~ah)

When mA introduces a clause that, along with the mA, can be replaced by a masdar/gerund derived from the main verb of the clause. This mA is called the gerundive mA (ما المَصْدَرِيَّة) and usually follows words like Einda, baEoda, mivola, etc.

(S (NP-TPC-1 (NP كثر::kaviyrN::many/much/numerous)
 (PP من::min::from
 (NP علماء::EulamA'i::scholars/scientists
 (NP الدين::Al+diyni::the+religion))))
 (VP اعترفوا::iEotaraf+uwA::acknowledge/recognize[verb]
 (NP-SBJ-1 *T*)
 (SBAR أن::->an~a::that
 (S (NP-TPC-2 هم::hum::they_[masc.pl.]
 (VP <axoTa&uw أخطأوا got wrong
 (NP-SBJ-2 *T*)
 (NP-TMP عند::Einoda-::when)
 (SBAR ما::mA::that
 (S (VP تولوا::tawal~+awoA::
 take_charge_of
 (NP-SBJ *)
 (NP-OBJ المناصب AlmanASiba الرسمية))))))
 وكثير من علماء الدين اعترفوا أنهم أخطأوا عندما تولوا المناصب الرسمية
 wa+kaviyrN + min + EulamA'i + Aldiyni + <iEotarafuwA + >an~ahum + <axTa<wA
 Einda+mA + tawal~awA + AlmanAsiba + Alrasmiy~apa.
 And many from scholars the religion confessed they that were wrong they when took the
 positions the official
 And many religious scholars confessed that they were wrong when they took on official
 (government) positions

7.1.4.3.2 the conditional mA (ما الشرطية)

- When mA introduces a subordinate clause with a conditional meaning. This is called the conditional mA (ما الشرطية)

(S (NP-TPC-1 حزب::Hizob+a::party/band+[def.acc.]
 الله::All~h::Allah/God)
 (VP (PRT لن::lan::[will]_not/never)
 يقبل::ya+qobal+a::he/it+accept/receive/approve+[sub.]
 (NP-SBJ-1 *T*)
 (NP-OBJ عدم::Eadama::absence_of/lack_of
 (NP توجيه tawojiyhi directing
 (NP المضادات AlmuDAdAti antimissiles)
 (SBAR-ADV ما::mA::as long as
 (S (VP دام dAma last
 (NP-SBJ الانتهاك Al<inotihAki transgression
 الاسرائيلي Al<isorA{iyliyi Israeli)
 (ADJP-PRD قائما:qA{imF: standing))))))
 حزب الله لن يقبل عدم توجيه المضادات ما دام الانتهاك الاسرائيلي قائما
 Hizbu + Al~ahi + lan + yaqobala + Eadama + tawojiyhi + AlmuD~AdAti + mA + dAma
 Al<intihAku + Al<isrA{iliy~u + qA{imAF.
 Hizbullah not accepts not aiming anti missiles as long as the violation the Israeli existing
 Hizbullah will not accept not aiming anti-missile missiles as long as the Israeli violations are
 ongoing

7.1.4.4 Particle mA (PRT)

7.1.4.4.1 mA negative particle

It negates the meaning of the constituent it precedes. In this context, the mA can be replaced by another negative particle with no change in meaning.

mA can negate a verb or a noun or any constituent.

```
(S (PP-PRD في::fiy::in
  (NP (NP معنى maEnAY meaning
    (NP (NP هذا::h`*A
      (NP الاسم Al<isomi name
        السامي AlsAmiy superior))))
  (NP-SBJ بداية bidAyapu beginning
    (NP (NP قداسة qadAsapK : holiness)
      (SBAR (WHNP-1 *0*)
        (S (VP (PRT ما::mA::not)
          عرفت Earifat knew
          (NP-SBJ-1 *T*)
          (NP-OBJ ظلالاً DilAlF shadows)))))))
```

في معنى هذا الاسم السامي بداية قداسة ما عرفت ظلالاً

Fiy + maEonaY + ha*A + Al<ismi + bidAyapu + qadAsapK + mA + Earifat + ZilalAF

In meaning this the name the holy beginning holiness not know it shadows

In the meaning of this sacred name there is the beginning of a holiness that never knew any flaws

```
(S (NP-TPC-1 (PRT ما mA not )
  (PRT من min of )
  >aHad+K no one)
  (VP يقدر yuqad~iru respect+he
    (NP-SBJ-1 *T*)
    (NP-OBJ عمل Eamal+a work
      (NP كم kum you ))
    (NP-ADV مثل mivol+a like
      (SBAR ما mA that
        (S (VP أقدر >uqad~iru respect+ I
          (NP-SBJ *)
          (NP-OBJ ه hu him))))))
```

ما من أحد يقدر عملكم مثل ما أقدره

NB: Need transliteration here

Not of one values he work your as that value I it

No one values your work the way I do

```

(S (NP-TMP (PRT ما mA not)
        (PRT من min of)
        يوم yawom+K day )
  (VP (PRT إلا il~A but)
      (PRT و wa and)
      تذكرت ta+*ak~ar+tu remebered
      (NP-SBJ *)
      (NP-OBJ ه hu him)))

```

ما من يوم إلا وتذكرته

NB: Need transliteration here

Not of day but and remembered I him

I remembered him every single day

7.1.4.4.2 mA AlzAidap (ما الزائدة)

mA is a particle and has the part-of-speech tag PART when it follows an adjective that has an adverbial function in the sentence. Such adjectives in the indefinite and in the accusative case like غالباً (gAlibAF/ often), كثيراً (kaviyrAF/much) are annotated in the tree as NP-ADV. What follows this particle mA must be a VP, and the mA particle links the NP-ADV to the VP.

(S (NP-TPC-2 (NP السلطات::Al+suluTAtu::the+authorities
النقدية::Al+naqodiy~apu::the+monetary/cash)
(PP في::fiy::in
(NP (NP عدد::EadadK::number/quantity/issue)
(PP من::min::from
(NP الدول Alduwali the countries))))))
(NP-ADV غالبا::gAlibAF::generally/commonly/typically
(PRT ما::mA)
(VP تعتمد taEotamidu rely+she
(NP-SBJ-2 *T*)
(NP-OBJ (NP اسعار::>asoEAra::prices/rates
(NP فائدة::fA{id+ap+K::benefit/use))
(ADJP مرتفعة::murotafiEapK::elevated/rising))
(PP-PRP ل::li-::to/for
(NP (NP لدفاع:AldifAEi :the defense)
(PP عن::Ean::from/about/of
(NP اسعار::>asoEAri::prices/rates
(NP الصرف AlSarofi the change)))))))))
السلطات النقدية في عدد من الدول غالبا ما تعتمد اسعار فائدة مرتفعة للدفاع عن اسعار الصرف
AlsuluTAtu + Alnaqdy~apu + fiy + EadadK + min + Alduali + gAlibAF + mA + taEotamidu
+ AsoEAra + fA{idapK + murotafiEapF + li+AldifAEi + Ean + AsoEAri + AlSarfi
The authorities the monetary in number from the countries often relies it prices interest high
to protecting on prices exchange
*It is often the case that monetary institutions in a number of countries employ high interest
rates to protect stock market prices*

7.1.4.5 mA verb ما المشبهة بليس

mA is a non-inflectional verb annotated as VERB when it plays the same syntactic role as *layosa*; that is to say, when it assigns the nominative to its subject and the accusative to its predicate. The difference between the non-inflectional verb mA and the negative particle mA is that the former negates the whole predication whereas the latter only negates verbs. mA as a non-inflectional verb is always followed by a noun in the nominative. It never occurs before a verb.

(S (VP ما:::-mA:::not
 (NP-SBJ تلك:::tiloka:::that/those)
 (NP-PRD سوى:::siwaY:::other_than/except_for)
 (NP بذور:::bu*uwru+u:::seeds+[def.nom.]
 (NP (NP اعتاق <inoEitAqi liberation
 (NP الفرد Alfarodi the individual))
 (PP من:::min:::from
 (NP قبيلته:::qabiyl+at+i-:::tribe
 (NP ه:::-hi:::his/its)))))))))
 ما تلك سوى بذور اعتاق الفرد من قبيلته
 mA + tiloka + siwaY + bu*uwru + <inoEitAqi + Alfarodi + min + qabiylati+hi
 not these but seeds liberation the individual from tribe his
Those are nothing but signs of the liberation of the individual from his tribe

(S (VP ما:::-mA:::not)
 (NP-SBJ نيل:::nayolu:::attainment/achievement/acquiring
 (NP المطالب:::Al+maTAlibi:::the+demands))
 (PP-MNR- PRD ب:::bi-:::with/by
 (NP التمني Altaman~iy the wishing))))
 ما نيل المطالب بالتمني
 mA + nayolu + AlmaTAlibi + bi+Altaman~iy
 not reaching the demands with the wishing
Reaching goals is not (just) by wishing

(S (VP ما mA not
 (NP-SBJ (NP هو huw~a he)
 (PP ل li to
 (NP ك ka you)))
 (PP-PRD ب bi with
 (NP جار jAr+K neighbor))))
 ما هو لك بجار

NB: Need transliteration here

Not he to you with neighbor
He is not a good neighbor

7.1.4.6 When to split “mA” and when not to split it

7.1.4.6.1 not split : mA is a part of SUB_CONJ, CONJ or PART

The two parts are always written together so we shouldn't split mA here.

Part of a SUB_CONJ: lam~A "when", lawomA "if", HiynamA "when", HAlamA "when", kayomA "in order to", rayvamA "until."

```
(S (SBAR-ADV لما lam~A when
    (S (VP أنهى >anohAY finished+he
        (NP-SBJ * )
        (NP-OBJ عمل Eamal+a work
            (NP ه hu his))))))
(VP غادر gAdar+a left+he
    (NP-SBJ *)))
```

لما أنهى عمله غادر

NB: Need transliteration here

When finshed he work his left he

When he finished his work, he left

```
(S (SBAR-ADV حالا HalamA as soon as
    (S (VP نحصل naHoSu+u get+we
        (NP-SBJ * )
        (PP-CLR على Ealay on
            (NP النتائج AlnatA{ij+i the results))))))
(VP (PRT س sa will)
    ناخذ na+t~axi*+u take
    (NP-SBJ * )
    (NP-OBJ الإجراءات Al+<ijorA'At+a the measures
        المناسبة Al+munAsibap+a the appropriate )))
```

حالما نحصل على النتائج سنأخذ الإجراءات المناسبة

NB: Need transliteration here

As soon as get we on the results will take we the measures the appropriate

As soon as we get the results, we'll take the appropriate measures

Part of a relative pronoun (part-of-speech tag REL_PRON): HayovumA “wherever”,
kayofamA “however”, mahomA “whatever”

```
(S (VP (PRT لن lano not)
    نستسلم na+sotasolim+a give up
    (NP-SBJ * )
    (SBAR-ADV (WHNP -1 مهما mahomA whatever )
        (S (VP كانت kAn+at is
            (NP-SBJ التكاليف Al+takAliyf+u the prices)
            (NP-PRD -1 *T* ))))))))
```

لن نستسلم مهما كانت التكاليف

NB: Need transliteration here

Not give up we whatever were the costs

We won't give up whatever are the results.

Part of a conjunction (part-of-speech tag CONJ): >im~A "either, or"

```

(S (PP-PRD علي EalayO on
  (NP ه hi him))
  (SBAR-SBJ أن >ano that
    (S (VP يختار ya+xtAr+a choose+he
      (NP-SBJ *)
      (NP-OBJ أحد >aHad+a one
        (NP أمرين >amor+ayoni thing+two
          (SBAR إما <im~A either
            (SBAR أن >ano that
              (S (VP يراهن yu+rAhin+a
                make a bet+he
                (NP-SBJ *)
                (PP-CLR علي Ealay on
                  (NP الفوز Al+fawoz+i
                    winning))))))
              (CONJP و wa and
                إما <im~A or)
              (SBAR أن >ano that
                (S (VP ينسحب ya+nosaHib+a
                  withdraw+ he
                  (NP-SBJ *)
                  (PP-TMP من min from
                    (NP الآن Al+!An+a
                      now))))))))))
            عليه أن يختار أحد أمرين إما أن يراهن على الفوز وإما أن ينسحب من الآن

```

NB: Need transliteration here

On him that chooses he one of two things either that makes a bet on the winning or that withdraws he from the now

He has either to make a bet on winning or to withdraw from now.

Part of a particle: >ay~umA “so much”, >am~A “as for”, <in~amA ”but indeed”, ka>an~amA “as if”

```

(S (VP كنت kunot+u was+I
  ( VP أقدر <uqad~ir+ u respect+I
    (NP-SBJ-1 *)
    (NP-OBJ ه hu him)
    (NP-ADV (PRT >ay~amA أيما very much)
      Taqodiyi+K تقدير respect ))))

```

كنت أقدره أيما تقدير

NB: Need transliteration here

Was i respect i very much respect

I respected him very much

7.1.4.6.2 Split it: mA is a SUB_CONJ, REL_PRON, INTEROG_PRON

Most of the time the two parts are written together and we should split them, and mA will have the POS tag **SUB_CONJ**.

EinodamA:

```
(S (VP (PRT لم lam not )
      تلفت ta+lotafit+o turn+she
      (NP-SBJ الفتاة Al+fatAap+u the young woman )
      (PP-CLR إلي <ilayo to
        (NP ه hi him)
      (NP-TMP (PRT إلا <il~A only)
        عند Einoda when
        (SBAR ما mA that
          (S (VP ذكر *akar+a mentioned+he
            (NP-SBJ *)
            (NP-OBJ اسم <isom+a name
              (NP (NP أبي >abiy father
                (NP ها hA her))
              (ADJP الراحل Al +rAHil+i
                late))))))))))
      لم تلفت الفتاة إليه إلا عندما ذكر اسم أبيها الراحل
```

NB: Need transliteration here

Not turned she the girl to him only when mentioned he name father her the late

The young woman turned only when he said the name of her late father.

```
(S (VP فقدنا faqado+nA lost+we
      (NP-SBJ *)
      (NP-OBJ (NP الأمل Al+>amal+a the hope)
        (PP في fiy in
          (NP عودت Eawodap+i going back
            (NP ه hi his))))
      (NP-TMP بعد baEod+a after
        (SBAR ما mA that
          (S (VP انتظرنا <inotaZaro+nA wait+we
            (NP-SBJ *)
            (NP-OBJ ه hu him)
            (NP-TMP طويلا TawiyL+F long)
            (NP-ADV دون duwn+a without
              (NP جدوى jadowAY feasibility ))))))))
      فقدنا الأمل في عودته بعدما انتظرناه طويلا دون جدوى
```

NB: Need transliteration here

Lost we the hope in coming back his after that waited we him long without meaning

*We lost hope in his coming back after we waited for him for so long with no avail***Relative pronoun (part-of-speech tag REL_PRON)**

baEodamA:

```

(S (NP-TMP بعد baEoda after
  (SBAR-NOM (WHNP-1 ما mA that)
    (S (VP حصل HaSal+a happen+it
      (NP-SBJ-1 * T*))))))
(VP (PRT لم lamo not)
  يعد yaEud+o return+it
  (NP-SBJ-2 *)
  (S (VP ينتاب Ya+notAb+u feel+It
    (NP-OBJ ني niy me)
    (NP-SBJ-2 (NP أي >ay~+u no
      (NP شك $ak~+K doubt))
      (PP في fiy in
        (NP نزاهة nazAhap+i honesty
          (NP ه hi his)))))))
  بعدما حصل لم يعد ينتابني أي شك في نزاهته

```

NB: Need transliteration here

After what happened it not come feel me any doubt in integrity his

After what happened, I haven't had any doubts about his honesty

```

(S (VP سقط saqaT+a fell+he
  (NP-SBJ *)
  (PP على EalAY on
    (NP وجه wajoh+i face
      (NP ه hi his))
  (PP مم mim from
    (SBAR-NOM (WHNP-1 ما mA what)
      (S (VP أثار >avAr+a excite+it
        (NP-SBJ-1 *T*)
        (NP-OBJ سخرية suxoriyap+a irony
          (NP المارة Al+mAr~ap+i passersby ))))))))
  سقط على وجهه مما أثار سخرية المارة

```

NB: Need transliteration here

Fell he on face his from what excite irony the passers by

He fell on his face, which made passers by laugh at him

Interrogative pronoun (with the part-of-speech tag INTERROG_PRON)

Eam~A "about that which":

```

(SBARQ (WHPP-1 ع Ea on
  (WHNP ما mA What))
  (S (VP تستفسرون ta+sotafosir+uwna ask+you
    (NP-SBJ *)
    (PP-CLR-1 *T*)
    (PP ب bi with
      (NP كل kul~+I all
        (NP (NP هذا ha*A this)
          (NP الإلحاح Al+<iloHAH+I insistence))))))
  عما تستفسرون بكلّ هذا الإلحاح

```

NB: Need transliteration here

On what ask you with all this insistence

What are you asking about with all this instance?

mim~A:

```
(SBARQ (WHPP-1 م mi from
      (WHNP مَا mA what))
  (S (VP تخافون ta+xAf+uwna be afraid of+you
      (NP-SBJ *)
      (PP-1 *T*)
      (SBAR-ADV و wa and
        (S (S (NP-SBJ أنتم >antum you)
              (PP-PRD على EalAY on
                (NP حق Haq~+K right))))
          و wa and
          (S (NP-SBJ هم hum they)
              (PP-PRD على EalAY on
                (NP باطل bATi+K wrong))))))))))
مما تخافون وأنتم على حق وهم على باطل
```

NB: Need transliteration here

From what are afraid you while you on right and they on wrong

What are you afraid of when you are in the right and they are in the wrong?

Note: mataY mA waqta mA are written separately and should be split also

7.1.4.6.3 Split or not split both are possible with different meaning

7.1.4.6.3.1 The case of fiymA

mA is a part of a SUB_CONJ, PART or CONJ: written together: NOT SPLIT	mA is a REL_PRON or SUB_CONJ: SPLIT
<p>FiymA : SUB_CONJ (while) (S(VP >aEolan+a: أعلن:announce+he (NP-SBJ >aEoDA'+u:: أعضاء members (NPAl+tanoZiym+i: التنظيم:group)) (S-NOM-OBJ (VP qubuw1+a: قبول:acceptance (NP-SBJ هم hum their) (NP-OBJ Al+mufAwaD+at+I المفاوضات negotiations))) (SBAR-ADV fiymA: فيما:while (S(VP نفت nafa+at denied+she (NP-SBJ Al+Hukuwmap+u الحكومة the government) (NP-OBJ >ay~ap+a أية any (NP(NP <imokAniy~ap+i إمكانية possibility (PP li ل to (NP(NP Al+tafAwuD+i التفاوض negotiating) (NP-ADV maEa مع with (NP Al+<irohAbiy~+iyna الإرهابيين The terrorists))))))))))</p> <p>أعلن أعضاء التنظيم قبولهم المفاوضات فيما نفت الحكومة أية إمكانية للتفاوض مع الإرهابيين Announced members the organization acceptance their the negotiations while denied =she the government any possibility of negotiating zith the terrorists the members of Al-Qaida announced their acceptance of the negotiations while the government denied such a thing.</p>	<p>In this case, mA and what comes before it are written separately (except for bimA and fiy+mA sometimes) Fiy + mA PREP + REL_PRON (in + what) (S(VP ta+nAqa\$onAlناقشنا discussed+we (NP-SBJ *) (PP-CLR fiy في in (SBAR-NOM(WHNP-1 mA ما what) (S(VP samiEo+nA سمعنا heard+we (NP-SBJ *) (NP-OBJ (NP hu ه it) (NP-1 *T*)) (NP-TMP Al+yawom+a اليوم Today))))))</p> <p>تناقشنا في ما سمعناه اليوم Discussed we in what heard we it the today. We discussed what we heard today.</p>

7.1.4.6.3.1.1 The case of TAlamA

mA is a part of a SUB_CONJ, PART or CONJ: written together: NOT SPLIT	mA is a REL_PRON or SUB_CONJ: SPLIT
<p style="text-align: center;">TAlamA SUB_CONJ (as long as)</p> <p>(S(VP(PRT لن lano will) yu+foliH+uw يفلحوا succeed+they (NP-SBJ *) (SBAR-ADV TalamA طالما as long as (S (NP-SBJ hum هم they) (PP-PRD fiy في in (NP tanAHur+K تناحر mutual slaughtering dA}im+K دائم permanent))))))</p> <p>لن يفلحوا طالما هم في تناحر دائم Not succeed they as long as they in mutual slaughtering permanent They won't succeed as long as they are permanently slaughtering each other.</p>	<p style="text-align: center;">Tala + mA VERB + REL_PRON</p> <p>(S(VP Tal+a طال take long (NP-SBJ(SBAR-NOM(WHNP-1 mA ما what) (S(VP na+notaZir+u ننتظر wait+we (NP-SBJ *) (NP-OBJ(NP hu ه it) (NP-1 *T*)) (PP من min of (NP(NP تحقيق تحقيق taHoqiyq+K realizing) (PP ل li of (NP لوعود Al+wuEuwd+ i the promises))))))))))</p> <p>طال ما ننتظره من تحقيق للوعود What we are waiting for with respect to keeping promises and making them true is taking so long.</p>

7.1.4.6.3.2 The case of bayonamA

mA is a part of a SUB_CONJ, PART or CONJ: written together: NOT SPLIT	mA is a REL_PRON or SUB_CONJ: SPLIT
<p>bayonamA SUB_CONJ (while) (S(SBAR-TMP bayonamA بينما while (S(VP kAnA كان be+he (NP-1 Al+Earab+u العرب the Arabs) (VP yu+tarojim+uwna يترجمون translate+they (NP-SBJ-1 *T*) (PP-MNR bi ب with (NP \$agaf+K شغف passion)) (NP-OBJ kutub+a كتب books (NP Al+yunAn+i اليونان the Greece)))))) (VP kAnA كان be+he (NP-1 الغرب Al+garob+u the West) (VP yu+Eaniy يعاني suffer+he (NP-SBJ-1 *T*) (PP-CLR min من from (NP sayoTarap+i سيطرة control (NP rijAl+i رجال (NP Al+kaniysap+i الكنيسة the church)))))))) بينما كان العرب يترجمون بشغف كتب اليونان كان الغرب يعاني من سيطرة رجال الكنيسة While was the arabs ttanslate with passion books the Greece was the West suffers from control men the church While Arabs were fervently translating Greeb books, the West were suffering from the tyranny of the church.</p>	<p>Bayna + mA NOUN + REL_PRON (between what) (S(VP شتآن \$at~Ana how different (NP-SBJ bayona بين between (SBAR-NOM(SBAR-NOM(WHNP-1 mA ما what) (S(VP yu+qAl+u يقال is said+it (NP-SBJ-1 *T*) (NP-OBJ-1 *) Wa و (SBAR-NOM(WHNP-2 mA ما what) (S(VP yu+nojaz+u ينجز is realized+it (NP-SBJ-2 *T*) (NP-OBJ -2 *) شتآن بين ما يقال وما ينجز How different between what is said it and is realized it There is a huge difference between what is said and what is achieved.</p>

7.1.4.6.3.3 The case of bimA

mA is a part of a SUB_CONJ, PART or CONJ: written together: NOT SPLIT	mA is a REL_PRON or SUB_CONJ: SPLIT
<p>bimA anna SUB_CONJ (since) (S(SBAR-ADV bimA بما since >an~a أَنْ that (S(NP-TPC-1 ك ka you) (VP qabilo+ta قبلت accepted+you (NP-SBJ-1 *T*) (NP-OBJ Al+taHad~iy التحدي the+challenge)))) (PRT fa ف so) (PP-PRD Ealayo علي on (NP ka ك you)) (NP-SBJ(NP Al+Sumuwd+u الصمود the+resistance) (PP-TMP <ilAY إلى to (NP Al+nihAyap النهاية the+end)))) بما أنك قبلت التحدي فعليك الصمود إلى النهاية Since that you accepted the challenge so on you the resistance to the end Since you accepted the challenge, you have to resist to the end.</p>	<p>Bi +mA PREP + REL_PRON (with what) (S (VP jAda جاد give (NP-SBJ Al+faqiyr+u الفقير the+poor) (PP-CLR bi ب with (SBAR-NOM (WHNP-1 mA ما what) (S(NP-SBJ-1 *T*) (NP-PRD عند Einoda at (NP hu •him)))))) جاد الفقير بما عنده Gave he the the poor with what with him The poor man gave what he has.</p>

7.1.4.6.3.4 The case of kamA

mA is a part of a SUB_CONJ, PART or CONJ: written together: NOT SPLIT	mA is a REL_PRON or SUB_CONJ: SPLIT
<p>kamA CONJ (and) (S kamA كما and (VP >aEolan+at أعلنت announced+she (NP-SBJ maSAdir+u مصادر sources >uxorAY أخرى other) (SBAR >an~a أن that (S(NP-TPC-1 Al+<ijotimAE+a الاجتماع the meeting Al+wizaAriy~+a الوزاري the+ministerial) (VP(PRT lam لم not) yu+sofir+o يسفر result (NP-SBJ-1 *T*) (PP-CLR عن Ean on (NP >it~ifAq+K اتفاق agreement tAm~+K تام total))))))))) كما أعلنت مصادر أخرى أن الاجتماع الوزاري لم يسفر عن اتفاق تام And announced she sources other that the meeting the ministerial not result on agreement total Other sources announced that the ministerial meeting didn't come to a total agreement.</p>	<p>Ka+ mA SUB_CONJ (as +that) (S(VP EAmala عامل treated (NP-SBJ *) (NP-OBJ hu ه him) (PP-MNR ka ك as (SBAR mA ما that (S(VP yu+EAmil+u يعامل treats (NP-SBJ Al+>ax+u الأخ the+brother+Nom) (NP-OBJ >axA أخا brother+Acc (NP hu ه his))))))))) عامله كما يعامل الأخ أخاه Treated he him as treats the brother brother his He treated him the way brothers treat each other.</p> <p>PREP+ REL_PRON(as +what) (S(NP-SBJ ha*A هذا this) (PP ka ك as (SBAR-NOM(WHNP-1 mA ما what) (S(VP yu+qAl+u يقال was said (NP-SBJ-1 *T*) (NP-OBJ-1 *)))))) (NP-PRD(NP Haq~+N حق right) (SBAR (WHNP-2 0) (S(VP >uriyda أريد was wanted (PP-CLR bi ب with (NP (NP hu ه it) (NP-2 *T*))) (NP-SBJ-1 bATil+N باطل mischief) (NP-OBJ-1 *))))))))) هذا كما يقال حق أريد به باطل This as what was said he right was wanted it with it evil This is as it is said a right deed that was originally meant to be mischievous.</p>

7.1.4.6.3.5 The case of kul~amA

mA is a part of a SUB_CONJ, PART or CONJ: written together: NOT SPLIT	mA is a REL_PRON or SUB_CONJ: SPLIT
<p>kul~amA + PERFECT-VERB : everytime that, whenever, as often as</p> <p>(S(SBAR-ADV kul~amA كلما whenever (S(VP Hawala حاول tried (NP-SBJ Al+>ab+u الأب the+father) (NP-OBJ(NP Al+taqar~ub+a التقرب the+ getting closer) (PP min من from (NP <ibon+i ابن son (NP hiه his))))))</p> <p>(VP >aHas~+a أحس felt (NP-SBJ *) (PP-CLR ب bi with (NP Eumoq+i عمق depth (NP(NP Al+fajowap+i الفجوة the+gap) (SBAR(WHNP -1 Al~atiy الي which)</p> <p>(S(VP >aSbaH+at أصبحت became+she (NP-SBJ-1 *T*) (NP-PRD bayona بين between (NP humAهما them))))))</p> <p>كلما حاول الأب التقرب من ابنه أحسّ بعمق الفجوة التي أصبحت بينهما</p> <p>Whenever tried he the father the getting closer from son his felt he with depth the gap that became itbetween them Every time the father tried to get closer to his son, he felt the depth of the gap coming between them.</p>	<p>Kul~a +mA(REL_PRON)</p> <p>(S(VP jamaEa جمع gather (NP-SBJ *) (NP-OBJ kul~a كل every (SBAR-NOM(WHNP-2 man من who) (S(VP ya+qodir+u يقدر be_capable_of (NP-SBJ-2 *T*) (PP-CLR EalaY على on (NP(NP Al+Eamal+i العمل the+work (PP fiy في in (NP(NP h`*ihi هذه this) (NP Al+qaloEap+i القلعة the+stronghold))))))))))</p> <p>جمع كل من يقدر على العمل في هذه القلعة</p> <p>Gathered he all who be able to the work in this the stronghold He gathered anyone who can work in this stronghold.</p>

7.1.4.6.3.6 The case of <in~amA

mA is a part of a SUB_CONJ, PART or CONJ: written together: NOT SPLIT	mA is a REL_PRON or SUB_CONJ: SPLIT
<p><in~amA is a PRT</p> <p>(S (VP (PRT لم lam not) يكن ya+kun+o is (NP-SBJ قصد qaSod+u intention (NP ه hu his)) (S-NOM-PRD (S-NOM (VP <iEolAm+a إعلام informing (NP-SBJ *) (NP-OBJ ه hu his (PP-DTV ب bi with (NP Al+mu\$okilap+i المشكلة the+problem)))) و wa and (S-NOM (VP (PRT إنَّما only) توريط taworiyT+a involving (NP-SBJ *) (NP-OBJ hu ه his) (PP-CLR في في in (NP hAها it)))))) لم يكن قصده إعلامه بالمشكلة وإنَّما توريطه فيها Not was he intention his informing his with the problem and only involving his in it He didn't mean to inform him about the problem but to get him involved in it.</p>	<p><in~a+mA (REL_PRON)</p> <p>(S (VP <in~a إن indeed (SBAR-NOM-SBJ-1 (WHNP-2 mA ما what) (S (VP قام qAma did (NP-SBJ *) (PP-CLR ب bi with (NP (NP hi ه it) (NP-2 *T*)))))) (S (VP ya+t~asim+u يتسم is characterized (NP-SBJ-1 *) (PP-CLR ب bi with (NP Al+\$ajAEap+I الشجاعة the courage)))))) إن ما قام به يتسم بالشجاعة Indeed what did he with it characterized with the courage What he did is courageous.</p>

7.1.4.6.4 When words related to “mA” need a trace and when not

7.1.4.6.4.1 There should be a trace, when it is a free relative and mA is a REL_PRON

Test : it can be replaced by Al*iy:

(S (SBAR-NOM-SBJ (WHNP-1 ما / الذي mA/ Al~*iy what)
(S (VP ta+quluw+u say+you
(NP-SBJ *)
(NP-OBJ (NP hu it)
(NP-1 *T*)))))
(NP-PRD gayor+u
(NP SaHiyH+K)))

ما تقوله غير صحيح

الذي تقوله غير صحيح Vs

NB: Need transliteration here, and Arabic script in the tree

What said you it not correct

What you said is not correct

7.1.4.6.4.2 There should not be a trace

7.1.4.6.4.2.1 When it is a SUB_CONJ

Test: It can be replaced by >an or by a gerund:

```
(S (NP-TMP بعد baEoda after
  (SBAR ما/ أن mA/ >an that
    (S (VP يأتي ya+>otiy come+it
      (NP-SBJ الصبح Al+SabAH+u the morning))))))
(VP نذهب na+*ohab+u go+we
  (NP-SBJ *)
  (PP-DIR إلى <ilAY to
    (NP السوق Al+suwq+I the market))))
بعد أن يأتي الصبح، نذهب إلى السوق / بعد إتيان الصبح، نذهب إلى السوق / نذهب إلى السوق
```

NB: Need transliteration here

After that the morning comes, we go to the market.

NB: Need translation here

7.1.4.6.4.2.2 when it is part of SUB_CONJ like kulamA, TalamA, etc.

NB: there is no test or example in this section

7.1.5 wa- (وَ)

Ara bic	POS Tag	Gloss	Syntacti c /Semanti c Context	TB
و [واو القسم]	PREP	By	It is used in oaths	(PP و::wa::by (NP الله::Al~ahi::God)) وَاللّٰهِ By God!
و [واو المعية]	PREP	Together with Along with	It expresse s the meaning of accompa niment	This wa expresses the meaning of togetherness. It assigns the accusative case to its complement. It can be replaced by مع (maEa/with). Most of the time, this wa is used with verbs with an embedded meaning of 'reciprocity'. (S (VP يتوافق yatawAfaqu comply (NP-SBJ *) (PP-CLR و wa-with (S-NOM (VP تصنيف tasniyfa classification (NP-SBJ *) (NP-OBJ جليل jubayol Jubail) (PP-CLR ك::ka- as (NP (NP موقع maowqiEK location) (PP ل: li for (NP لتراث: AlturAvi the heritage العالمي AlEAlamy~i the international))))))))) يتوافق وتصنيف جليل كموقع للتراث العالمي.

				<p>yatawAfaqu + wa+tasniyfa + jubayol + ka+maowqiEK +li+lturAvi + AIEAlamy~i</p> <p><i>This complies with the classification of Jubail as a site of world heritage.</i></p>
<p>و-</p> <p>[الربط]</p>	CONJ	And	It coordinates two constituents	<p>Both the wa of coordination (واو العطف / wAw AlEatf) and the wa of connection (واو الربط / wAw AlrabT) are annotated as conjunctions in the following contexts:</p> <p>- Coordination of two constituents of the same syntactic category and function.</p> <p>(NP (NP القرار Al+qarAri the+decision) (ADJP التاريخي Al+tAriyxiy~I the+historical :wa-::and الفريد Al+fariydi the+unique)) القرار التاريخي والفريد</p> <ul style="list-style-type: none"> Coordination of two constituents of different syntactic categories (NP-PP, ADJP-PP, ADJP-SBAR, S-SBAR etc.) but with the same syntactic function. In this case the higher node that ties the two constituents together should be UCP. <p>(S :wa-::and (VP قالوا qAl+uwA said+they_[verb] (NP-SBJ *) (S " (NP-SBJ هذا :h`*A::this_[masc.sg.] (NP-PRD (NP غير :gayor+a::not/other (NP مسموح masomuwhK permissible)) :wa-::and (NP غير :gayor+a::not/other (NP مقبول maqobuwLk welcome)) (PP-LOC في :fiy::in (NP (NP منطقة minoTaq+apK area/zone) (UCP (ADJP مكتظة munkta~apK populated (PP ب :bi-::with/by (NP السكان Al+suk~Ani the+residents و wa and المتاجر Al+matAjiri the+stores))) :wa-::and (SBAR (WHNP-1 *0*) (S (VP تستحق ta+sotaHiq~u it+deserve (NP-SBJ-1 *T*) (NP-OBJ الاهتمام</p>

				<p>Al+{ihotimAma</p> <p>the+interest</p> <p>))))))</p> <p>”</p> <p>)))</p> <p>وقالوا هذا غير مسموح وغير مقبول في منطقة مكتظة بالسكان والمتاجر وتستحق الإهتمام.</p> <p>wa + qAluwA + “ + ha*A + gayoru + masmuwHK + wa+gayoru + maqbuwlK + fiy + mintaqapK + mukotaZ~apK + bi+Alsuk~Ani + wa+AlmatAjjiri + wa+tastaHiq~u + Al<ihtimAma ”</p> <p><i>And they said “this in not allowable and not acceptable in a zone populated with residents and shops and (that) needs a lot of attention”</i></p>
				<p>- when it occurs at the beginning of the paragraph functioning like a discourse connector.</p> <p>(S (S (NP-TTL-TPC-1 (PRT حتى ::-Hat~aY::even) واشنطن ::wA\$inoTun::Washington بوست ::buwsot::Post nogloss) (VP (PRT لم ::lam::did_not) تتردد ::ta+tarad~ad+o::it/they/she +be_hesitant +[jus.] (NP-SBJ-1 *T*) (PP-CLR في ::fiy::in (NP (NP إثارة ::<ivAr+api::provocation (SBAR-NOM (WHNP-2 ما ::mA::what) (S (VP يشبه ::yu+\$obihi he/it+resemble (NP-SBJ-2 *T*) (NP-OBJ السخرية Al+suxoriy~apa the+sarcasm) (NP-TMP (NP يوم ::yawoma::day) (NP السبت ::Al+saboti::the+Saturday الماضي ::Al+mADiy::the+past))))))</p> <p>وَحَتَّى وَاشْنِطْنَ بُوسَتْ لَمْ تَتَرَدَّدْ فِي إِثَارَةِ مَا يُشْبِهُ السُّخْرِيَّةَ يَوْمَ السَّبْتِ الْمَاضِي</p> <p>wa + Hat~aY + wA\$inTun + buwsot + lam + tatarad~ad + fiy+<ivArapi + mA + yu\$bihi Alsuxriyapa + yawma + Alsabti + AlmADy</p>

				<p><i>And even Washington Post did not hesitate to provoke what looks like sarcasm last Staurday.</i></p> <p>- when wa, as a discourse connector, introduces a modifier which is almost always an SBAR or a PP. The wa along with the modifier will be included under an NAC node.</p> <pre>(S (NP-SBJ هذا :-h` *A::this_[masc.sg.]) (SBAR-NOM-PRD (WHNP-2 ما ::mA::what) (S (VP يعطي ::yu+EoTiy he/it+give (NP-SBJ-2 *T*) (NP-DTV الفلسطينيين Al+filasoTiyiny~iyna the+Palestinian) (NP-OBJ حقا ::Haq~AF truth/right اكثر >akovar+a more) (NAC و ::wa-::and (PP-TMP ل ::-li-::for/to (NP مدة mud~apK period_of_time أطول >aTowala longer/longest))))))) هذا ما يعطي الفلسطينيين حقا أكثر ولمدة أطول.</pre> <p>ha*A + mA + yuEoTiy + li+AlfalasTynyyna + Haq~AF + <akvara + wa+li + mud~apK + <ATowala <i>This is what gives the Palestinians more right and for a longer time</i></p>
- [واو الحال]	SUB_ CONJ	{Subordinating conjunction}	It introduce s a clause describin g the manner with which an event took place	<pre>(S (S wa and (VP لكنْ lakin~a but (NP-SBJ-1 نيْ niy I) (NP-ADV زغمْ ragoma despite (NP ذلك *'lika that)) (S (VP أحكي >aHokiy tell I (NP-SBJ-1 *) (NP-OBJ ها hA her/it) (SBAR-ADV و wa while (S (NP-SBJ أنا >anA I) (SBAR-NOM-PRD (WHNP-2 0) (S (VP واثقْ wAviqN certain (NP-SBJ-2 *T*) (SBAR أَنْ >an~a that (S (NP-TPC-3 هُ hu him/it) (VP (PRT لنْ lan not) بغضبْ (NP-SBJ-3 *T*)))))))))))))</pre>

				<p>وَلَكِنِّي رَغْمَ ذَلِكَ أَحْكِيهَا وَأَنَا وَاثِقٌ أَنَّهُ لَنْ يَغْضَبَ</p> <p>But in spite of that I tell it while I'm certain that he won't be angry.</p>
وَ (واو الزائدة)	PART		It is a constituent particle.	<p>(S (VP سَبَقَ sabaqa preceded (SBAR-SBJ (PRT و wa no gloss) 0 (S (VP رَأَيْتُ ra>ayota see you (NP-SBJ *) (NP-OBJ ذَلِكَ *'lika that) (PP-TMP مِنْ min from (NP قَبْلُ qabolu before))))))</p> <p>سَبَقَ وَرَأَيْتُ ذَلِكَ مِنْ قَبْلُ</p> <p>It happened that you saw that before.</p>

7.1.5.1 Tests distinguishing the wa of coordination, NAC clauses introduced by wa, and the circumstantial wa

7.1.5.1.1 wa: SUB_CONJ or CONJ?

A simple coordinated main clause referring to an antecedent in a preceding main clause usually will have additional information, regardless of the first event. The additional information is not about the circumstances that surround the event expressed in the first clause, as in the following examples:

وزين النعش بزهور على شكل رقم 7 وهو رقم القميص الذي كان يرتديه ماثيوز خلال المباريات.
wazu~yna AlnaEo\$u bizuhuwK EalaY \$akoli raqomi 7 wahuwa raqmu Alqamiysi
Al~a*iy kAna yarotadiyhi mAviuwz xilAla AlmubArAti.
And the gasket was decorated with flowers in the shape of the number 7 and this was Matthews' shirt's number.

هَذِهِ الْفَنَةُ مِنَ الْعَمَالِ تَعْمَلُ فِي مَجَالِ عَمَلٍ وَاحِدٍ وَهُوَ مَجَالُ صِنَاعَةِ الْأَحْذِيَّةِ
ha*ihi Alfi}apu min AlEum~Ali taEomalu fiy majAli EamalK wAHidK wahuwa majAlu
SinAEapi Al>aH*iyapi
This category of factory workers works in one job area and it the area of shoe manufacturing.

As a subordinating conjunction, wa introduces a clause that “relates a condition or action simultaneous with an event” (W. Fisher 1972). Below are tests to differentiate the two wa:

- Look for simultaneity in the events of the first and the second clause. For example, looking for the patterns:
 - “something happens to X, while X is/is doing something”

- “X does something to Y, while Y is/is doing something”
- “something happens at/before/after/under Z, while Z is/is doing something”

ماتت ابنته وهي تلد

mAtat <ibonatuhu wahya talidu

She died while she was giving birth

If any of these patterns happen, this is a SUB_CONJ wa.

- Omit the wa and replace the predicate, which is usually an adjective or a verb, in the wa clause by a participle (derived from the same verb) in the indefinite accusative. The substitution will lead to an S-ADV or S-MNR. If this substitution is successful without changing the meaning of the sentence in any way, this is a SUB_CONJ wa.

جاء ضاحكاً → جاء وهو يضحك / جاء وهو ضاحك

NB: Need transliteration and translation here

- Add extra arguments or modifiers after the wa clause, to the verb of the first clause. If this is possible and an additional constituent can be placed after the clause introduced by wa, this clause is a circumstantial modifier of the verb in the main clause. The clause is an SBAR, and the POS tag for wa is SUB_CONJ.

ماتت ابنته وهي تلد بسبب إهمال الأطباء

NB: Need transliteration and translation here

7.1.5.1.2 wa: CONJ (under NAC) or SUB_CONJ?

The following are tests to differentiate between the NAC structures and the subordinating clauses introduced by the circumstantial wa:

- The wa of the NAC structures can be omitted without any other accompanying structural modification. This test is especially valid for prepositional phrases and relative clauses with an overt relative pronoun, preceded by wa. The wa in an NAC structure is always CONJ.

Example with PP-MNR:

7.1.6 <i*/إذ/ ‘because’, ‘as’, ‘suddenly’

<i* has 2 POS values and 2 matching syntactic representations.

7.1.6.1 The subordinating conjunction <i*/ ‘because,’ ‘since,’ ‘as’ (SUB_CONJ)

<i* is a subordinating conjunction annotated as SUB_CONJ when it introduces an adverbial subordinating clause with most of the time an explanatory meaning related to the content of the main clause.

```
(S (PP ك::ka::as
(NP ذلِكَ*`lika::that)
(VP شكَا::$akA+-LRB-null-RRB-::complain/suffer+he/it_[verb]
(NP-SBJ خالد::xAlid::Khalid القحطاني::Al+qaHoTAniy~::the+Qahtani)
(PP-CLR من::min::from
(SBAR ان::>an~a::that
(S "::::nogloss
(NP-TPC-3 (NP الطريقة::Al+Tariyqapa::the+method)
(SBAR (WHNP-1 التي::Al~atiy::which_[fem.sg.])
(S (VP قبض::qubiD+a::be_arrested/
be_seized+he/it
(NP-SBJ-2 *)
(PP في::fiy-::in
(NP (NP ها::-hA::it/her)
(NP-1 *T*))
(PP-CLR على::EalaY::on/above
(NP-2 (NP شقيق::
$aqiyq-::
brother)
(NP يي:: iy::
my))
و::wa-::and
(NP زوجت::
-zawojati-::
wife
(NP هـ::-hi::
his))))))
(VP كانت::kAn+at::be/was/were+it/they/she_[verb]
(NP-SBJ-3 *T*)
(ADJP-PRD وحشية::waHo$iy~apF::brutal/savage))
"::::nogloss) ,::,::nogloss))
(SBAR-ADV إذ::<i*::as
(S (S (VP قيدَا::quy~id+A::be_bound/+they_[both]_[verb]
(NP-SBJ-4 *)
(NP-OBJ-4 *)))
و::wa-::and
(S (VP نزع::-nuziE+a::be_removed/+he/it_[verb]
(NP-SBJ-5 الحجاب::Al+HijAb+u::the+veil+[def.nom.])
(NP-OBJ-5 *)
(PP عن::Ean::from/about/of
(NP رأس::ra>os+i::head/top+[def.gen.]
(NP الأخيرة::Al+>axiyrapi::
the+last/latest/latter)))))))))
```

كَذَلِكَ شكا خالد من أن " الطريقة التي قبض فيها على شقيقي وزوجته كانت وحشية" إذ قيدوا ونزع الحجاب عن رأس الأخيرة
 ka+*alika + \$akaY + khAlid + min + >an~a + AlTariyqapa + qubiDa + fiy+hA+ EalaY +
 \$aqiyq+iy + wa+zawojati+hi + kAnat + waHo\$iy~apF + >i*o + quw~idA + wa+nuziEa +
 AlHijAbu + EalaY + ra>osi + Al>axiyrapi

NB: Need gloss here

Also, Khaled complained about the fact that "the way in which my brother and his wife were treated was brutal", in view of the fact that they were bound and that the wife's scarf was removed from the top of her head

Syntactically, <i* in its general adverbial usage can also be followed by <an~a which is annotated as SUB_CONJ:

(S (S (VP (اعتبرت::wa-::and
 (VP (اعتبرت::wa-::and
 (NP-SBJ *)
 (SBAR (ان::>an~a::that
 (S (NP-TPC-1 رئيس::ra}iysa::president/head/chairman
 (NP (الوزراء::Al+wuzarA'i::the+ministers))
 (VP (ذهب::*ahab+a::go/depart+he/it_[verb]
 (NP-SBJ-1 *T*)
 (NP-ADV (بعيداً::baEiydAF::far_away_[from])
 (PP (في::fiy::in
 (NP (NP (بيان::bayAni-::communique/statement
 (NP (هو::hi::its/his))
 (ADJP (الرسمي::Al+rasomiy~i::the+official))
 (SBAR-PRP (إذ::<i*::because/since/
 (ان::>an~a::that
 (S (NP-TPC-2 موسى::muwsaY::Mousa/Moussa)
 (VP (PRT (لم::lam::did_not)
 يطلب::ya+Tolub+o::he/it+request+[jus.]
 (NP-SBJ-2 *T*)
 (PP-CLR (من::min-::from
 (NP (هو::hu::it/him))
 (NP-OBJ (تسمية::tasomiyapa::
 designation/naming
 (NP (القذافي::Al+qa*~Afiy~::
 the+Qaddafi))))))))))
 واعتبرت أن رئيس الوزراء ذهب بعيداً في بيانه الرسمي إذ أن موسى لم يطلب منه تسمية القذافي
 wa+<iEotabarato + >an~a + ra}iysa + AlwuzarA'i + *ahaba + baEidF + fiy + taqoriyri+hi +
 Alrasomiy~i + <i*o + >an~a + muwsaY + lam + yaTolubo + min+hu + tasomiyapa +
 Alqa*~Afiy

NB: Need gloss here

And she considered that the prime minister went too far in his official report, since Musa did not ask him to mention Gaddafi

7.1.6.2 The particle <i* /‘suddenly’ (PART)

<i* is a particle annotated as PART when it is an equivalent of the particle <i*A. It is called in Arabic the <i* of surprise (إذ الفجائية) and begins a main clause conveying the meaning of ‘suddenly.’

(S (S (VP fataHat فَتَحَتْ opened
 (NP-SBJ nabiylapu نَبِيلَةُ Nabila)
 (NP-OBJ Eayonayo عَيْنَيَّ eyes
 (NP hA هَا her))))
 fa فَ so
 (S (PP (PRT <i*A إِذَا suddenly)
 bi بِ with
 (NP hA هَا))
 (NP-SBJ *)
 (PP-PRD fy فِي in
 (NP nafosi نَفْسٍ same
 (NP algurofapi أَلْغُرْفَةُ the room))))))
 فتحت نبيلة عينيها فإذا بها في نفس الغرفة
 fatahat nabylapu EayonayhA fa >i*A bi hA fy nafsi algurofapi
 opened + nabila + eyes her + so + if + with + her + in + same + room
Nabila opened her eyes and all of a sudden she was still in the same room.

7.1.7 Hat~Y (حَتَّى)

Arabic	POS Tag	Gloss	Syntactic /Semantic Context	TB
حتى	PART	Even	Hat~aY is a particle and is annotated as PART when it is used to emphasize or stress the role of an entity in an event. It has the meaning of inclusion like the conjunction Hat~Y but does not coordinate 2 constituents	<p>(S (VP (PRT لم يجرؤ (NP-SBJ *) (PP (PRT حتى) على (NP التفكير))))))</p> <p>لم يجرؤ حتى على التفكير Not dared he even on the thinking He didn't even dare to think.</p>
حتى	CONJ	And	It coordinates two constituents with the same function Very rare	<p>(S (VP darsoṭu دَرَسْتُ (NP-SBJ *) (NP-OBJ (NP kul~a كُلُّ (NP الإِحْتِمَالَاتِ) حتى Hat~aY (NP (ADJP الضَّغِيرَةِ) (PP min مِنْ</p>

				<p>(NP ها))))))</p> <p>درستُ كلَّ الإِحتِمالاتِ حتَّى الصَّغِيرَةِ مِنْهَا</p> <p>Studied I all the possibilities and the insignificants of its</p> <p>I studied all the possibilities even the insignificant ones.</p>
حتَّى	PREP	Until/ up to	It indicates the end point of an action or its extent.	<p>(S ::wa-::and (VP استكملت (NP-SBJ سيول::siyuwl::Seoul) (NP-OBJ بناء::binA'+a (NP خط::xaT~i (NP السكة::Al+sik~api))) (PP-LOC حتى::Hat~aY::until/up_to (NP (NP المنطقة::Al+minoTaqapi (ADJP المنزوعة (NP السلاح::Al+silAHi))) وَإِسْتَكْمَلْتُ سَيُولَ بِنَاءَ خَطِّ السِّكَّةِ حَتَّى الْمِنَظَّةِ الْمَنْزُوعَةِ السِّلَاحِ</p> <p>Finished it Seoul construction the railway to the area the removed the weapons Seoul finished the construction of the railways to the disarmed area.</p> <p>(PP-TMP من::min::from (NP الآن::Al+ n+a)) (PP-TMP حتى::Hat~aY::until/up_to (NP الصيف::Al+Sayofi::the+summer (NP المقبل::Al+muqobili))))))</p> <p>من الآن حتى الصيف المقبل</p> <p>From the now to the summer the coming</p> <p>From now to next summer</p>
حتَّى	SUB_CO NJ	So that Unless	it introduces an adverbial subordinating clause with the meaning of purpose, causality, or condition	<p>(S (VP (PRT لن:::-lan) يَحْلُمُوا:::ya+Holum+uwA (NP-SBJ (-NONE- *))) (PP-CLR ب::bi-:::with/by (NP الامن)) (SBAR-ADV حتى::Hat~aY::until/up_to (S (VP يرفعوا (NP-SBJ (-NONE- *))) (NP-OBJ أيدي:::ayodiy+a- (NP هم:::-hum) (PP-CLR عن:::Ean (NP أمت (NP نا:::-nA)))) لَنْ يَحْلُمُوا بِالْأَمْنِ حَتَّى يَرْفَعُوا أَيْدِيَهُمْ عَنْ أُمَّتِنَا</p> <p>Not dream they of security until lift hands their on nation ourthey won't dream of security until they lift their hands over our nation.</p>

7.1.8 1A (لا)

Arabic	POS Tag	Gloss	Syntactic /Semantic Context	TB
لا [لا النافية للجنس]	PSEUDO_VERB	{PSEUDO_VERB}	It negates an equational sentence.: لا بد لا مَنَصَّ , لا شَكَّ , (لا حاجة , etc.	<p>(S (VP لا:: -lA::no/not/non- (NP-SBJ (NP اَحد:: >aHada::one) (SBAR (WHNP- 0) (S (VP يتجرأ ya+tajar~a>+u dare (NP-SBJ-1 *T*) (PP-CLR على::EalaY on (NP رفض::rafoDi refusing (NP (NP هذا h`*A this) (NP العمل: Al+Eamali the job)))))</p> <p>لا أَحَدَ يَتَجَرَّأُ عَلَى رَفْضِ هَذَا الْعَمَلِ</p> <p>No one dare on refusing this work</p> <p>Nobody dares to refuse this job</p>
لا لا...لا	NEG_PART	Not Neither..nor..	<p>It negates the meaning expressed by a verb in the present</p> <p>It negates persons or things. There must be two negated elements for 1A to have the value of a negative particle in this context.</p>	<p>- Negation of the verb</p> <p>(S (NP-TPC-1 (NP (NP هذا h`*A this) (NP النوع AlnawoEu the kind)) (PP من min of (NP (NP الحشائش AlHa\$A}i\$}i the herbs))))</p> <p>(VP (PRT لا::lA::no/not/non-) ينمو:yanomuw grows it (NP-SBJ- 1*T*) (PP-LOC في fiy in (NP (NP المناطق AlmanATiqi the areas الحارة AlHArapi the hot))))</p> <p>هَذَا النَّوْعُ مِنَ الْحَشَائِشِ لَا يَنْمُو فِي الْمَنَاطِقِ الْحَارَّةِ</p> <p>This kind of herbs not grows it in the areas the hot</p> <p>This type of herbs doesn't grow in hot areas</p> <p>(S (VP عادَ EAda came back (NP-SBJ *) (PP-DIR (PP (PRT لا::lA:: not/non-) من::min::from (NP ليبيا liybiyA Libya)) بل balo but (PP من::min::from (NP فلسطين falasoTiyin Palestine))))</p>

				<p>عادَ لا من ليبيا بل من فلسطين</p> <p>Came back he not from Libya but from Palestine He returned not from Libya but from Palestine.</p> <p>Note: In this example, 1A is annotated as PART because the assertion ‘bal min falaSotiyn’ follows the negation. When the assertion precedes the negation, 1A is annotated as CONJ as shown in the section 7.2.6.2. above.</p>
لا العطف]	CONJ	And not	it coordinates two adjacent constituents sharing the same syntactic function with the second conjunct being negated	<p>(S (NP-SBJ المسألة (ADJP-PRD سياسية::siyAsiy~apN لا::1A::no/not (قضائية)</p> <p>المسألة سياسية لا قضائية</p> <p>The matter political not judiciary The issue is not political but judiciary.</p>
لا	INTERJ	No!	Direct speech	<p>(S wa فو (VP >ajAba أجاب (NP-SBJ *) (PP-MNR bi ب (NP Hid~apK (جدة) : (S " (INTJ 1A لا (VP (PRT 1A لا >uwAfiqu أوافق (NP-SBJ *) (PP-CLR EalaY على (NP ha*A (هذا)) ")))</p> <p>وأجاب بحدة : لا، لا أوافق على هذا</p> <p>And answered he with violence : No, not agree I on this And he answered sharply: “No, I don’t agree to this”.</p> <p>The interjection 1A can occupy a nominal position as in the example below. In this case its POS value will be still INTERJ but at the tree level, it will take an NP head with whatever relevant function dashtag.</p>

				<p>(S (NP-TPC-2 جواب::jawAb+u (NP خبائ::xubarA}+i- (NP نا::nA::our)) (VP كان::kAn+a (NP-SBJ-2 (-NONE- *T*)) (NP-PRD لا::lA::no))))))</p> <p>جَوَابُ خُبْرَانِنَا كَانَ لَا</p> <p>Answer experts our was No The answer of our experts was no.</p>
--	--	--	--	---

A useful test to differentiate between *lA* as a conjunction and *lA* as a negative particle is to look at the syntactic context where both occur.

- When the *lA* is used before two negated elements; that is to say, when there is a *lA* before the first element and another *lA* before the second, then the two *lA*s are annotated as negative particles. In these contexts, the second *lA* is preceded by a *wa-*, a phenomenon that is not possible when *lA* is a conjunction.

رَأَيْتُ لَا زَيْدًا وَلَا عَمْرًا

ra>aytu + *lA* + zaydAF + wa + *lA* + EamrAF

I saw neither zayd nor Eamr

- In the same contexts mentioned above, *lA* as a conjunction appears only before the second conjunct with the meaning of “this **not** this”. This serves as the confirmation of a fact over another.

رَأَيْتُ زَيْدًا لَا عَمْرًا

Saw I zayd not Amr

I saw Zayd not Amr

7.1.9 The particle and the preposition *إِلَّا* <il~A

7.1.9.1 <il~A as a Restrictive Particle

<il~A as a restrictive particle restricts the effect of an event to the constituent that <il~A precedes. In other words, the restriction applies to that constituent excluding other previously mentioned constituents. This restriction happens only when what was mentioned before is negated via the use of negative particles or verbs of negation such as *layosa* and *lA*. In all the examples below, one can substitute أَحَدٌ (>aHad/nobody) with the phrase *إِلَّا* (<il~A zayd/except for zayd) and vice versa.

(S (NP-TPC-4 سراب
(NP (التفاوض))
(VP (PRT لن::lan::[will]_not/never)
يصبح::
(NP-SBJ-4 *T*)
(NP-PRD حقيقة)
(NP-TMP (PRT لا::<il~A)
بعد::baEoda::after
(S-NOM (VP عودة::Eawodapi
(NP-SBJ الجولان
السوري
(المحتل)
(PP-CLR الى::<ilaY
(NP اصحاب
(NP (ه)))))))))
سَرَابُ التَّفَاوُضِ لَنْ يُصْبِحَ حَقِيقَةً إِلَّا بَعْدَ عَوْدَةِ الْجُولَانِ السُّورِيِّ الْمُحْتَلِّ إِلَى أَصْحَابِهِ

NB: Need transliteration here

Mirage the negotiations not become reality only after coming back the Golan the Syrian the colonized to owners its

The phantom of negotiations won't be a reality until the Syrian colonized Golan comes back to its owners

(S (و::wa-::and
(VP (PRT ما::mA::not)
كان::kAn+a::be/was/were+he/it_[verb]
(VP يلعب::ya+loEab+u
(NP-OBJ ها::hA::it/them/her)
(NP-SBJ (NP (PRT لا::<il~A)
سياسي::siyAsiy~N)
(PP ب::bi-::by/with
(NP حجم::Hajom+i
(NP الملك::Al+maliki
(NP الراحل::Al+rAHili)))))))))
وَمَا كَانَ يَلْعَبُهَا إِلَّا سِيَّاسِيٌّ بِحَجْمِ الْمَلِكِ الرَّاحِلِ

NB: Need transliteration here

And not was plays it it only politician with size the king the late

And none could have played it if it was not someone with the importance of the late king

- when <il~A modifies a constituent that is **in apposition with another one**, this constituent takes necessarily the case ending of the constituent it is in apposition with.

(S (VP (PRT لا::lA::no/not/non-)
يبقى::ya+boqaY
(NP-SBJ (NP أحد::AHadN
(NP (PRT لا::<il~A)
الكبار::Al+kibAr+u::))
(PP-TMP في::fiy::in
(NP نهاية::nihAy+ap+i
(NP اللعبة::Al+luEobapi))))))
لَا يَبْقَى أَحَدٌ إِلَّا الْكِبَارُ فِي نِهَآيَةِ اللَّعْبَةِ

NB: Need transliteration here

Not remain he one except the oldest at end the game

Only the big ones will remain till the end of the game

- <il~A also occurs before prepositions, even when what precedes the prepositional phrase is not negated.

```
(S (و :wa-::and
  (VP (PRT لن :lan::[will]_not/never)
    (يُسمَح :yu+somaH+a
      (NP-SBJ-1 (-NONE- *))
      (PP-CLR ب :bi-::with/by
        (NP-1 (NP الدخول :daxwul-
          (PP الي :ilay
            (NP ها :ha- )))
        (PP-MNR (PRT لا :la-::<il~A)
          ب :bi-::by/with
          (NP اذن :adhin-::<i*onK
            رسمي :rasomiy~+K))))))
```

وَلَنْ يُسَمَّحَ بِالْدُخُولِ إِلَيْهَا إِلَّا بِإِذْنٍ رَسْمِيٍّ

NB: Need transliteration here

And will not be allowed with the entry to her except with permission official

Its access won't be allowed except with official permission

7.1.9.2 <il~A as a Preposition

<il~A as a preposition heading a noun phrase occurs in affirmative constructions as a modifier of another noun phrase. It assigns the accusative to the noun it heads.

- <il~A + noun in the accusative

```
(S (VP جاء :ja'a
  (NP-SBJ الجميع :aljami'a
    (PP إلَّا :illa
      (NP زيدا :zayda))))
```

جاء الجميع إلَّا زيدا

NB: Need transliteration here

Came it the all except Zayd

Everybody came except Zayd

- <il~A + an SBAR introduced by >an~a':

(S (PP-PRD من::min::from
 (NP الواجب::Al+wAjibi::the+duty/obligation/requirement))
 (NP-SBJ توافر::tawAfuru::availability/abundance/existence
 (NP الحوار السياسي::Al+siyAsiy~i::the+political))
 (PP إلا::<il~A::however/except/other
 (SBAR ان::>an~a-::that
 (S (NP-SBJ هو::hu::it/he)
 (ADJP-PRD معدوم::maEoduwmN::absent/vanished)
 (NP-TMP اليوم::Al+yawoma::the+today))))))
 مِنْ الْوَاجِبِ تَوَافُرُ الْحِوَارِ السِّيَاسِيِّ إِلَّا أَنَّهُ مَعْدُومٌ الْيَوْمَ

NB: Need transliteration here

NB: Need gloss here

The presence of a political dialogue is essential except that it is absent today.

- <il~A as preposition can also occur with the meaning of ‘minus’:

(S و::wa-::and
 (PP-TMP في::-fiy::in
 (NP (NP الأولى::Al+>uwlaY::the+first)
 (PP إلا::<il~A::except)
 (NP ربعاً::ruboEAF::quarter/fourth)))
 (VP اجتمع::{ijotamaE+a::meet/confer+he/it_[verb]
 (NP-SBJ الأعضاء::Al+>aEoDA'u::the+members)))
 وَفِي الْأُولَى إِلَّا رُبُعًا اجْتَمَعَ الْأَعْضَاءُ

NB: Need transliteration here

NB: Need gloss here

NB: Need translation here

Note: Watch out for structures involving <il~A followed by an SBAR introduced by >an, in which case <il~A can either be a restrictive particle or a preposition. This is different from SBARs introduced by >an~a where <il~A there is always a preposition.

(FRAG (NP-ADV رغم
 (NP مرضه
 (NP له)))
 (PP إلا
 (SBAR أن
 (S (NP-TPC-1 له)
 (VP جاء
 (NP-SBJ-1 *T*))))))
 رغم مرضه إلا أنه جاء

NB: Need transliteration here and in the tree

Despite illness his except that he came

Despite his illness, he came

7.1.10 mun*u and mu*

mun*u and mu* are prepositions and should be tagged as PREP except in those very rare cases where they are followed by a graphemically present nominative on the word that follows them notably the >alif of the dual as in yawmAni (يومان). In these cases, mun*u and mu* have the

POS value of noun. They should also be annotated as nouns when they have an equational sentence as a complement.

7.1.11 EalY : PREP or NOUN

EalY should always be considered as a preposition except when it is preceded by the preposition “min” in which case it should have the tag NOUN In the example من على الطاولة (min EalaY AlTAwilapi/from top (the)table), EalY is a noun and should be tagged as NOUN.

7.1.12 The particle and preposition la-

Arabic	POS Tag	Syntactic context	TB (not available)
لـ -li	PREP To For	Always governs a nominal constituent (noun, for example) or an SBAR.	(PP li لـ (NP baHv+I بَحْثُ (NPAl+mawoqif+i (المَوْقِفِ))) لِيَحْثُ المَوْقِفِ To discuss the situation
لـ li	JUS_PART Lets	It expresses a meaning of request or command. It assigns the jussive to the verb it cliticizes to.	(S(VP (PRT li لـ) Na*habo نَذْهَبُ (NP-SBJ*))) لِنَذْهَبِ Let's go
لـ - la	EMPHATIC_PART Definetly Certainly	Always preceding a verb and indicates reinforcement of the action the verb is describing	(S(VP (PRT la لـ) >aDriban~a أَضْرِبَنَّ (NP-SBJ*) (NP-OBJ -ka كَ)) لَأُضْرِبَنَّكَ Certainly will beat I you I will certainly hit you.
لـ la	RC_PART So	It is used in conditional sentences introducing the apodosis sentence/main clause	(S (SBAR-ADV إِنْ <ino أَعْطَيْتَ (NP-SBJ *))) (VP (PRT la لـ) أَكْرَمْتُ (NP-SBJ *) (NP-OBJ كَ -ka)) إِنْ أَعْطَيْتَ لَأَكْرِمَنَّكَ If gave you so honor I you If you gave, I would have honored you.

7.1.13 layosa/ليس/ 'not be', 'not' : layosa has 2 POS values and 2 matching syntactic annotations.

7.1.13.1 The perfective verb layosa/ 'not be' (PV)

layosa is annotated as a perfective verb (PV). Like kAna, layosa takes a subject and a predicate in the accusative. It inflects for gender and number to agree with its subject.

- When it requires a subject and a predicate, layosa is annotated as in the example below:

```
(S (VP ليست::layos+at::not_be+it/they/she [verb]
  (NP-SBJ المسألة::Al+maso>alapu::the+issue/affair/matter)
  (ADJP-PRD خطيرة::xaTiyrapF::serious/grave/dangerous/significant))
ليسَت المسألة خطيرة
```

NB: Need transliteration here

Wasn't the issue important

The matter wasn't important

- layosa is annotated like an auxiliary when it has a VP complement.

```
(S (VP ليس-lays+a::- /was/ +he/it_[verb]
  (VP يقيم::yu+qim+u he/it+live+[ind.]
    (NP-SBJ *)
    (PP-LOC فيfiy-:in
      (NP القاهرةAlqAhirapi: Cairo))))
ليس يقيم في القاهرة
```

NB: Need transliteration here

Wasn't he resides he in Cairo

He doesn't reside in Cairo

7.1.13.2 The negative particle layosa 'not' (PART)

layosa is a negative particle annotated as NEG_PART in the following contexts:

- layosa is not involved in a subject-predicate relationship.
- layosa does not inflect for gender or number to agree with a preceding constituent.
- layosa is part of a second conjunct in a relationship of coordination (with or without the presence of a conjunction) with a previous constituent or conjunct. These conjuncts have the same syntactic function but are not necessarily of the same syntactic category in which case they need to be headed by a UCP node.


```
(S (ADVP-PRD هناك::hunAka::there_is/are)
  (NP-SBJ (NP اهتمام::AhtmAmN::nogloss
    متزايد::mutazAyid+AF::increasing/growing+[acc.indef.])
  (PP لـ::li-::for/to
    (NP (NP حل::-Hal~i::solution/dissolution/cancellation
      (NP النزاعات::Al+nizAEAti::the+conflict/struggle))))
  (PP-MNR من::min::from
    (NP خلال::xilAli::during/through
      (NP (NP الحوار::Al+HiwAri)
        و::wa-::and
        (NP (PRT ليس::-layos+a::not+[it] )
          المواجهة::Al+muwAjahapa::the+confrontation))))))
  هناك إهتمامٌ مُتزايدٌ لِحَلِّ النزاعاتِ مِنْ خِلالِ الحوارِ وَلَيْسَ المُواجهةُ
```

NB: Need transliteration here

NB: Need gloss here

NB: Need translation here

7.2 li heads PP only not SBAR

The preposition “li” always functions as a preposition and always heads a PP. It can no longer be treated as heading an SBAR.

```
(S (VP -ya+jib+u::يَجِبُ::he/it+be_necessary/be_incumbent+[ind.]
  (NP-SBJ Al+{inotiZAr+u::نَتَظَرُ::the+waiting/anticipating+[def.nom.]
  (PP-PRP li-::لـ::for
    (SBAR *0*
      (S (VP -ta+tagay~ar+a::تَتَغَيَّرُ::it/them/her_to+
        be_changed +[sub.]
        (NP-SBJ *))))))
```

يجب الانتظار للتغير

yajib+u + Al+<inotiZAr+u + li+tatagay~ar+a
 necessary + the+waiting + to+change
Waiting is necessary for it to change

7.3 Dates

c.f. in 2.3.10.

7.4 Compass directions

Compass directions are basically calques in Arabic, and they are done flat:
 715-11-1 (24-26) south east

NB: this section needs an example. Also potentially discussion with Mohamed.

7.5 Sports scores

Sports scores are now annotated in a parallel way to the English Treebank annotation. In previous Arabic Treebank corpora, scores were annotated as single token ADVPs, but now the tokens will be split and they will be annotated as NP-ADV.

The Phillies won 6-5.

```
(NP-ADV (NP 6)
        (PP -
          (NP 5) ) )
```

The Phillies won 6 5.

```
(NP-ADV (NP 6)
        (NP-ADV 5) )
```

Any tokenization that needs to be revised should be labeled as "X" so that the tokenization can be fixed during the publication process.

```

(S and |wa- | و-
  (NP-SBJ this |-h`*A | -ه`ذ)
  (NP-PRD arrangement/organization/preparation |tarotiybu | تَرْتِيْبُ
    (NP the+groups/teams/troupes |Alfiraqi | الفِرَقِ))
  : )
(FRAG (NP collection/group/bloc |majomuWapu | مَجْمُوعَةُ
  (NP the+Atlantic/NATO |Al>aTolasiy~ | الأَطْلَسِي))
  :)
(FRAG (LST 1)
  -
  (NP Boston |buwsoTun | بُوْسْطُنْ
    Celtics |salotiks | سَلْتِكْس)
  (NP-ADV (NP 6)
    (PP -
      (NP 2)))
  (FRAG (LST 2)
  -
  (NP New_Jersey |niyuwjiyroziy | نِيُوجِيْرُزِي
    Nets |nits | نِتْس)
  (NP-ADV (NP 6)
    (PP -
      (NP 2)))
  (FRAG (LST 3)
  -
  (NP Orlando |>uwrolAnoduw | أُورْلَانْدُو
    Magic |mAjiyk | مَاجِيْك)
  (NP-ADV (NP 5)
    (PP -
      (NP 3)))
  (FRAG (LST 4)
  -
  (NP Philadelphia |fiylAdilofiyA | فِيلَادِلْفِيَا
    76)
  (NP-ADV (NP 4)
    (PP -
      (NP 3))))

```

وهذا ترتيب الفرق: مجموعة الأطلسي: 1. بوسطن سلتيكس 6-2. نيو جيرزي نتس 6-2. 3. أورلاندو ماجك 5-3. 4. فلادلفيا 4-3

wa+h`*A + tarotiyb+u + Al+furaq+i + majomuWapu+u + Al>aTolasiy~ + 1 + buwsoTun +
 salotiks + 6-2 + 2 + niyuwjiyroziy + nits + 6-2 + 3 + >uwrolAnduw + mAjiyk + 5-3 + 4 +
 fiylAdilofiyA + 4-3
 and+this + arrangement + the+teams + group + the+Atlantic + 1 + Boston + Celtics + 6-2 + 2
 + New Jersey + nits + 6-2 + 3 + Orlando + Magic + 5-3 + 4 + Philadelphia + 4-3
*Here is the teams arrangement: The Atlantic Group: 1. Boston Celtics 6-2 2.New Jersey Nits
 6-2 3.Orlando Magic 5-3 4.Philadelphia 76ers 4-3*

7.6 Internal combined with external modification, both are annotated (both times count!)

Cf. 2.3.12.8.

8 Arabic terminology

NB: need glosses throughout this section 8

8.1 FAEil, nA&ib fAEil and mubtada'

These three categories are annotated as subjects in the Treebank.

8.1.1 The fAEil الفاعل (doer)

This function is reserved to the subject of an active verb.

```
(S and |wa- | و-
  (VP accuse+he/it |-{it~ahama | اتَّهَمَ
    (NP-SBJ the+president/head/chairman |Alr~a}iysu | الرَّئِيسُ
      the+Venezuelan |Alfinizowiyl~iy~u | الْفِنْزَوِيلِيُّ
    (NP-OBJ the+striker |AlmuDoribiyne | الْمُضْرِبِينَ
    (PP-CLR with/by |bi | بِ
      (NP attempt/effort |muHawalapi | مُحَاوَلَةً
        (S-NOM (VP robbing/dispossession |salobi | سَلَبَ
          (NP-SBJ *)
          (NP-DTV the+country/countries |AlbilAdi | الْبِلَادُ
          (NP-OBJ feast/holiday/festival |Eiyda | عَيْدُ
            (NP the+birthday/birth
              |AlmiylAdi | الْمِيلَادُ))))))))
        واتهم الرئيس الفنزولي المضربين بمحاولة سلب البلاد عيد الميلاد
wa+<it~aham+a + Al+ra}iys+u + Al+finizowiyl~iy~u + Al+muDoribiyne +
  bi+muHawalapi + salobi + Al+bilAdi + Eiyda + Al+miylAdi
and + accused + the+president + the+Venezuelan + the+striker + with+attempt +
  dispossession + the+country + feast + the+birthday
The Venezuelan president accused the strikers of an attempt to deprive the country from
Christmas
```

8.1.2 The nA&ib fAEil نائب الفاعل (the substitute of the doer)

This function is reserved to the subject of a passive verb, which would be the object in the active form.

```
(S (VP قتل::qutil+a::be_killed+he/it_[verb]
  (NP-SBJ-1 ثلاثة::valAv+ap+u::three+[fem.sg.]+[def.nom.]
    (NP junuwd+K::soldiers+[indef.gen.]
    (NP-OBJ-1 *)
    قتلت ثلاثة جنود
qutil+a + valAv+ap+u + junuwd+K
be killed + three + soldiers
Three soldiers were killed
```

8.1.3 The mubtada' المبتدأ (what is started with)

This function is reserved for the subject of an equational sentence or a topicalized subject or object in a verbal sentence or a topicalized NP with a resumptive pronoun.

```
(S (wa-::wa-::and
  (NP-SBJ هذه::h`*ihi::this/these)
  (NP-PRD (NP ظاهرة::ZAhir+ap+N::phenomenon+[fem.sg.]+[indef.nom.])
    (SBAR (WHNP-1 *0*)
      (S (VP (PRT قد::qad::may/might)
        يسبب::yu+sab~ib+u-::he/it+cause/produce/provoke+[ind.]
        (NP-OBJ (NP ها::hA::it/them/her)
          (NP-1 *T*))
        (NP-SBJ الخوف::Al+xawof+u::the+fear+[def.nom.])))
      .)
    )
  )
)
```

وهذه ظاهرة قد يسببها الخوف

wa + h`*ihi + ZAhirap+N + qad + yusab~ib+u+hA + Al+xawof+u
 and + this + phenomenon + may + cause+it + the+fear
And this is a phenomenon that fear might cause

```
(S (fa-::fa-::and/so
  (NP-TPC-1 الجليد::Al+jaliyd+u::the+ice+[def.nom.])
  (VP سيطر::sayoTar+a::dominate/seize/command+he/it_[verb]
    (NP-SBJ-1 (-NONE- *T*))
    (PP-CLR على::EalaY::on/above
      (NP الطرق::Al+Turuq+i::the+roads/ways/methods+[def.gen.])))
  )
)
```

فالجليد سيطر على الطرق

fa + Al+jaliyd+u + sayoTar+a + EalaY + Al+Turuq+i
 so + the+ice + seize + the+roads
So the ice seized the roads

```
(S (PRT >am~A::أما::as_for/concerning )
  (NP-TPC-1 Al+Tul~Ab+u::الطلاب::the+students+[def.nom.] )
  (PRT fa-::فـ::and/so )
  (VP (PRT -qad::قد::[has/have] )
    waqaf+a::وَقَفَ::stop/cease/stand+he/it_[verb]
    (NP-SBJ (NP Eadad+N::عَدَدُ::number/quantity/issue+[indef.nom.] )
      (PP min-::من::from
        (NP (NP -hum::هم::them_[masc.pl.] )
          (NP-1 *T*))
        )
      )
    (NP-LOC Einoda::عِنْدَ::with/at
      (NP madoxal+i::مَدْخَلُ::entrance/introduction+[def.gen.]
        (NP Al+jAmiE+ap+i::الجامعة::the+university+
          [fem.sg.]+[def.gen.] )
        )
      )
    )
  )
)
```

أما الطلاب فقد وقف عدد منهم عند مدخل الجامعة

<am~A + Al+Tul~Ab+u + fa+qad + waqaf+a + Eadad+N + Einoda + madoxal+i +
 Al+jAmiEap+i
 concerning + the+students + so+have + quantity + from + them + at + entrance +
 the+university
Concerning the students, some of them have stood at the entrance of the university

8.2 the mafEuwl bihi المفعول به (the recipient of the action)

This term in Arabic grammar covers any second obligatory argument of the verb. In ATB it is tagged in several ways:

- -OBJ when it is the direct object

(S (VP it/they/she+reward/repay/compensate |tukAfi}u | تكافئ
(NP-SBJ the+state/country |Ald~awolapu | الدَّوْلَةُ)
(NP-OBJ (NP the+discipline/precision |AlAinoDibATa | الإنضباط
and |wa- | و-
(NP (NP the+fidelity/sincerity |-Al<ixolASa | الإخلاص-
(PP in |fiy | في
(NP the+work/action |AlEamali | العمل))))))
تكافئ الدولة الإنضباط والإخلاص في العمل
tukAfi}+u + Al+dawolap+u + Al+<inoDibAT+a + wa+Al+<ixolAS+a + fiy + Al+Eamal+i
reward + the+state + the+discipline + and+the+sincerity + in + the+work
The state rewards for discipline and sincerity at work

- -DTV when it is the indirect object of a ditransitive verb that can be either an NP or a PP

(S (VP you+forbid/deprive |taHorimuwna | تَحْرِمْوُنَا-
(NP-SBJ *)
(NP-DTV us |-nA | -نَا)
(NP-OBJ (NP the+benefit/advantage |Al<ifAdapa | الإفادة
(PP from |min | مِنْ
(NP (NP the+law/statutes/regulations |AlqAnuwni | القانون
(NP 161))))))
تَحْرِمْوُنَا الإفادة من القانون
taHorimuwna+nA + Al+<ifAdap+a + min + Al+qAnuwn+i
deprive+us + the+benefit + from + the+law
You deprive us from the benefit of the law

- -CLR when it is an indirect object with an obligatory preposition or prepositional noun like maEa, bayna, Hawla, etc.

(S (VP >adolaY+[null]::أَذْلَى[نُلل]::provide/present/grant+he/it_[verb]
(NP-SBJ Al+muHAfiZ+u::المُحَافِظ::the+governor+[def.nom.]
Al+mawolaY::المَوْلَى::the+Mawla)
(PP-CLR bi-::بِ::by/with
(NP -taSoriyH+K::تَصْرِيح::declaration/statement+[indef.gen.])))
أَذْلَى المُحَافِظ المولى بتصريح
>adlaY + Al+muHAfiz+u + Al+mawolaY + bi+taSoriyH+K
provide + the+governor +the+Mawla + with+declaration
The Mawla governor provided a declaration

```

(S (VP تعامل
    (NP-SBJ *)
    (NP-CLR مع
        (NP الأزمة ) )
    (PP-MNR ب
        (NP كل
            (NP (هدوء ) ) ) ) ) ) )

```

تعامل مع الأزمة بكل هدوء

NB: Need transliteration here and in the tree

Dealt he with the crisis with all calm

He dealt with the crisis calmly

8.3 The mafEuwl muTlaq المفعول المطلق (the absolute object)

This constituent emphasizes the action or specifies the way in which it took place or the number of times it happened.

This modifier is tagged -MNR when it specifies the way in which the action took place .

```

(S (VP ضرب
    (NP-SBJ *)
    (NP-OBJ هـ)
    (NP-MNR ضربا
        (NP (عنيفا ) ) ) ) )

```

ضربه ضربا عنيفا

NB: Need transliteration here

hit(he) him hiting violent

He hit him very violently

```

(S (VP ضرب
    (NP-SBJ *)
    (NP-OBJ هـ)
    (NP-MNR (ضربا ) ) ) )

```

ضربه ضربا

NB: Need transliteration here

hit(he) him hiting

He hit him violently

It is tagged -ADV when the modifier just emphasizes the action or specifies the number of times it took place.

```

(S (VP ضرب
    (NP-SBJ *)
    (NP-OBJ هـ)
    (NP-ADV (كثيرا ) ) ) )

```

ضربه كثيرا

NB: Need transliteration here

hi(he) him several

He hit him many times.

8.4 The mafEuwl li>ajlih (the purposive object)

This modifier introduces the cause or the result of the action.

It is tagged in ATB as –PRP.

```
(S (VP (PRT qad::قَدْ::[has/have] )
  taEaT~al+at::تَعَطَّلت::be_hindered/be_interrupted/
                                be_obstructed+it/they/she_[verb]
  (NP-SBJ Al+madAris+u::المدارس::the+schools+[def.nom.] )
  (PP-LOC fiy::في::in
    (NP Al+jibAl+i::الجبال::the+mountains+[def.gen.] ))
  (PP-PRP min::من::from
    (NP jar~A'+i::جَرَاء::because_of/as_a_result_of+[def.gen.]
      (NP Al+vuluwj+i::الثلوج::the+snow/ice+[def.gen.]
        wa-::وَ::and
        -Al+jaliyd+i::الجليد::the+ice+[def.gen.] )))))
  قد تعطلت المدارس في الجبال من جرّاء الثلوج والجليد
qad + taEaT~alat + Al+madAris+u + fiy + Al+jibAl+i + min + jar~A'+i + Al+vuluwj+i +
  wa+Al+jaliyd+i
has + be hindered + the+schools + in + the+mountains + from + because of + the+snow +
  and+the+ice
The schools on the mountains were hindered because of the snow and the ice
```

8.5 The HAL (state)

Traditionnaly the HAL is defined as the description of the subject or the object when the action took place. It can also express concession. In these cases we tag the HAL in ATB with the dashtag –ADV.

```
(S-NOM (VP -qiyAm+i-::قيام-::undertaking/carrying_out/setting_up+[def.gen.]
  (NP-SBJ-2 -hi::هـ::its/his )
  (PP-CLR bi-::بـ::by/with
    (NP (NP ->aEomAl+i::أعمال::actions/activities/work+[def.gen.]
      (NP salob+K::سَلَب::robbing/dispossession+[indef.gen.] ))
      (ADJP Eid~+ap+K::عِدَّة::several/numerous/
        many+[fem.sg.]+[indef.gen.] )))
  (S-ADV (VP musotaxodim+AF::مُسْتَخْدِمًا::employing/using+[acc.indef.]
    (NP-SBJ-2 *)
    (NP-OBJ say~Ar+ap+a::سَيَّارَةٌ::car/automobile/
      vehicle+[fem.sg.]+[def.acc.]
      (NP riynuw::رينو::Renault )))))
  قيامه باعمال سلب عدة مستخدماً سيارة رينو
qiyAm+i+hi + bi+>aEomAl+i + salob+K + Eid~ap+K + musotaxodim+AF + say~Arp+a +
  riynuw
undertaking+his + activities + robbing + numerous + car + Renault
His numerous activities of robbery using a Renault car
```


(S (NP-SBJ هو huwa he
 (NP-PRD رجل rajulN man
 محترم muHotaramN respectable
 (NP-ADV رغم ragoma despite
 (NP ماضي mADiy past
 (NP له hi his))))

هو رجل محترم رغم ماضيه

NB: Need transliteration here

He man respectable despite past his

He is a respectable man in spite of his past

The HAl can be also tagged as -MNR when the modifier expresses the way the verb is done.

(S (VP he/it+constitute/form/compose |yu\$ak~ilu | يُشكِّلُ
 (NP-SBJ *)
 (PP-MNR by/with |bi- | بـ
 (NP (clarity/plainness |-wuDuwHK | (وُضُوح-
 (NP-OBJ (NP call/invitation/supplication |daEowapF | دَعْوَةٌ
 (PP to/towards |<ilaY | إلى
 (NP sedition/dissension/discord |fitonapK | فِتْنَةٌ
 sectarian/factional |TA}ifiy~apK | طَائِفِيَّةٌ))))))

يشكل بوضوح دعوة الى فتنة طائفية

yu\$ak~il+u + bi+wuDuwH+K + daEowap+F + <ilaY + fitonap+K + TA}ifiy~ap+K

form + with+clarity + call + to + sedition + sectarian

It clearly forms a call for a sectarian sedition

The HAl also can be an SBAR headed by a subordinating conjunction called wAw Al HAl, and this SUB_CONJ is used only when the HAl is

- Equational sentence

(S (VP دخل daxala came in he
 (NP-SBJ الأستاذ Al>usotA*u the teacher)
 (SBAR-ADV و wa while
 (S (PP-PRD في fiy in
 (NP يد yadi hand
 (NP له hi his))
 (NP-SBJ كتاب kitAbN book))))))

دخل الأستاذ وفي يده كتاب

NB: Need transliteration here

Came in the teacher while in hand his book

He came in, the book in his hands.

- A sentence beginning with a topicalized NP

```
(S (VP دخل daxala came in he
  (NP-SBJ الأستاذ Al>usotA*u the teacher)
  (SBAR-ADV و wa while
    (S (NP-TPC-1 هو huwa he)
      (VP يضحك yaDoHaku laughs he
        (NP-SBJ-1 *T*)))))
```

دخل الأستاذ وهو يضحك

NB: Need transliteration here

Came in the teacher while he laughs

The teacher came in laughing.

- A verbal sentence with a perfect verb preceded by the PRT قد qad

```
(S (VP دخل
  (NP-SBJ الأستاذ )
  (SBAR-ADV و
    (S (VP (PRT قد )
      علت
      (NP-SBJ الابتسامة )
      (NP-OBJ وجهه
        (NP له ) ) ) ) ) ) )
```

دخل الأستاذ وقد علت الابتسامة وجهه

NB: Need transliteration here and in the tree

Came in he the teacher while probably came up it the smile face his

The teacher came in with a smile on his face.

8.6 The tamyiyz

The tamyiyz construction is used mostly in counting and other measurements. It is a complement structure (a mutam~im مَتَمِّم) and it always assigns the ACC (verb argument and/or noun complement).

8.6.1 Noun complement when the noun that completed is...

8.6.1.1 NOUN_NUM (from 11 to 99) → treebanked as Complementation

ثلاثة وأربعون بلدا

valAvapN & wa
>arbaEuwna & baladAF
mA}apN &
wa+xamosuwna &
\$ahiydAF

Forty three country

مائة وخمسون شهيدا

(One)hundred and+fifty
martyr

8.6.1.2 Nouns of measurement (meter, liter, kilo...)→ treebanked as NP-ADV

لتران حلبيا	litraAni & HaliybAF	Two liters (of) milk
خمسون مترا قماشا	xamosuwna & mitrAF & qumA\$AF	Fifty meters fabric
كيلو غرامين شعيرا	kiluwgrAmayni & \$aEiyraf	Two+kilo (of) barley
حفنة ترابا	HafnapN turAbAf	(a) handful (of) dust
قيراطين ذهب	qiyrATayni *ahabAF	Two+karat (of) gold

8.6.1.3 Kam in the interrogative sentence

?كم كتابا قرأت	Kam kitAbAF qara'ta?	How many book(s) read you?
Ka*A when it means an approximative quantity		
اشتريت بكذا دينار افساتين	Ai\$tarayotu & bi+ka*A & diynArAF & fasAtiyna	(I) bought this+much dinar dresses
أعطاني كذا أقلاما وكذا كتب	>aETAniy ka*A >aqlAmAF wa ka*A kutubAF	(he) gave me this+much pens and this+much books

8.6.2 The tamiyz is a verb modifier

S (VP تصبب
(NP-SBJ زيد)
(NP-ADV عرقا))

تصبب زيد عرقا

taSab~ba zaydN EaraqAF
poured he Zayd sweat
Zayd sweated a lot

9 Arabic Treebank Notation

9.1 Node labels and functional "dashtags"

Node (bracket) labels are syntactic (S, NP, VP, ADJP, etc.)

"Dashtags" are more or less semantic function (-SBJ subject, -OBJ object, -ADV adverbial, -TMP temporal, -PRD predicate, etc.). Dashtags are used only if they are relevant, not on every node label.

See section ????? on VP and NP arguments and adjuncts. See also section ????? describing each node label and section ????? describing each dashtag.

9.2 Empty categories

The empty categories are essentially the same as in the Penn English Treebank. The most common being

*	Pro-drop subjects and passive traces
T	WH- traces, NP-TPC trace to subject
ICH	Rightward movement
RNR	Right Node Raising
0	Null complementizer, null WH- word
?	Ellipsis

As in the Penn Treebank, we are not showing any pronominal coreference. Coreference will be indicated only for syntactically controlled empty categories and for exceptional cases such as VP gapping structures.

9.3 Coordination

Coordination is done as adjunction ($Z(Z)$ and (Z)); coordination has the same structure at all phrase levels.

When constituents of different types are coordinated, the outer coordination-level node label is UCP (Unlike Coordinated Phrase). Any shared function tags are put on the UCP label, and not on the lower labels.

See section ????? on coordination (chapter 6?).

9.4 VP template gapping

The technicalities of gapping coreference are different in the Arabic Treebank from the original Penn Treebank, although gapping coreference is done this way in the current English Treebanks as well.

All indices are on the node label itself, and gapping co-reference is shown with '=' on both the template and the filler node labels.

```

(S (VP (VP -fAz+a::فاز::win/be_victorious+he/it_[verb]
  (PP fiy-::في-::in
    (NP -hi::هو::it/him ))
  (NP-SBJ=1 (NP Al+biriyTAniy~+u::البريطاني::the+British+[def.nom.])
    (NP Ady::ادي::Ed
      <iyrfAyn::إيرفاين::Irvine ))
  (NP-TMP=2 EAm+a::عام::year+[def.acc.]
    ,::,::nogloss
    (NP 1999::1999::nogloss )))
  wa-::و-::and
  (VP (NP-SBJ=1 -$uwmAxa::شوماخر::Schumacher )
    (NP-TMP=2 EAm+ayo::عامي::year+two_[acc.]
      (NP 2000::2000::nogloss
        wa-::و-::and
        .2001::.2001::nogloss ))))
    فاز فيه البريطاني ادي إيرفاين عام 1999 وشوماخر عامي 2000 و2001
fAz+a + fiy+hi + Al+biriyTAniy~+u + >adiy + <iyrfAyn + EAm+a + 1999 + wa+$uwmAxa
+ EAm+ayo + 2000 + wa+2001
won + in+it + the+British + Eddie + Irvine + year + 1999 + and+Schumacher + two years +
2000 + and+2001
In which the Brit, Eddie Irvine, won in 1999 and Schumacher in 2000 and 2001

```

9.5 Coreference

Coreference is shown always as a ‘-’ on the node label, never on the empty category token itself. This is a difference from the original Penn English Treebank, although coreference is also shown in this way for the current English Treebanks.

In this treebank, we show syntactic coreference through coindexing, but we do not show discourse coreference. This means that when two items are coreferenced, one of them must be an empty category. It also means that we do not show the coreference of pronouns.

10 References

Bies, A., Ferguson, M., Katz, K., and MacIntyre, R. (1995). *Bracketing Guidelines for Treebank II Style Penn Treebank Project*. University of Pennsylvania, Department of Computer and Information Science Technical Report MS-CIS-95-06.

Blachere, R. and Gaudefroy-Demombynes, M. (1975). *Grammaire de l'arabe classique*. Editions Maisonneuve & Larose. Paris, France.

Fischer, W. (2002). *A Grammar of Classical Arabic* (Translated into English by Jonathan Rodgers). Yale University Press. New Haven & London.

Kremers 2003

Badawi

Schultz

Wright

Mohamed Mohamed

Fassi Fehri

Ibn Yaish

Maamouri and Bies, 2004

11 Index

Index to be created as time permits.

NB: To be added to section ??? From email with/to Dalila (6-21-11)