

ARABIC TREEBANKING

MORPHOLOGICAL ANALYSIS & POS ANNOTATION

Version 3.8

08 June 2009

TABLE OF CONTENT

0. FOREWORD	8
1.0 GENERAL INTRODUCTION	8
1.1 INTRODUCTION.....	9
1.2 OUTLINE	9
2.0 ORGANIZATION & OUTLINE	10
2.1 ORGANIZATION	10
2.2 RATIONALE.....	11
2.2.1 Syncretism.....	11
2.3 MISCELLANEOUS	12
2.3.1 Use of the English Gloss.....	12
2.3.2 Verbal Nouns and Adjectives	12
3.0 SPECIFICATIONS OF POS ANNOTATION	12
3.1 SPECIFICATIONS FOR THE PENN ARABIC TREEBANK MORPHOLOGICAL ANALYSIS & POS ANNOTATION GUIDELINES.....	12
3.1.1 Lexicon and Morphological Analyzer	12
3.1.2 POS Annotation Tool.....	13
3.1.3 POS Annotation Stages and Tool	14
3.1.4 Human Intervention is Necessary in POS	16
3.1.5 POS Annotator Decision Process.....	17
3.2 THE BUCKWALTER TRANSLITERATION TABLE	19
4.0 TAGS, TOKENIZATION, AND MORPHOLOGICAL AND POS ANNOTATION	20
4.1 POS TAGS.....	20
4.1.1 List of Parts of Speech with their Tags	20
4.1.2 List of Tags with their Corresponding Parts of Speech.....	23
4.2 TOKENIZATION.....	25
4.2.1 List of Cliticized Tokens	26

4.2.2 List of Tokens with Cliticized and Non-Clitic Versions.....	28
4.3 MORPHOLOGICAL ANALYSIS AND POS ANNOTATION.....	32
4.3.1 Parts of Speech.....	32
4.3.1.1 Nouns.....	32
4.3.1.1.1 Common Noun (اسم جامد): NOUN.....	32
4.3.1.1.2 Proper Noun (اسم علم): NOUN_PROP.....	33
4.3.1.1.3 Noun Quantifiers: NOUN_QUANT.....	34
4.3.1.1.4 Cardinal numbers: NOUN_NUM.....	37
4.3.1.1.5 Annotation of Nouns.....	38
4.3.1.1.5.1 Gender Inflections.....	38
4.3.1.1.5.2 Number and Definiteness.....	39
4.3.1.1.5.3 Case Marking.....	39
4.3.1.1.5.4 Illustration of the POS Annotation for Nouns.....	39
4.3.1.2 Pronouns.....	39
4.3.1.2.1 Personal Pronouns.....	40
4.3.1.2.1.1 Unbound Pronouns: PRON.....	40
4.3.1.2.1.2 Bound Pronouns.....	40
4.3.1.2.1.2.1 Object Pronouns: PRON.....	40
4.3.1.2.1.2.2 Possessive Pronouns: POSS_PRON.....	41
4.3.1.2.2 Demonstrative Pronouns: DEM_PRON.....	42
4.3.1.2.2.1 Demonstratives Indicating Proximity.....	42
4.3.1.2.2.2 Demonstratives Indicating Distance.....	43
4.3.1.2.3 Interrogative Pronouns: INTERROG_PRON.....	45
4.3.1.2.4 Interrogative Adverbs: INTERROG_ADV.....	45
4.3.1.2.5 Relative Pronouns: REL_PRON.....	46
4.3.1.2.6 Exclamative Pronoun: EXCLAM_PRON.....	49
4.3.1.3 Adjectives.....	49
4.3.1.3.1 Adjective: ADJ.....	49
4.3.1.3.1.1 Qualitative Adjectives (صفة مُشَبَّهة): ADJ.....	49
4.3.1.3.1.2 Adjectives of Intensity (صيغة المُبالغة): ADJ.....	52
4.3.1.3.1.3 Relational Adjectives (اسم النسبة): ADJ.....	53
4.3.1.3.2 Ordinal Numbers: ADJ_NUM.....	54
4.3.1.3.3 Comparative Adjective: ADJ_COMP.....	54
4.3.1.3.3.1 From the Base Trilateral Verbs (الفعل الثلاثي المجرد).....	55

4.3.1.3.3.2 Important Remarks	56
4.3.1.3.4 Illustration of the POS Annotation for Adjectives.....	56
4.3.1.4 Adjectives and Nouns: Principles for Decision Making (Under Experimentation and Review)	57
4.3.1.4.1 Treatment of Active and Passive Participles.....	57
4.3.1.4.2 Treatment of the Manufactured Gerund (مَصْدَرٌ صِنَاعِي).....	57
4.3.1.5 Verbs	58
4.3.1.5.1 Inflectional Verbs.....	58
4.3.1.5.2 Pseudo-Verbs: PSEUDO_VERB (الحُرُوفُ الْمُشَبَّهَةٌ بِالْأَفْعَالِ).....	59
4.3.1.5.3 Non-Inflectional Verbs: VERB (الأفعال الجامدة).....	60
4.3.1.6 Adverbs: ADV.....	62
4.3.1.6.1 Regular Adverbs: ADV	62
4.3.1.6.2 Relative Adverb: REL_ADV	65
4.3.1.7 Prepositions: PREP.....	67
4.3.1.8 Interjections: INTERJ.....	74
4.3.1.9 Particles.....	75
4.3.1.9.1 Vocative Particles: VOC_PART	75
4.3.1.9.2 Emphatic Particle: EMPHATIC_PART.....	76
4.3.1.9.3 Restrictive Particles: RESTRIC_PART.....	76
4.3.1.9.4 Negative Particles: NEG_PART.....	77
4.3.1.9.5 Verb Particles: VERB_PART	78
4.3.1.9.6 Future Particle: FUT_PART.....	79
4.3.1.9.7 Interrogative Particles: INTERROG_PART.....	79
4.3.1.9.8 Focus Particle: FOCUS_PART	79
4.3.1.9.9 Response Conditional Particles : RC_PART	80
4.3.1.9.10 Connective Particle: CONNEC_PART	81
4.3.1.9.13 Jussive Particle: JUS_PART	82
4.3.1.9.12 Other Particles: PART	83
4.3.1.10 Coordinating Conjunctions: CONJ.....	84
4.3.1.11 Subordinating Conjunctions: SUB_CONJ.....	87
4.3.2 POS Summary Chart.....	93
5.0 ADDITIONAL INFORMATION ON ANNOTATION & POLICIES	103
5.1 POS ANNOTATION.....	103
5.1.1 List of Former Prepositions and/or Adverbs	103

5.1.2 List of Former Pronouns	107
5.1.3 List of Former Particles.....	108
5.2 POS POLICIES.....	109
5.2.1 Cardinal Numbers vs. Ordinal Numbers	109
5.2.2 The Pseudo-Verb <in~a (إِنَّ) & the Subordinating Conjunction >an~a (أَنَّ).....	111
5.2.2.1 In Direct Speech.....	111
5.2.2.2 In Indirect Speech.....	111
5.2.2.3 Default Policy	111
5.2.3 Al- : The Determiner and Relative Pronoun	111
5.4 ANNOTATION OF LAYOSA	112
5.5 MISCELLENOUS POS VALUES [UNDER CONSTRUCTION].....	112
6.0 PARTS OF SPEECH WITH MORE THAN ONE VALUE	113
6.1. >AY/>AY~ (أي/أي)	113
6.2 XALA, EADA, AND HASA (خَلا, عَدا, حاشا)	115
6.3 FA- (فأ-).....	116
6.4 MA (ما).....	120
6.5 WA- (و-).....	122
6.6 HAT~Y (حتى).....	123
6.7 LA (لا)	124
6.8 THE PARTICLE AND THE PREPOSITION إنا <il~A.....	124
6.8.1 <il~A as a Restrictive Particle	125
6.8.2 <il~A as a Preposition	125
6.9 MUN*U AND MU*	125
6.10 THE PREPOSITION EALY	126
6.11 THE PARTICLE AND PREPOSITION LA-	126
7.0 INTERACTION BETWEEN POS AND TREEBANK [UNDER CONSTRUCTION]	127
7.1 ADJECTIVES AND NOUNS: TESTS [UNDER EXPERIMENTATION AND REVIEW]	127
7.1.1 Insertion of a Missing Lexical Head.....	127
7.1.1.1 The Case of Relational Adjectives	127
7.1.1.2 The Case of Active and Passive Participles	129
7.1.1.3 The Case of إسم الصفة (<ism AlSifap)	129
7.1.2 Homographs: Semantic Tests	130

7.1.2.1 The Case of Manufactured Gerunds	130
7.1.2.2 The Case of صفة /Sifap	131
7.1.2.3 The Case of Active and Passive Participles	131
7.1.3 Morphological Tests.....	132
7.1.4 Adjectives in Adverbial Positions	132
7.1.4.1 Nouns.....	133
7.1.4.2 Adjectives	133
7.1.5 Adjectival vs. Nominal Modification.....	133
7.1.5.1 Verbal Corroboration/Lexical Corroboration (توكيد لفظي).....	133
7.1.5.2 Paraphrased/Periphrastic Corroboration توكيد معنوي.....	134
7.1.5.3 Modification with Numerical Nouns	134
7.1.5.4 Modification with maSdars	134
7.1.5.5 Special cases	134
7.2 TREES FOR POS WITH MORE THAN ONE POS VALUE	135
7.2.1 fa- /فـ ‘and, so, then. so that’	135
7.2.1.1 The conjunction fa- (فاء العطف) /fA’ AIEatf) ‘and/so’ (CONJ).....	135
7.2.1.2 The particle fa-.....	138
7.2.1.2.1 fa- as a connective particle ‘and/so’ (CONNec_PART).....	138
7.2.1.2.2 fa- as a response conditional particle ‘then’ (RC_PART)	139
7.2.1.3 The subordinating conjunction fa- (فاء السببية) /fA’ Alsababy~ap) ‘ so that’ (SUB_CONJ).....	140
7.2.2 <il~A/إلا/’except’, ‘except for’.....	141
7.2.2.1 The restrictive particle <il~A (RESTRIC_PART).....	142
7.2.2.2 The preposition <il~A (PREP).....	146
7.2.3 EadA عدا/ ‘except’, xalA خلا/ ‘except’, and HASA حاشا/’except’	149
7.2.3.1 The prepositions EadA, xalA, and HASA (PREP).....	149
7.2.3.2 The verbs EadA, xalA, and HASA (VERB)	150
7.2.4 layosa/ليس/ ‘not be’, ‘not’	152
7.2.4.1 The perfective verb layosa/ ‘not be’ (PV).....	152
7.2.4.2 The negative particle layosa ‘not’ (PART).....	153
7.2.5 Hat~aY/حتى/ ‘and even’, ‘so that’, ‘even’, ‘until’.....	154
7.2.5.1 The conjunction Hat~aY ‘and even’ (CONJ).....	154
7.2.5.2 The subordinating conjunction Hat~aY ‘so that’ (SUB_CONJ).....	155
7.2.5.3 The particle Hat~aY ‘even’ (PART).....	156
7.2.5.4 The preposition Hat~aY ‘until/up to’ (PREP)	157

7.2.6	IA/ لا/ ‘not’, ‘no’, ‘not be’	158
7.2.6.1	The pseudo verb IA ‘not be’ (PSEUDO_VERB).....	158
7.2.6.2	The conjunction IA ‘not’ (CONJ).....	160
7.2.6.3	The negative particle IA ‘no/not’ (NEG_PART).....	161
7.2.6.4	The interjection IA ‘no’ (INTERJ).....	163
7.2.6.5	Tests to distinguish IA as a conjunction from IA as a negative particle	164
7.2.7	wa/ و/ ‘and’, ‘while’, ‘by’, ‘with’.....	165
7.2.7.1	The conjunction wa ‘and’ (CONJ).....	165
7.2.7.2	The subordinating conjunction wa ‘while’ (SUB_CONJ)	169
7.2.7.3	The preposition wa (PREP).....	171
7.2.7.3.1	The wa of oath ‘by’	171
7.2.7.3.2	The wa of accompaniment ‘with’	171
7.2.7.5	Tests distinguishing the wa of coordination, NAC clauses introduced by wa, and the circumstantial wa.....	173
7.2.7.5.1	wa: SUB_CONJ or CONJ?	173
7.2.7.5.2	wa: CONJ (under NAC) or SUB_CONJ?.....	175
7.2.8	mA/ ما/ ‘not’, ‘that’, ‘as long as’, ‘as soon as’, ‘what’, ‘some’, ‘not be’	178
7.2.8.1	The negative particle mA/ ‘not’ (NEG_PART).....	178
7.2.8.2	The subordinating conjunction mA/ ‘that’, ‘as long as’, ‘as soon as’ (SUB_CONJ)	179
7.2.8.3	The relative pronoun mA/ ‘what’ (REL_PRON).....	181
7.2.8.4	The interrogative pronoun mA/ ‘what’ (INTERROG_PRON).....	182
7.2.8.5	The exclamative pronoun mA/ ‘what’ (EXCLAM_PRON)	183
7.2.8.6	The noun mA / ‘some’ (NOUN).....	183
7.2.8.7	The non-inflectional verb mA/ ‘not be’ (VERB).....	184
7.2.8.8	The particle mA (PART)	185
7.2.9	<i*A/ إذا/ ‘if’, ‘whether’, ‘suddenly’	186
7.2.9.1	The subordinating conjunction <i*A ‘if’, ‘whether’ (SUB_CONJ).....	186
7.2.9.2	The particle <i*A/ ‘suddenly’ (PART)	190
7.2.10	<i*/ إذ/ ‘because’, ‘as’, ‘suddenly’	192
7.2.10.1	The subordinating conjunction <i*/ ‘because,’ ‘since,’ ‘as’ (SUB_CONJ)	192
7.2.10.2	The particle <i*/ ‘suddenly’ (PART).....	196
7.2.11	turaY/ ترى/ ‘I wonder’.....	197
7.2.11.1	The interjection turaY (INTERJ)	197
7.2.11.2	The non-inflectional verb turaY ‘I wonder’ (VERB)	197
7.2.12	Syntactic Annotation of Particles	198

7.2.12.1 Particles annotated under constituents	198
7.2.12.1.1 Vocative particles (VOC_PART).....	198
7.2.12.1.2 Emphatic particle (EMPHATIC_PART).....	199
7.2.12.1.3 Negative particles (NEG_PART).....	199
7.2.12.1.4 Future Particle (FUT_PART) and Verb Particles (VERB_PART).....	200
7.2.12.1.5 Other Particles	201
7.2.12.2 Particles Annotated under the Sentence Level.....	202
7.2.12.2.1 Interrogative Particles (INTERROG_PART).....	202
7.2.12.2.2 Focus Particle (FOCUS_PART).....	203
7.2.12.2.3 Other Particles	203
7.2.12.3 Summary Chart	205
7.2.12.3.1 Particles under Constituents.....	205
7.2.12.3.2 Particles under the Sentence.....	206
7.2.13 >akvar/ أكثر and >aglab/أغلب.....	207
7.2.13.1 The noun quantifiers >akvar ‘most of’ and >aglab ‘majority of’ (NOUN_QUANT).....	207
7.2.13.2 The comparative adjectives >akvar ‘more’/’more than’/’most’ and >aglab ‘most’ (ADJ_COMP).....	207
8.0 LINGUISTIC INFORMATION	212
8.1. CONSTRUCTIONS OF NUMERALS	213
8.1.1 Cardinal numbers: NOUN_NUM	213
8.1.1.1 Prenominals	213
8.1.1.1.1 Construct State Structure (<iDAfap/إضافة).....	213
8.1.1.1.2 tamyyz (تَمييز) Structure.....	214
8.1.1.2 Postnominals	214
8.1.2 Ordinal numbers: ADJ_NUM	215
8.1.2.1 Prenominals	215
8.1.2.2 Postnominals	216
9. REFERENCES	216
0. FOREWORD	
1.0 GENERAL INTRODUCTION	

1.1 Introduction

The morphological and Part of Speech (POS) annotation guidelines are a reference document for POS annotators at the Linguistic Data Consortium (LDC). The guidelines have been designed to provide common ground and principles for an annotation of Arabic newswire or broadcast news, corpora that reflects the morphological nature and properties of the Arabic Language.

1.2 Outline

Section 2 provides the rationale behind the organization of the present guidelines.

Section 3 gives information about the POS annotation tool as well as the annotation process, describes the stages of the annotation decision making process as well as the annotator's role in this process.

Section 4 lists the parts of speech used in ATB and their tags in two formats. The section on tokenization sets the principles of word segmentation and illustrates its process for a set of function words that present tokenization issues. Tokenization is an important stage prior to POS tagging.

The morphological analysis and POS annotation section takes each part of speech and provides a sample list of entries for open-class categories such as nouns and adjectives as well as an exhaustive list of entries for the closed-class categories of pronouns, pseudo-verbs, adverbs, prepositions, particles, conjunctions, and subordinating conjunctions.

This section also embodies the interaction between morphological (POS) and syntactic (Treebank) annotations. In the Arabic morphological and POS annotation, this morphosyntactic interface is deemed essential as it imparts the annotator with the syntactic context that will lead to the disambiguation of meaning and will determine which of the token's POS values has to be selected. In other words, it provides POS and Treebank annotators information about the syntactic context that governs the assignment of one particular tag rather than another to one single token.

The POS annotation methodology is exemplified in the sections for the parts of speech of nouns, adjectives, and verbs as these represent the parts of speech which carry inflections of features such as person, gender, number, tense, case, and/or mood.

A summary chart for the closed-class parts of speech is provided at the end of the chapter.

Section 5 provides information on new annotation decisions. The chapter also articulates policies that have been adopted concerning the annotation of some complex or problematic tokens.

Section 6 lists tokens with more than two POS values as a way to illustrate their different syntactic distributions and interpretations. Although the values of each token have been discussed individually in their corresponding section, we think that grouping them into tables which illustrate and explain their various values, will provide a faster and more meaningful access to annotators during the decision making process.

Section 7 represents the interaction between POS and Treebank annotations.

Section 8 provides additional linguistic information on some basic aspects of Arabic grammar such as specific syntactic constructions which we think will assist annotators during annotation.

Section 9 provides a list of references.

2.0 ORGANIZATION & OUTLINE

2.1 Organization

A detailed outline for the guidelines is provided in Section 1. The general principle behind the current organization has been the defining of standards and procedures for a POS annotation of Arabic Treebank corpora based on a modern approach to Modern Standard Arabic (MSA). This will be achieved through the provision of key linguistic information as well as a standard methodology.

It is important to note that the present guidelines will be used by both POS and Treebank (TB) annotators. Some parts of the guidelines, which may not directly pertain to POS annotation, have been kept as they include POS information that is considered relevant to TB annotation.

2.2 Rationale

A morphosyntactic framework for the organization of the POS guidelines has been adopted based on the nature and specificities of MSA. Some parts of speech sections also incorporate semantic subdivisions that are thought to inform and assist certain annotation decisions and policies whether at the level of POS or TB. Additional linguistic information has also been appended to the guidelines as a joint reference for POS and TB annotators as well as other users of the guidelines.

2.2.1 Syncretism

Arabic is a type of synthetic languages as it is highly inflectional. Arabic is also a derivational language. Syncretism further adds to the difficulty of analyzing Arabic both for annotators and parsers. On the one hand, this is due to the fact that single morphemes can embed more than one meaning. This is the case of homonyms. On the other hand, one morpheme may have different morphological forms depending on its syntactic distribution. For example, pronouns in Arabic have different morphological realizations depending on whether they occupy a subject or object position. This adds complexity into the analysis of the language.

Because of this high level of syncretism, which for the most part can only be resolved through context, the POS annotation of Arabic had to be based on a morphosyntactic analysis that takes into account the syntactic distribution and function of the part of speech at hand. In other words, the incorporation of the syntactic context for POS annotators will contribute to the disambiguation of the meaning of a token in context during annotation. This will lead to an accurate and reliable POS annotation.

A special section has also been devised to revisit those parts of speech that are very frequent in the corpora but that present some challenges at the level of annotation. The inclusion of syntactic trees has been decided on the ground that the POS and morphological guidelines document also serve as a reference for Treebank annotators whose annotation decision making process is also based on an awareness of the different syntactic behaviors of a single token.

2.3 Miscellaneous

2.3.1 Use of the English Gloss

Whenever possible, glosses were provided in alignment with their corresponding English parts of speech. Curly brackets are used when there is no equivalent English word that matches the POS of the Arabic word. Parentheses were used whenever the gloss does has a different part of speech in English.

2.3.2 Verbal Nouns and Verbal Adjectives

Masdars/gerunds, active and passive participles are annotated as nouns and adjectives at the POS annotation level. A VN (verbal noun) tag will be added to the NOUN or ADJ tag during the TB release. A masdar that has a verbal reading during TB annotation will be given the tag NOUN.VN. An adjectival participle with a verbal reading will get the tag ADJ.VN

3.0 SPECIFICATIONS OF POS ANNOTATION

3.1 Specifications for the Penn Arabic Treebank Morphological Analysis & POS Annotation Guidelines

Details about the POS annotation tools, principles, and procedures are laid out in this section and illustrations and examples of the POS tools and the process are provided as a general overview.

3.1.1 Lexicon and Morphological Analyzer

The Penn Arabic Treebank uses a level of annotation that uses morphological analysis as well as part-of-speech tagging. The automatic Arabic morphological analysis and/or part-of-speech tagging is performed with the Buckwalter Arabic Morphological Analyzer (BAMA), an open-source software package distributed by the Linguistic Data Consortium (LDC catalog number LDC2002L49).

The analyzer consists primarily of three Arabic-English lexicon files: prefixes (299 entries), suffixes (618 entries), and stems (82158 entries representing 38600 lemmas). The lexicons are supplemented by three morphological compatibility tables used for controlling prefix-stem combinations (1648 entries), stem-suffix combinations (1285 entries), and prefix-suffix combinations (598 entries).

3.1.2 POS Annotation Tool

The annotation tool used for the Arabic Morphological Analysis and POS tagging is SelectPOS, a GUI software developed at LDC for the Arabic POS annotation. Written in Python/Qt, it displays the source Arabic text and the POS choices generated from the Tim Buckwalter Morphology Analyzer (see Figure 1). Human annotators can click and select the most suitable POS.



Figure 1: Screen shot of SelectPOS annotation tool

3.1.3 POS Annotation Stages

The current procedure for POS annotation includes the following steps:

1. We begin by segmenting the raw input text and we apply the Buckwalter Arabic Morphological Analyzer to generate a list of candidates for each Arabic segment (i.e., token or word).
2. Human annotators then go through the alternatives for each word and select the appropriate tag if it is present in the list provided (pass1).
3. Next, human annotators check the work of other annotators (pass2).
4. We may, as needed, perform a third pass to correct some particular errors and improve overall quality.

The output from the Buckwalter Arabic Morphological Analyzer (Buckwalter, 2002) is used as the starting point for the morphological annotation and POS tagging of Arabic text. For each input string, the Analyzer provides a fully vocalized solution (in Buckwalter Transliteration), including the word's unique identifier or lemma ID, a breakdown of the constituent morphemes (prefixes, stem, and suffixes), and their POS values and corresponding English glosses, as in the following example:

```

INPUT STRING: ثمن
LOOK-UP WORD: vmn
Comment:
INDEX: P6W30
SOLUTION 1: (vam~ana) [vam~an_1] vam~an/PV+a/PVSUFF_SUBJ:3MS
(GLOSS): appraise/estimate + he/it [verb]
SOLUTION 2: (vam~an~a) [vam~an_1] vam~an/PV+na/PVSUFF_SUBJ:3FP
(GLOSS): appraise/estimate + they [fem.pl.] [verb]
SOLUTION 3: (vaman) [vaman_1] vaman/NOUN
(GLOSS): value/price
* SOLUTION 4: (vamanu) [vaman_1] vaman/NOUN+u/CASE_DEF_NOM
(GLOSS): value/price + [def.nom.]
SOLUTION 5: (vamana) [vaman_1] vaman/NOUN+a/CASE_DEF_ACC
(GLOSS): value/price + [def.acc.]
SOLUTION 6: (vamani) [vaman_1] vaman/NOUN+i/CASE_DEF_GEN
(GLOSS): value/price + [def.gen.]
SOLUTION 7: (vamanN) [vaman_1] vaman/NOUN+N/CASE_INDEF_NOM
(GLOSS): value/price + [indef.nom.]
SOLUTION 8: (vamanK) [vaman_1] vaman/NOUN+K/CASE_INDEF_GEN
(GLOSS): value/price + [indef.gen.]
SOLUTION 9: (vumon) [vumon_1] vumon/NOUN
(GLOSS): one-eighth

```

SOLUTION 10: (vmonu) [vmon_1] vmon/NOUN+u/CASE_DEF_NOM
 (GLOSS): one-eighth + [def.nom.]
 SOLUTION 11: (vmona) [vmon_1] vmon/NOUN+a/CASE_DEF_ACC
 (GLOSS): one-eighth + [def.acc.]
 SOLUTION 12: (vmon_i) [vmon_1] vmon/NOUN+i/CASE_DEF_GEN
 (GLOSS): one-eighth + [def.gen.]
 SOLUTION 13: (vmonN) [vmon_1] vmon/NOUN+N/CASE_INDEF_NOM
 (GLOSS): one-eighth + [indef.nom.]
 SOLUTION 14: (vmonK) [vmon_1] vmon/NOUN+K/CASE_INDEF_GEN
 (GLOSS): one-eighth + [indef.gen.]
 SOLUTION 15: (vmon) [vmon_1] vmon/NOUN
 (GLOSS): eighth
 SOLUTION 16: (vmonu) [vmon_1] vmon/NOUN+u/CASE_DEF_NOM
 (GLOSS): eighth + [def.nom.]
 SOLUTION 17: (vmona) [vmon_1] vmon/NOUN+a/CASE_DEF_ACC
 (GLOSS): eighth + [def.acc.]
 SOLUTION 18: (vmon_i) [vmon_1] vmon/NOUN+i/CASE_DEF_GEN
 (GLOSS): eighth + [def.gen.]
 SOLUTION 19: (vmonN) [vmon_1] vmon/NOUN+N/CASE_INDEF_NOM
 (GLOSS): eighth + [indef.nom.]
 SOLUTION 20: (vmonK) [vmon_1] vmon/NOUN+K/CASE_INDEF_GEN
 (GLOSS): eighth + [indef.gen.]
 SOLUTION 21: (vmn) [DEFAULT] vmn/NOUN_PROP
 (GLOSS): NOT_IN_LEXICON

The morphological annotation process is performed by means of *SelectPOS*, a POS Annotation tool developed at LDC by Huaichuan (Hubert) Jin, that displays the various morphological analysis solutions provided by the Morphological Analyzer. The human annotator must carefully review the available choices (cf. POS Annotator's Decision Process), and then accept one of the solutions, but only if it meets the following criteria: (a) the POS tag is correct; (b) the identified sequence of morphemes (word segmentation) is correct; (c) the vocalization (short vowels and diacritics) is correct; and (d) the English gloss is accurate. **MSA words which are not in BAMA automatically receive the tag NO_FUNC.**

3.1.4 Human Intervention is Necessary in POS

Human annotators perform the necessary task of disambiguating many orthographically identical forms. For example, active verbs may have the same input string as passive verbal forms, and prepositions cliticized with nouns (*bi*-noun) may have the same input string as pure nouns or verbs (noun or verb starting with *b*), as in 3 below. The morphological analyzer will give all of the possibilities allowed by the orthography (nine potential solutions, only two of which are shown below).

```
INPUT STRING: باسم

SOLUTION 1: bAsim

LEMMA ID: bAsim_1

POS: bAsim/NOUN_PROP

GLOSS: Basem/Basim

SOLUTION 2: biAismi

LEMMA ID: {ism_1

POS: bi/PREP+{isom/NOUN+i/CASE_DEF_GEN

GLOSS: by/with + name + [def.gen.]
```

In this example, the correct choice is the proper name “bAsim” (SOLUTION 1) but the choice of “bi-ism” with the prepositional clitic “bi” is also available (SOLUTION 2), since that is one of the possible analyses of the text string. The POS annotator must choose the correct analysis. An incorrect choice will lead to drastically different syntactic tree structures (a noun phrase vs. a prepositional phrase), and one of them is clearly (to the human reader or annotator) incorrect.

At the moment, these are distinctions that are hard for automatic tools to make, so human annotation is necessary.

3.1.5 POS Annotator Decision Process

The guidelines for the POS annotators are relatively straightforward, since the task essentially involves choosing the correct analysis from the list of alternatives provided by the morphological analyzer and adding the correct case ending. The difficulties encountered by annotators in assigning POS and case endings are somewhat discussed above.

Annotators use the following criteria for making the following five POS decisions:

1. Correctness/acceptability is a decision concerning each of the following ordered set, provided by the Morphological Analyzer's output: (a) POS tag, (b) morphological segmentation, (c) vocalization including case and mood endings, and (d) English gloss.
2. If all four criteria are met in one of the displayed solutions, the annotator chooses that solution and is automatically moved on to the next item in the displayed text.
3. If the first three criteria are acceptable but the English gloss is defective, the annotator may still choose the appropriate solution, but should also choose the "Gloss Problem" option and type an explanation in the Comment field (e.g., "Gloss Problem: should be...").
4. If any of the first three criteria is unacceptable (i.e., wrong POS tag, or wrong segmentation, or wrong vocalization), the annotator must choose the "No match" option, and then enter the appropriate explanation in the Comment field.
5. The automatic solution is preferred over the "no match" solution as the latter leads to a proliferation of NO_FUNC tokens in the corpus. As a result, exceptions to point 4 are outlined below:
 - 5.1 If the word is a proper name → Choose the "X-Solution" option, which displays one or more morphological analyses based on prefix/suffix analyses only.
 - 5.2 If the passive form of a verb is not in the output → Choose the active form solution then hit PASSIVE FORM.
 - 5.3 If the word is a comparative adjective → Choose the ADJ solution then hit ADJ_COMP.
 - 5.4 If a word is a typo (NW) or a transcription error (BN), hit the TYPO or TRANSERR button. For transcription errors, type in the correct word along with its morphological analysis and gloss in the comment field.
 - 5.5 If there is a grammatical problem related to one word (E.g. a word is in nominative instead of accusative) → Select the solution that matches the writer/speaker's intended word then hit GRAMMAR PROBLEM.

5.6 If a word is transcribed without case ending (E.g. an indefinite accusative that is not graphemically represented) → **DO NOT** hit GRAMMAR PROBLEM. Whether in newswire or BN, writer and frequently omit case endings.

5.7 If a token is a partial segment of a word → Hit PARTIAL and select the solution that will come in the solution field. This situation is relatively frequent in speech-based BN corpora.

5.8 If unsure whether the token is a transcription error or grammar problem from the part of the speaker → Check the AUDIO

5.9 If a selected token has two distinct meanings one of which is absent from BAMA, do not select the available analysis and hit GLOSS PROBLEM. For example, if the POS analysis of the token >ahlAF as NOUN with the meaning of “welcome” is absent from BAMA, do not select the analysis with the gloss “family” by default. Instead, hit the no-match button and add the analysis and gloss that match the meaning of the word in that context.

3.1.6 Issues related to the annotation of dialect

1. If a word is dialectal and has no equivalent in MSA (E.g. *مُوش*/muw\$) → Hit DIALECT.
2. If a word is dialectal in pronunciation only and has an equivalent in MSA → Choose the correct solution from the output then hit DIALECT.
3. If a word is dialectal and has an equivalent morphological form in MSA but not the same function and/or meaning, the word should be annotated as DIALECT. In the example, *ما في شي* (mA fiy \$iy), fiy should not be annotated as fiy/PREP → Hit DIALECT

3.1.7 Reporting on tool issues

Annotators are required to report on the situations such as a mismatch between the output and the guidelines. For example, a misspelled tag PSEUDO-VERBS instead of PSEUDO_VERB. Corruption of files should also be reported.

3.2 The Buckwalter Transliteration Table

ب b	ا A	ى I	إ <	ؤ &	أ >	ا A	ء e
ذ *	د d	خ x	ح H	ج j	ث v	ت t	ة p
ظ z	ط T	ض D	ص s	ش S	س s	ز z	ر r
م m	ل l	ك k	ق q	ف f	- _	غ g	ع E
ك K	ن N	ف F	ي Y	ى Y	و w	ه h	ن n
	أ {	ا }	و o	و ~	ي i	أ u	ا a
			ر R	گ G	ف v	ج J	پ P

4.0 TAGS, TOKENIZATION, AND MORPHOLOGICAL AND POS ANNOTATION

4.1 POS Tags

Section 4.1.1 provides an alphabetical list of the parts of speech with their corresponding tags whereas **Section 4.1.2** matches tags with their related parts of speech.

4.1.1 List of Parts of Speech with their Tags

In this section, the name of the part of speech is set first and the tag used for it is provided second.

Abbreviation - ABBREV
Accusative - ACC
Adjective - ADJ
Adverb - ADV
Case - CASE
Command Verb - CV
Cardinal Number - NOUN_NUM
Comparative Adjective - ADJ_COMP
Connective Particle - CONNEC_PART
Conjunction - CONJ
Definite - DEF
Demonstrative Pronoun - DEM_PRON
Determiner – DET
Dialect - DIALECT
Direct Object - DO
Dual-D
Dual - DU
Emphatic Particle - EMPHATIC_PART
Feminine - F
Feminine - FEM
Focus Particle - FOCUS_PART
Foreign Word – FOREIGN
Foreign Script -FOREIGN_SCRIPT
Future - FUT
Future Particle - FUT_PART
Genitive - GEN
Imperfect Verb - IV
Indefinite - INDEF
Indicative - I
Interjection – INTERJ

Interrogative Adverb- INTERROG_ADV
Interrogative Particle - INTERROG_PART
Interrogative Pronoun – INTERROG_PRON
Jussive Particle – JUS_PART
Masculine - M
Masculine-MASC
Mood – MOOD
Negative Particle – NEG_PART
Noun - NOUN
Noun Suffix- NSUFF
Nominative - NOM
Noun Quantifier - NOUN_QUANT
Ordinal Number - ADJ_NUM
Passive - PASS
Particle – PART
Partial Word-PARTIAL
Perfect Verb - PV
Plural - PL
Possessive - POSS
Preposition - PREP
Pronoun- PRON
Proper Noun - NOUN_PROP
Pseudo Verb - PSEUDO_VERB
Punctuation - PUNC
Relative Adverb - REL_ADV
Relative Pronoun - REL_PRON
Response Conditional Particle - RC_PART
Restrictive Particle- RESTRIC_PART
Singular - S
Singular - SG
Subject -SUBJ

Subjunctive - S
Subordinating Conjunction - SUB_CONJ
Suffix - SUFF
Transcription error - TRANSERR
Typo - TYPO
Verb - VERB
Verb Particle -VERB_PART
Verbal Noun – NOUN.VN
Verbal Adjective- ADJ.VN
Vocative Particle -VOC_PART

4.1.2 List of Tags with their Corresponding Parts of Speech

This section puts the tag as given and provides the part of speech that it refers to.

ABBREV-Abbreviation
ACC- Accusative
ADJ-Adjective
ADJ_COMP- Comparative Adjective
ADJ_NUM- Ordinal number
ADJ.VN - Verbal Adjective
ADV- Adverb
CASE – Case
CONJ- Conjunction
CONNEC_PART- Connective Particle
CV - Command Verb
DEF - Definite
DEM_PRON-Demonstrative Pronoun
DET – Determiner
DIALECT - Dialect
DO - Direct Object

D – Dual
DU – Dual
EMPHATIC_PART - Emphatic Particle
F-Feminine
FEM- Feminine
FOCUS_PART - Focus Particle
FOREIGN-Foreign Word
FOREIGN_SCRIPT-Foreign Script
FUT-Future
FUT_PART -Future Particle
GEN-Genitive
I- Indicative
INDEF - Indefinite
INTERJ –Interjection
INTERROG_ADV – Interrogative Adverbs
INTERROG_PART - Interrogative Particle
INTERROG_PRON – Interrogative Pronoun
IV-Imperfect Verb
JUS_PART – Jussive Particle
M-Masculine
MASC-Masculine
MOOD – Mood
NSUFF- Noun Suffix
NOM- Nominative
NOUN - Noun
NOUN_NUM- Cardinal Numbers
NOUN_PROP- Proper Nouns
NOUN_QUANT- Noun Quantifier
NOUN.VN – Verbal Noun
NEG_PART - Negative Particle
PART –Particle

PARTIAL-Partial Word
PASS –Passive
PL-Plural
POSS- Possessive
PREP- Preposition
PRON -Pronoun
PSEUDO_VERB – Pseudo-Verb
PUNC - Punctuation
PV-Perfect Verb
RC_PART - Response Conditional Particle
RESTRIC_PART – Restrictive Particle
REL_ADV - Relative Adverb
REL_PRON - Relative Pronoun
S -Subjunctive
SUBJ –Subject
S- Singular
SG –Singular
SUB_CONJ - Subordinating Conjunction
SUFF- Suffix
TRANSERR-Transcription error
TYPO - Typo
VERB-Verb
VERB_PART - Verb Particle
VOC_PART -Vocative Particle

4.2 Segmentation

In Arabic, some function words that constitute separate words tend to be merged in writing. The reverse phenomenon also occurs where morphologically complex words have fossilized; and therefore, have to be interpreted and tagged as one single word.

Complexity is further added when the two variants, the non-segmented and the segmented, coexist in the language. In this case, annotators should use context as a disambiguator to decide on the appropriate annotation of the token in that particular context and in conformity with the present guidelines.

The two sections below present exhaustive lists of tokens that should be segmented and tokens that have a **segmented** and a **non-segmented** version. We assume that words outside the two lists constitute one single indivisible unit or token.

4.2.1 List of **Segmented** Tokens

In this section, the most common and problematic tokens that have to be cliticized are listed. The part of speech should be cliticized and each part annotated separately based on the POS cliticization procedure below.

Arabic	Buckwalter	POS Segmentation
لَئِنْ	la}n~a	la- (EMPHATIC_PART) + <in~a (PSEUDO_VERB)
لِإِنْ	li>an~a	li- (PREP) + >an~a (SUB_CONJ)
لَئِنْ	la}in	la- (EMPHATIC_PART) + <in (SUB_CONJ)
لَأُبَدُّ	lAbud~a	lA (PSEUDO_VERB) + bud~ (NOUN) + a (CASE_DEF_ACC)
لِمَنْ	liman	li- (PREP) + man (INTERROG_PRON)
لِمَنْ	liman	li- (PREP) + man (REL_PRON)
لِإِنَّا	li>al~A	li- (PREP) + >an (SUB_CONJ) + lA (NEG_PART)
لِكَيْ	likay	li- (PREP)+ kay (SUB_CONJ)
هَلَّا	hallA	hal (INTERROG_PART) +lA (NEG_PART)
مِمَّنْ	mim~an	min (PREP) + man (INTERROG_PRON)
مِمَّنْ	mim~an	min (PREP) + man (REL_PRON)
مَالَمْ	mAlam	mA (SUB_CONJ) + lam (NEG_PART)
قَلَّمَا	qal~amA	qal~/PV+-a PVSUFF_SUBJ :3MS + mA (SUB_CONJ)
قَبْلَمَا	qablamA	qabol(NOUN)+a (CASE_DEF_ACC) + mA (SUB_CONJ)
مِثْلَمَا	mivlamA	mivol (NOUN)+a (CASE_DEF_ACC) + mA (SUB_CONJ)

مِثْلَمَا	mivlamA	mivol (NOUN)+a (CASE_DEF_ACC) + mA (REL PRON)
حَسْبَمَا	HasbamA	Hasob (NOUN) +a (CASE_DEF_ACC) + mA (REL PRON)
وَقْتَمَا	waqtamA	waqot (NOUN) + a (CASE_DEF_ACC)+mA (SUB CONJ)
حَيْنَمَا	HynamA	Hyn (NOUN) + a (CASE_DEF_ACC)+ mA (SUB CONJ)
بِمَا	bimA	bi- (PREP) + mA (REL PRON)
بِمَا	bimA	bi- (PREP) + mA (INTERROG PRON)
بِمَا	bimA	bi- (PREP) + mA (SUB CONJ)
بِمْ	bima	bi- (PREP) + ma (INTERROG PRON)
بِمْ	bima	bi- (PREP) + ma (REL PRON)
بِمْ	bima	bi- (PREP) + ma (SUB CONJ)
لِكَيْمَا	likaymA	li- (PREP) + kaymA (SUB CONJ)
عَمَّا	Eam~A	Ean (PREP) + mA (INTERROG PRON)
عَمَّا	Eam~A	Ean (PREP) + mA (REL PRON)
عَمَّا	Eam~A	Ean (PREP) + mA (SUB CONJ)
عَمَّ	Eam~a	Ean (PREP) + ma (INTERROG PRON)
عَمَّ	Eam~a	Ean (PREP) + ma (REL PRON)
عَمَّ	Eam~a	Ean (PREP) + ma (SUB CONJ)
مِمَّا	mim~A	min (PREP) + mA (INTERROG PRON)
مِمَّا	mim~A	min (PREP) + mA (REL PRON)
مِمَّا	mim~A	min (PREP) + mA (SUB CONJ)
مِمَّ	mim~a	min (PREP) + ma (INTERROG PRON)
مِمَّ	mim~a	min (PREP) + ma (REL PRON)
مِمَّ	mim~a	min (PREP) + ma (SUB CONJ)
عَلَام	EalAma	EalY (PREP) + ma (INTERROG PRON)
عَلَام	EalAma	EalY (PREP) + ma (REL PRON)
عِنْدَمَا	EindamA	Einod (NOUN) +a (CASE_DEF_ACC) + mA (SUB CONJ)
عِنْدَمَا	EindamA	Einod (NOUN) +a (CASE_DEF_ACC) + mA (REL PRON)

بَعْدَمَا	baEdamA	baEod (NOUN) + a (CASE_DEF_ACC) + mA (SUB_CONJ)
بَعْدَمَا	baEdamA	baEod (NOUN) + a (CASE_DEF_ACC)+mA (REL PRON)
كَذَلِكَ	ka*alika	ka- (PREP) + *alika (DEM PRON_MS)
لِذَلِكَ	li*alika	li- (PREP) + *alika (DEM PRON_MS)
لِذَا	li*A	li- (PREP) + *A (DEM PRON_MS)
لَطَالَمَا	laTAlamA	la- (EMPHATIC_PART) + TAlamA (ADV)
أَنْتَ لَا	>anlA	>an (SUB_CONJ) + lA (NEG_PART)
أَلَا	>al~A	>an (SUB_CONJ) + lA (NEG_PART)
لَمَّا	lamA	la- (RC_PART) + mA (NEG_PART)
كَأَنَّ	ka>ano	ka/PREP+>ano/SUB_CONJ

4.2.2 List of Tokens with Segmented and Non-Segmented Versions

Below is a listing of words the **segmentation** of which has to be interpreted and decided upon by the annotator based on context. In other words, the annotator has to figure out whether these entries have been merged for pure stylistic purposes. As such, those entries carry two distinct meanings in the sentence. On the other hand, the **segmented** entries express one single meaning. In other words, the morphemes making up the entry do not contribute to separate meanings in the sentence. The table represents both the **non-segmented** and the **segmented analyses** of the entries. Annotators should refer to the examples given for the **non-segmented** versions in their corresponding sections.

Arabic	Buckwalter	Non-Segmented Analysis	Syntactic/Semantic Context	Example	Segmentation	Syntactic/Semantic Context	Example
إِذَا	<in~amA	RESTRIC_PART	It is used after a negative construction to mark a restriction.	يَعْتَقِدُونَ أَنْ ذَلِكَ إِذَا يَتِمُّ عَلَى حِسَابِ الاهْتِمَامِ بِالْحَدَاثَةِ (yaEtaqidwna >an~ *lika <in~amA yatim~u EaY HisAbi AlAhtimAmi biAlHadAvapi	<in~a (PSEUDO_VERB) + mA (REL_PRON)	It typically introduces a verbal never an equational sentence. The relative pronoun introduces a free relative clause that has a nominal function as either a subject or a predicate.	

				/they believe that this indeed is accomplished at the cost of modernism)			
أَمَّا	>an~amA	SUB_CONJ	It can be replaced by the subordinating conjunction >an~a		>an~a (SUB_CONJ) + mA (REL_PRON)	The relative pronoun stands for and refers to a noun in the sentence.	
فِيمَا	fiymA	SUB_CONJ	It introduces a dependent clause the event of which is simultaneous with another event in the main clause.	قَامَ الرَّجُلُ مُتَوَعِّدًا فِيمَا اسْتَمَرَ هَذَا التَّعْبِيرُ (qAma Alrajulu mutawaEidAF fiymA Astamar~a h*A AltaEbyru/ the man stood up threatening while the former continued, saying)	fiy- (PREP) + mA (REL_PRON)	The relative pronoun stands for or refers to a noun in the sentence	وَفِيمَا يَتَعَلَّقُ بِعَمَلِيَّاتِ السَّلَامِ فَقَدْ اَلْكَدَّ بَرْنَامَجُ الحِزْبِ (wa fiymA yataEal~aqu biEamaliy~api AlsAlAmi faqad >ak~ada barnAmaju AlHizbi.../ and in what is connected with the peace process, the party's programme has emphasized...)
فِيمَا	fiymA	SUB_CONJ	It introduces a clause the event of which is simultaneous with another event in the main clause.		fiy- (PREP) + mA (INTERROG_PRON)	In direct questions	
فِيمَا	fiymA	SUB_CONJ	It introduces a dependent clause the event of which is simultaneous with another event in the main clause.		fiy- (PREP) + mA (SUB_CONJ)		
أَلَا	>aIA	PART	It is used in declarative structures		>a (INTERROG_PART) + IA (NEG_PART)	In direct questions	أَلَا يَكْفِي بِأَنَّكَ عَامِلٌ؟ (>aIA yakfy bi>an~aka EATiIN ?/ is it not enough that you are out of work?)

أَيُّهَا	>ay~uhA	VOC_PART	It is used in vocative constructions		>ay~ (NOUN_QUANT) + -hA (POSS_PRON_3F S)	The meaning of >ay~ is « which » and it has the pronoun as a complement that can be replaced by a noun in genitive.
أَيُّهَا	>ay~uhA	VOC_PART	It is used in vocative constructions		>ay~ (INTERROG_PRON) + -hA (POSS_PRON_3F S)	The meaning of >ay~ is « any » and it has the pronoun as a complement that can be replaced by a noun in genitive.
أَيُّهَا	>ay~uhA	VOC_PART	It is used in vocative constructions	أَيُّهَا الْفَارِيُّ الْكَرِيمُ (>ay~uhA AlqAri}u Alkarymu/Oh dear writer	>ay~ (REL_PRON) + -hA (POSS_PRON_3F S)	The meaning of >ay~ap is « which » and it has the pronoun as a complement that can be replaced by a noun in genitive.
أَيُّهَا	>ay~atuhA	VOC_PART	It is used in vocative constructions		>ay~ (NOUN_QUANT) + +at (NSUFF_FEM_S G)+u/CASE_DEF_NOM + -hA (POSS_PRON_3F S)	The meaning of >ay~ap is « which » and it has the pronoun as a complement that can be replaced by a noun in genitive.
أَيُّهَا	>ay~atuhA	VOC_PART	It is used in vocative constructions		>ay~ (INTERROG_PRON) + +at (NSUFF_FEM_S G)+u/CASE_DEF_NOM + -hA (POSS_PRON_3F S)	The meaning of >ay~ap is « any » and it has the pronoun as a complement that can be replaced by a noun in genitive.
أَيُّهَا	>ay~atuhA	VOC_PART	It is used in vocative constructions		>ay~ (REL_PRON) +at (NSUFF_FEM_S G)+u/CASE_DEF_NOM+ -hA (POSS_PRON_3F S)	The meaning of >ay~ap is « which » and it has the pronoun as a complement that can be replaced by a noun in genitive.
كما	kamA	CONJ	It coordinates two constituents. It has the same function as “and”		ka- (PREP) + mA (REL_PRON)	Comparison where the meaning of the preposition is “like”. If mA has a nominal reference in the sentence, it is annotated as a relative pronoun.
كما	kamA	CONJ	It coordinates two constituents. It has	يَقُومُ الْبَنْكُ يَفْتَحُ الْحِسَابَاتِ	ka- (PREP) + mA (SUB_CONJ)	

			the same function as “and”	الشخصية... كما يعمل البنك على تمويل المشروعات المختلفة (yaqwmu Albanku bifatHi AlHisAbAti Al\$axSiy~api ...kamA yaEmilu Albanku EaLY tamwylI Alma\$rwEAti Almuxtalifapi/ the bank undertakes the opening of personal accounts... and the bank will endeavour to finance various projects)			
دُونَمَا	duwnamA	NOUN	It has the meaning of “without”. تَصَدَّرَ لِقَاتِهِ ال-62 دُونَمَا تَهْدِيدٍ جَدِيٍّ Al-62 duwnamA tahdiyK jidiy~K He was first in his 62 turnings without any serious threat.	تَصَدَّرَ لِقَاتِهِ ال-62 دُونَمَا تَهْدِيدٍ جَدِيٍّ (taSad~ara laf~Atihi Al-62 duwnamA tahdiyK jidiy~K He was first in his 62 turnings without any serious threat)	duwn (NOUN) + a (CASE_DEF_ACC) + mA (REL_PRON) (When duwna + mA comes as duwnamA, it should be split) duwn (NOUN) + i (CASE_DEF_GEN) + mA (SUB_CONJ)	It has the meaning of « without + what” It has the meaning of ‘without + that’	دُونَ مَا ذَكَرَهُ أَحْمَدُ (duwna mA vakarahu >aHmadN/ without what >aHmad mentioned) كَيْفَ صُرِفَ مِنْ دُونِ مَا نَعْلَمُ (kayofa Surifa mino duwni mA naEolama /how was it spent without that we know)
لِمَا	limA	INTERROG_PRON	It has the meaning of “why”	لِمَا كُلُّ هَذَا الْجَدَلِ؟	li- (PREP) + mA +(REL_PRON)	It has the meaning of for + what	عَبَّرَ عَنْ تَقْدِيرِهِ الْمَعْنَوِيَّ وَالسِّيَاسِيَّ لِمَا تُقَدِّمُهُ إِيرَانَ لِحَرَكَةِ

				(limA kul~u ha*A Aljadali?/ <u>why</u> is this all this discussion?)			المقاومة في لبنان (Eab~ara Ean taqdiyrihi AlmaEnawiy wa AlsiyAsiy~ <u>limA</u> tuqadimuhu <yraAn liHarakapi AlmuqAwamap fiy lubnAn/He expresses his moral and political respect for what Iran gives to the resistance in Lebanon)
كذا	ka*A	NOUN	It used to denote an indefinite number of something. ka*A is called in Arabic إسم كناية (<isom kinAyap)	وَصَحَّحَهُ كَذَا مَرَّةً (wa waD~aHahu ka*A mar~apK/ and he clarified it so/such occasion/and he clarified it so many times.	ka/PREP+*A/DE M_PRON_MS	It is used in comparisons and analogies	كذا كذا (*A ka*A/this like this)

4.3 Morphological Analysis and POS Annotation

Parts of speech and their morphological analysis are described in this section.

4.3.1 Parts of Speech

The parts of speech that are covered are nouns, pronouns, adjectives, verbs, adverbs, prepositions, interjections, particles, conjunctions, and subordinating conjunctions.

4.3.1.1 Nouns

Nouns are divided into two main categories: Common nouns (إسم الجنس) and proper nouns (إسم علم).

4.3.1.1.1 Common Noun (إسم جامد): NOUN

Common nouns refer to entities and concepts that have a more general reference than proper nouns. Common nouns are hosts of prefixes and suffixes of person, gender, definiteness, number, and case. They are given the tag NOUN in POS annotation.

Common nouns include derived and non-derived nouns. Non-derived nouns include words such as أُم (>um~/mother), يَوْم (yawm/day), رَأْس (ra>s/head), ماء (mA'/water), رَجُل (rajul/man). Derived nouns can be abstraction nouns such as كَيْفِيَّة (kayfiy~ap/manner) and نَظْرِيَّة (naZariy~ap/theory), or diminutives (تَصْغِير) like شَجِيرَة (\$ujayrap/'small' tree) and دُوَيْلَة (duwaylap/small country), or abstract concepts such as فَوْز (fawz/victory), دَعْم (daEm/support).

Titles are also annotated as common nouns and should, therefore; be tagged as NOUN. Example, جنرال (jinirAl/General), عَمِيد (Eamiyd/Lieutenant), الأَخ (Al>ax/Brother), الرئيس (Alra}iys/The president).

Arabic also includes a set of borrowed nouns such as مُوسِيقَى (muwsiyqY/music), بَنْك (bank/bank), and كَامِيرَا (kAmiyrA/camera). These should be annotated as nouns.

Words such as رَأْسَمَال (ra>small/capital), لَامَرْكَزِيَّة (lAmarkaziy~ap/decentralization), and يَأْنَصِيْب (yAnaSiyb/lottery) are nouns and should be treated as one unit even when they occur as two separate tokens in the corpus. Whenever that occurs, each token should be tagged as noun. Such compound nouns should not be confused with other nouns that occur in the form of two nouns such as جَوَازُ سَفَر (jawAzu safar/passport) and غُرْفَة نَوْم (gurfap nawm/bedroom). These should be annotated as separate nouns.

Nouns expressing possession and exception which were formerly treated as pronouns and particles are nouns; and therefore, should also be tagged as NOUN. Please refer to Sections 5.1.3 and 5.1.4 for a list of former pronouns and particles.

4.3.1.1.2 Proper Noun (إِسْم عَلَم): NOUN_PROP

Proper nouns are nouns that have a unique referential meaning in context that is mutually exclusive with other entities. In the example, نَشَرَ زَيْدُ الْكِتَابَ (na\$ara zaydN AlkitAba/zayd published the book) “zayd” is a proper noun because it has a single reference in the context of writing even though other people may have the same name.

Proper nouns refer to names of people, geographical entities, and acronyms. Personal names may be one word such as أَحْمَد (>aHmad/Ahmed) and مَرْيَم (maryam/Mary); or two or more words as in أَبُو أَحْمَد (>abuw >aHmad/Abu Ahmed), ابْنُ خَالِدُونَ (<ibn xalduwn/Ibn Khalduwn), and عَلَاءُ الدِّين (Eala' Aldiyn/Ala Eddine). When such nouns are made up of two or more words, both words should have the tag NOUN_PROP. Geographical entities are cities, countries, and geographical features such as mountains and deserts are annotated as proper nouns. Acronyms such as أَوْبِك (>wbk/OPEC) and يُونِسْكُو (yuwniskuw/UNESCO) are proper nouns and should have the tag NOUN_PROP.

Months should be annotated as proper nouns even if they are made up of more than one token as in تشرين الثاني (ti\$ryn AlvAniy/November) where each token is annotated as a proper noun.

Titles of newspapers, magazines, and news agencies, sport teams are annotated as proper nouns. Names of parties of the type of حزب الله (Hizb All~ah) and البعث (AlbaEv) should be annotated as proper nouns. Parties with names such as الإتحاد الديمقراطي (Al<it~iHAD AldiyuwqrATiy/the Democratic Union) should be annotated as common nouns.

Proper nouns can be confused with common nouns. A case in point is the word جنوب إفريقيا (januwb <ifriyyA/ South Africa) the two parts of which are tagged as NOUN_PROP when it refers to the country South Africa. Each word; however, must be annotated as common nouns when the token refers to a geographical part; that is to say, to the southern part of Africa. In this case, context should be used as a disambiguator on the basis of which the correct interpretation should be selected.

The policy of the guidelines with regard to case marking on proper nouns is to annotate them with no case ending. Proper nouns are also assumed to be inherently definite; and therefore, are not marked for definiteness.

Exceptions to the two points above do exist; however, though these are very rare. The exception on case assignment for proper nouns occurs in some specific writing context. In the sentence, قابلت علياً (qAbaltu Ealiy~AF/I met Ali) عَلِيًّا (Ealy~AF/Ali) gets the accusative case as it is the object in the sentence. The exception to the inherent definiteness of proper nouns occurs in examples such as هُنَاكَ خَمْسُونَ مُحَمَّدًا فِي هَذَا الْإِجْتِمَاعِ (hunAka xamsuwna muHAM~adAF fiy ha*A Al<ijtimAEi/There is 50 Mohamed in this meeting) where مُحَمَّدًا (Mohamed) is tagged as NOUN_PROP with an indefinite case. Another exception is the annotation of the proper noun الله (All~ah/ God) which is annotated in the same line as common nouns in terms of case ending rules.

4.3.1.1.3 Noun Quantifiers: NOUN_QUANT

The noun quantifiers listed below express either quantity or approximation.

Arabic	Buckwalter	Gloss	Example	Transliteration	Gloss	Translation
--------	------------	-------	---------	-----------------	-------	-------------

أي	>ay~	Any	ليس لهما أي استخدامات عسكرية	laysa lahumA >ay~u <istixdAmAtK Easkariy~apK	Not for them <u>any</u> usages military	They don't have any military usages
أية	>ay~ap	Any [fem]	و في أي من هذه الخلافات لم يحثكم أي من لحود و الحريري إلى مجلس الوزراء	wa fiy >ay~K min ha*ihi AlxilAfAt lam yaHtakim >ay~N min laHud wa AlHariyriy <iLY majlisi AlwuzarA'	And in <u>any</u> of these disputes did not appeal any of Lahood and al- Hariri to the cabinet	And in any of these disputes neither Lahood nor al-Hariri appealed to the cabinet
كل	kul~	All, the entire				
كلا	kilA (Masculine Dual)	Both				
كلتا	kiltA (Feminine Dual)	Both				
جميع	jamiyE	All, the entire	جميع الناس أتى الجميع	jamiyEu AlnAsi >atY AljamiyE	<u>All</u> the people Came <u>all</u>	All the people All came
بعض	baED	Some				

بضع	biDE	Some, a few, little,				
بضعة	biDEp	Some				
جَلّ	jul~	Most, the majority	جَلّ الكتب	jul~u Alkutubi	<u>Most</u> the books	Most of the books
كافة ¹	kAf~ap	All				
معظم	muEDam	Most				
غالبية	gAlibiy~ap	Most of				
أغلب ²	>aglab	Most, the majority of				
أكثر ³	>akvar	Most, the majority of				
نصف	niSf	Half	نصف ساعة	niSofu sAEapK	<u>Half</u> hour	Half an hour
شطر	\$aTr	Half				
ضعف	DiEf	Double				
ربع	rubuE	Quarter	ربع المدّة	rubuEu Almud~api	<u>Quarter</u> the duration	The quarter of the duration
ثلث	vuluv	Third				
خمس	xumus	Fifth	خُمْسُ الأَرْضِي	xumusA Al>arADiy	<u>Fifth</u> the lands	The fifth of the earth
سدس	sudus	Sixth				
سبع	subE	Seventh				
ثمان	vumn	Eighth				
تسع	tusuE	Ninth				
عشر	Eu\$ur	Tenth				
حوالي	HawAlay	About, around	حوالي سبعين جنديًا	HawAlay sabEyna jundy~A	<u>Around</u> seventy soldiers	Around seventy soldiers
زهاء	zuhA'	About, approximately, around	زهاء ثلاثة وعشرين طفلًا	zuhA'a valAvapK wa Ei\$ryna TiflAF	<u>Around</u> twenty-three child	Around twenty three children
قُرابة	qurAbap	Approximately, almost, about, around	قُرابة مائة ألف دولار	qurAbapa mA}api >alfi duwlArK	<u>Around</u> a hundred thousand dollar	Around one hundred thousand dollars

¹ Compare with kAf~ap as the active participle of the verb kaf~a meaning “cease to”. This active participle is annotated as ADJ in POS.

² It also has the POS value of a comparative adjective in specific contexts. Please refer to section 5.2.2 for important discriminatory tests.

³ It also has the POS value of a comparative adjective in specific contexts. Please refer to section 5.2.2 for important discriminatory tests.

قاب	qAb	(short) distance, (small) space, very close, very near				
قيس	qays	Around	قيس ثلاثة أمتار	qaysa valAvapi >amtArK	Around three meter	Around three meters
قيد	qayd	Around				
كامل	kAmil ⁴	(the) entirety				

4.3.1.1.4 Cardinal numbers: NOUN_NUM

Cardinal numbers are tagged as NOUN_NUM. Unlike ordinal numbers which are adjectives and tagged as ADJ_NUM, cardinal numbers quantify rather than rank. They answer the question “How many?” Cardinal numbers from 11 to 19 are considered in POS as one word though there is a space between the first and the second token.

Compound numerals (from 21 to infinite) that are coordinated with the conjunction wa- follow the POS tag of the first element. If the first element is a cardinal number, the second element should be annotated as such.

Cardinal numbers can occur in pre-nominal and post-nominal positions. Pre-nominal cardinal numbers from 11 to 99 assign the accusative to the noun they are in complementation with. The remaining numeral numbers assign the genitive to their complement.

There is a total agreement in gender with the number 1, 2 and 10. Reverse gender agreement occurs from number 3 to 9. That is to say, the number gets feminine when followed by masculine and vice versa.

10's (except 10), hundreds, thousands ... etc are invariable; that is to say, they do not carry any inflections in order to agree with the noun they occur with. In post-nominal positions, they are tagged as NOUN_NUM when they are preceded by a noun in plural.

- Post-nominal cardinal numbers occur in a construction called apposition (بتل) when they are preceded by a plural noun. There is reverse gender agreement with that noun.

Please refer to Section 8.1.1 for the interpretation of digits as well as for tests distinguishing cardinal numbers from ordinal numbers.

Arabic	Buckwalter	Gloss	Example	Transliteration	Gloss	Translation
عشرة	Ea\$rap	Ten	عشر كتب	Ea\$ru kutubK	Ten books	Ten books

⁴ kAmil has another POS value. It is an adjective (ADJ) with the gloss ‘perfect’/‘complete’.

عشرة	Ea\$rap	Ten	عشرة ساعات	Ea\$rapu sAEAtK	Ten (FEM) hours (FEM)	Ten hours
خمسة	xamsap	Five	خمسة كتب	xamsapu kutubK	Five (FEM) books (MASC)	Five books
خمسة	xamsap	Five	خمسة ساعات	xamsu sAEAtK	Five (MASC) hours (FEM)	Five hours
مائة	mA}ap	One hundred	مائة كتاب	mA}apu kitAbK	One hundred book	One hundred books
مائة و خمسون	mA{ap wa xamsuwna	One hundred and fifty	السنوات المائة والخمسون	AlsanawAtu Alma{apu wa Alxamsuwna	The year the hundredth and the fiftieth.	The hundred and fiftieth year
ستة وثلاثون	sit~ap wa valAvwn	Thirty six	ستة وثلاثون طالبا	sit~apN wa valAvwna TalibAF	Thirty six student	Thirty six students
أربعة	>arbaEap	Four	غرف أربع	gurafN >arbaEN	Rooms four	Four rooms
ستة وخمسون	sit~ap wa xamsuwn	Fifty six	ستة وخمسون عاما	sit~apN wa xamsuwna EamAF	Fifty-six year	Fifty-six years

4.3.1.1.5 Annotation of Nouns

An illustration of the way noun inflections should be decided on and represented during annotation is provided in this section.

4.3.1.1.5.1 Gender Inflections

Gender inflections in Arabic are restricted to masculine and feminine markings though some nouns can fit either category. Annotators should mark gender based on the following criteria:

- Masculine is tagged as MASC for gender inflections on nouns and adjectives. **M** is used as part of the tag for masculine. The tag MASC is only applicable for the plural of masculine nouns. The masculine feature that pertains to an inherent masculine property in the noun as in رَجُل (rajul/man) and أَسَد (>asad/lion) or to a figurative meaning where words such as بَدْر (badr/moon) or لَيْل (layl/night) are masculine are not represented in POS annotation.
- Feminine is tagged FEM for gender inflections on nouns and adjectives. **F** is used as part of the tag for feminine pronouns. Feminine refers either to an inherent feminine property in the noun as in إِمْرَأَة (<imra>ap/woman) and نَاقَة (nAqap/she-camel) or to a figurative meaning where nouns such as شَمْس (\$ams/Sun) and عَيْن (Eayn/Eye) are feminine. The latter meaning is not marked during POS annotation. Unlike masculine, particular inflections on the noun mark the feminine. These are the letter “p” as in سَلَة (sal~ap/basket), the letter “Y” as in مَأْوَى (ma>wY/shelter), and the letter “ ‘ ” as in صَحْرَاء (SaHrA’/desert). Other

nouns are feminine though they do not carry any of the previously mentioned feminine inflections as in يَد (yad/Hand). These are not considered in POS annotation for gender inflections.

4.3.1.1.5.2 Number and Definiteness

Properties of number are indicated for the singular (SG or S for pronouns), dual (DU or D as part of the tag for pronouns), and plural (PL and P as part of the tag for pronouns) only when the plural is marked through inflection not through pattern derivation as in الأَيَّامُ (Al>ay~Amu/the days). Definiteness in Arabic is marked morphologically using the determiner Al-

4.3.1.1.5.3 Case Marking

Nouns generally can inflect for three cases: The nominative tagged as NOM, the accusative tagged as ACC, and the genitive tagged as GEN. Common nouns along with proper nouns can take those cases depending on their function and position in the sentence. Case is annotated as definite or indefinite.

4.3.1.1.5.4 Illustration of the POS Annotation for Nouns

Arabic	Buckwalter	Gloss	POS Annotation
اليوم	Al+yawm+u	The + day	DET+NOUN+NSUFF_MASC_SG+CASE_DEF_NOM
يومان	yawm+Ani	Days two	NOUN+NSUFF_MASC_DU+CASE_INDEF_NOM
اليومين	Al+yawm+ayni	The + days two	DET+NOUN+NSUFF_MASC_DU+CASE_DEF_ACC OR DET+NOUN+NSUFF_MASC_DU+CASE_DEF_GEN
الجمعة	Al+jumuE+ap+a	Friday	DET+NOUN+NSUFF_FEM_SG+CASE_DEF_ACC
خالد	xAlid	xAlid	NOUN_PROP
القاهرة	Al+qAhirap+u	Cairo	DET+NOUN_PROP+CASE_DEF_NOM
الأمس	Al>ams+i	Yesterday	DET+NOUN + CASE_DEF_GEN

4.3.1.2 Pronouns

There are five categories of pronouns: personal, demonstrative, interrogative, relative, and exclamative pronouns.

4.3.1.2.1 Personal Pronouns

Personal pronouns are split into two major categories: Unbound and bound pronouns. Unbound pronouns are free morphemes that occur as separate words. Bound pronouns are morphemes that cannot occur independently of another morpheme. They are related to other words called their hosts.

4.3.1.2.1.1 Unbound Pronouns: PRON

Arabic	Buckwalter	POS Tag	Gloss	Examples	Transliteration	Gloss	Translation
أنا	>anA	PRON_1S	I	أنا مَعَكُمْ	>anA maEakum	I am with you	I am with you
نحن	naHnu	PRON_1P	We	نَحْنُ مُنْتَظِرَاتَانِ	nahnu muntaZiratAni	We are waiting	We are waiting
أنت	>anta	PRON_2MS	You	أنتَ مَنْ يُحْرِجُنِي	>anta man yuHrijuniy	You are who embarrasses me	You are the one who is embarrassing me
أنتِ	>anti	PRON_2FS	You	أنتِ جَمِيلَةٌ	>anti jamiylapN	You are beautiful	You are beautiful
أنتم	>antumA	PRON_2D	You	أنتم في المَدْرَسَةِ	>antumA fiy Almadrasapi	You are in the school	You are in the school
أنتم	>antum	PRON_2MP	You	أنتم جالسونَ	>antum jAlisuwna	You are sitting	You are sitting
أنتنَ	>antun~a	PRON_2FP	You	أنتنَ جالساتُ	>antun~a jAlisAtuN	You are sitting	You are sitting
هو	huwa	PRON_3MS	He	هُوَ صَدِيقِي	huwa Sadiyqiy	He is my friend	He is a friend of mine
هي	hiya	PRON_3FS	She	هِيَ صَدِيقَتِي	hiya Sadiyqatiy	She is my friend	She is a friend of mine
هما	humA	PRON_3D	They	هُمَا صَدِيقَانِ	humA SadiyqAni	They are friends	They are friends
هم	hum	PRON_3MP	They	هُم يُعَالُونَ مِنَ الْفَقْرِ	hum yuEAnuwna mina Alfaqri	They are suffering from poverty	They are suffering from poverty
هنَ	hun~a	PRON_3FP	They	هُنَّ مُتَابِرَاتُ	hun~a muvAbirAtN	They are perseverant	They are perseverant

4.3.1.2.1.2 Bound Pronouns

4.3.1.2.1.2.1 Object Pronouns: PRON

The entries in the table below function as objects or complements of prepositions. Pronouns that are objects of verbs are annotated as suffixes to the verbs they attach to. For example, the token **وَضَعَهُ** (waDaEahu/put) is annotated as PV+a/PVSUFF_SUBJ: 3MS+hu/**PV SUFF DO:3MS**.

Arabic	Buckwalter	POS Tag	Gloss	Examples	Transliteration	Gloss	Translation
ني-	-niy	PRON_1S	Me	شاهدتني	\$Ahadatniy	Saw (she) <u>me</u>	She saw me
ي	-y	PRON_1S	Me	لي	liy	For <u>me</u>	For me
ي-	-ya	PRON_1S	Me	لي	liya	For <u>me</u>	For me
نا -	-nA	PRON_1P	Us	شاهدتنا	\$AhadatnA	Saw (she) <u>us</u>	She saw us
ك -	-ka	PRON_2MS	You	شاهدتك	\$Ahadatka	Saw (she) <u>you</u>	She saw you
ك -	-ki	PRON_2FS	You	شاهدتك	\$Ahadatki	Saw (she) <u>you</u>	She saw you
كما -	-kumA	PRON_2D	You	شاهدتكما	\$AhadatkumA	Saw (she) <u>you</u> (two)	She saw you
كم -	-kum	PRON_2MP	You	شاهدتكم	\$Ahadatkum	Saw (she) <u>you</u>	She saw you
كن -	-kun~a	PRON_2FP	You	شاهدتكن	\$Ahadutkun~a	Saw (she) <u>you</u>	She saw you
هُ-	-hu	PRON_3MS	Him	شاهدته	\$Ahadathu	Saw (she) <u>him</u>	She saw him
ها-	-hA	PRON_3FS	Her	شاهدتها	\$AhadathA	Saw (she) <u>her</u>	She saw her
هما-	-humA	PRON_3D	Them	شاهدتهما	\$AhadathumA	Saw (she) <u>them</u> (two)	She saw them
هم -	-hum ⁵	PRON_3MP	Them	شاهدتهم	\$Ahadathum	Saw (she) <u>them</u>	She saw them
هن -	-hun~a	PRON_3FP	Them	شاهدتهن	\$Ahadathun~a	Saw (she) <u>them</u>	She saw them

4.3.1.2.1.2.2 Possessive Pronouns: POSS_PRON

⁵ A phonological variant is –him as in من بعدهم (min baEdihim/from after them)

Arabic	Buckwalter	POS Tag	Gloss	Examples	Transliteration	Gloss	Translation
ي -	-niy	POSS_PRON_1S	My/me	عَسَائِي أَوْفَقٌ فِي أَعْمَالِي	EasAniy >uwaf~aqu fiy >aEmAliy	May I succeed in my endeavors	May I succeed in my endeavors
ي	-y	POSS_PRON_1S	My	كِتَابِي	kitAbiy	Book <u>my</u>	My book
ي -	-ya	POSS_PRON_1S	My	كِتَابِي	kitAbiya	Book <u>my</u>	My book
نَا -	-nA	POSS_PRON_1P	Our	كِتَابُنَا	kitAbunA	Book <u>our</u>	Our book
كَ -	-ka	POSS_PRON_2MS	Your	كِتَابُكَ	kitAbuka	Book <u>your</u>	Your book
كَ -	-ki	POSS_PRON_2FS	Your	كِتَابُكَ	kitAbuki	Book <u>your</u>	Your book
كُمَا -	-kumA	POSS_PRON_2D	Your	كِتَابُكُمَا	kitAbukumA	Book <u>your</u>	Your book
كُمْ -	-kum	POSS_PRON_2MP	Your	كِتَابُكُمْ	kitAbukum	Book <u>your</u>	Your book
كُنَّ -	-kun~a	POSS_PRON_2FP	Your	كِتَابُكُنَّ	kitAbukun~a	Book <u>your</u>	Your book
هُ -	-hu	POSS_PRON_3MS	His	كِتَابُهُ	kitAbuhu	Book <u>his</u>	His book
هَا -	-hA	POSS_PRON_3FS	Her	كِتَابُهَا	kitAbuhA	Book <u>her</u>	Her book
هُمَا -	-humA	POSS_PRON_3D	Their	كِتَابُهُمَا	kitAbuhumA	Book <u>their</u>	Their book
هُمْ -	-hum	POSS_PRON_3MP	Their	كِتَابُهُمْ	kitAbuhum	Book <u>their</u>	Their book
هُنَّ -	-hun~a	POSS_PRON_3FP	Their	كِتَابُهُنَّ	kitAbuhun~a	Book <u>their</u>	Their book

4.3.1.2.2 Demonstrative Pronouns: DEM_PRON

Demonstrative pronouns are pronouns used for proximal or distal reference. They are invariable. The subdivision below is purely semantic.

4.3.1.2.2.1 Demonstratives Indicating Proximity

Arabic	Buckwalter	POS Tag	Gloss	Examples	Transliteration	Gloss	Translation
هَذَا	h`*A	DEM_PRON_MS	This	هَذَا الرَّجُلُ	ha*A Alrajulu	<u>This</u> the man	This man
هَذِي	h`*iy	DEM_PRON_FS	This				
هَاتِيهِ	hAtihi	DEM_PRON_FS	This				

هَات	hAti	DEM_PRON_FS	This	هَاتِ الْفَتَاةُ	hAti AlfatApu	<u>This</u> the girl	This girl
هَا	hA	DEM_PRON	This	هَا هُوَ الرَّجُلُ	hA huwa Alrajulu		Here comes the man
هَذِهِ	h`*ihi	DEM_PRON_F	This	هَذِهِ الْفَتَاةُ هَذِهِ الْبُيُوتُ	ha*ihi AlfatApu ha*ihi Albuyuwtu	<u>This</u> the girl <u>These</u> the houses	This girl These houses
هَذَانِ	h`*Ani	DEM_PRON_MD	These	هَذَانِ الرَّجُلَانِ	ha*Ani AlrajulAni	<u>These</u> the two men	Those (two) men
هَذَيْنِ	h`*ayn						
ذَانِ	*aAn	DEM_PRON_MD	These				
ذَيْنِ	*ayn						
هَاتَانِ هَاتَيْنِ	hAtAni hAtayn	DEM_PRON_FD	These	هَاتَانِ الْفَتَاتَانِ	hatAni AlfatAtAni	<u>These</u> the two girls	Those (two) girls
تَانِ	tAni	DEM_PRON_FD	These				
تَيْنِ	tayna						
هَؤُلَاءِ / أُولَئِكَ	h'a&uLA'i/>uwIA'	DEM_PRON_P	These	هَؤُلَاءِ الرَّجَالِ هَؤُلَاءِ الْفَتَاتِ	h'&uLA'I AlrijAl h'&uLA'I AlfatayAt	<u>These</u> the men <u>These</u> the girls	These men These girls

4.3.1.2.2 Demonstratives Indicating Distance

Arabic	Buckwalter	POS Tag	Gloss	Examples	Transliteration	Gloss	Translation
ذَا	*aA	DEM_PRON_MS	That				
ذَلِكَ	*alika	DEM_PRON_MS	That	ذَلِكَ الرَّجُلُ	*alika Alrajulu	<u>That</u> the man	That man
ذَٰلِكَ	*Aka	DEM_PRON_MS	That				

ذَلِكَمَا	*alikumA	DEM_PRON_D	Those				
ذَلِكُمْ	*alikum	DEM_PRON_MP	Those				
ذَلِكَنَّ	*alikuln~a	DEM_PRON_FP	Those				
تَاكَ	tilka	DEM_PRON_F	That Those	تَاكَ الْفَتَاةُ تَاكَ الْبُيُوتُ	tilka AlfatApu tilka Albuyuwtu	That the girl Those the houses	That girl Those houses
تَاكَ	tAka	DEM_PRON_FS	That				
تِيكَ	tiyka						
تَلِكُمْ	tilkum	DEM_PRON_FP	Those				
ذَانِكَ	*Ank	DEM_PRON_MD	Those	ذَانِكَ الرَّجُلَانُ	*Anaka AlrajulAni	Those men	Those (two) men
ذَيْنَاكَ	*aynaka						
تَانَاكَ	tAnika	DEM_PRON_FD	Those	تَانَاكَ الْفَتَاتَانُ	tAnika_AlfatAtAni	Those girls	Those (two) girls
تَيْنَاكَ	taynaka						
أُولَئِكَ	>uwla}ika	DEM_PRON_P	Those	أُولَئِكَ الرَّجَالُ أُولَئِكَ الْفَتَاتُ	>uwla}ka AlrijAl >uwla}ka AlfatyAt	Those men Those the girls	Those men Those girls
أُولَئِكُمْ	>uwla}ikum	DEM_PRON_MP	Those				

4.3.1.2.3 Interrogative Pronouns: INTERROG_PRON

Entries in the table below are tagged as INTERROG_PRON when used to form direct questions. These entries are relative pronouns and should be tagged as REL_PRON when used to introduce indirect speech clauses.

Arabic	Buckwalter	Gloss	Syntactic/Semantic Context	Examples	Transliteration	Gloss	Translation
كم	kam	How much/many	In direct questions	كم قرشًا دَفَعْتَ في هَذَا	kam qir\$FA dafaEta fiy ha*A?	<u>How many</u> piaster paid you in this	How many piastres did you pay for this?
أيّ	>ay~	What Which	In direct questions	من أيّ بلد أنتَ	min >ay~ baladK >anta?	From <u>which</u> country you	From which country are you?
أيّة	>ay~ap						
ماذا	mA*A	What	In direct questions	ماذا تَقُولُ الوثائقُ	mA*A taquwlu AlwawA}iqu ?	<u>What</u> say the documents	What do the documents say ?
ما	mA	What	In direct questions	ما الذي أوقَعَهَا في ذلكَ المَصِيرِ المُخِيفِ	mA Ala*iy >awqaEahA fiy *alika AlmaSiy Almuxiyf ?	<u>What</u> which plunged her in that the fate the scary	What was it that plunged her into that fearful fate?
مَنْ	man	Who Whom	In direct questions	مَنْ هُوَ الأفضَلُ بَيْنَهُمْ	man huwa Al>afDal baynahum?	<u>Who</u> he the best among them	Who is the best among them?

4.3.1.2.4 Interrogative Adverbs: INTERROG_ADV

Entries in the table below are tagged as INTERROG_ADV when used to form direct questions. These entries are relative adverbs and should be tagged as REL_ADV when used to introduce indirect speech clauses.

Arabic	Buckwalter	Gloss	Syntactic/Semantic Context	Examples	Transliteration	Gloss	Translation
متى	mataY	When	In direct questions	متى هذا الحفل	matY hA*A AlHafI?	<u>When</u> this the party	When is the party?
كيف	kayfa	How	In direct questions	كيف هي استعداداتكم	kayfa hiya <istiEdAdAtukum?	<u>How</u> it is your preparation	How are your plannings going?
لماذا	awn*A	Why	In direct questions	لماذا أكون مختلفًا	lima*A >akuwnnu muxtaliFA?	<u>Why</u> be (I) different	Why should I be different?
لم	lima	Why	In direct questions	لم أكون مختلفًا	lima >akuwnnu muxtaliFA?	<u>Why</u> be (I) different	Why should I be different?
لما	limA	Why	In direct questions				
أين	>ayna	Where	In direct questions	أين أنا الآن	>ayna >anA Al na?	<u>Where</u> I am now	Where am I now?
أنى	>an~Y	How Where	In direct questions				
أين	>ay~Ana	Where	In direct questions				

4.3.1.2.5 Relative Pronouns: REL_PRON

Arabic	Buckwalter	Gloss	Syntactic context	Examples	Transliteration	Gloss	Translation
الذي	Ala*y	Who/whose/	It introduces a	الرجل الذي أفعاله تسبق	Alrajulu Al*iy	The man <u>who</u> his	the man whose actions

		which (masculine)	relative clause modifying a noun	أقواله	>afEAluhu tasbiq >aqwAlahu	actions precede his words	precede his words
التي	Alaty	Who/whose/which (feminine)	It introduces a relative clause modifying a noun	المقابلة التي حضرها	AlmuqAbalapu Alatiy HaDaraha	The interview <u>which</u> attend (he) it	The interview which he attended.
الذاتين	Alla*Ani Alla*ayni (accusative or genitive case)	Who/whose/which (Dual Masculine)	It introduces a relative clause modifying a noun	لم تنس أبداً حالة الفلق و عذم الثقة الذين رافقا خطواتها الأدبية الأولى	Iam tansa >abadFA Halapa Alqalaq wa Eadam Alviqap Ala*yn rAfaqA xuTuwAtihA Al>adabiy~ap Al>uwIY	Did not forget never state the anxiety and lack confidence <u>which</u> accompanied her steps the literay the first	She never forgot the state of anxiety and lack of confidence which accompanied her first literary steps.
اللتان اللتين	AllatAni Allatayni (accusative or genitive case)	Who/whose/which (Dual Feminine)	It introduces a relative clause modifying a noun	في لندن و باريس اللتين زرتُهُما	fiy lunduN wa bAriys Alatayn zurtuhumA	In London and Paris <u>which</u> visited (I) them	In London and Paris which I visited
الذين	Ala*yna	Who/whose/which (Plural Masculine)	It introduces a relative clause modifying a noun	الناس الذين لا يكثرثون للإشارة السير	Aln~As Al*iyin lA yaktarivuwna li<i\$ArApi Alsayr	The people <u>who</u> not care for the traffic lights the movement (walking/driving)	The people who do not care for the pedestrian lights
اللاتي اللواتي	AllAty AllawAty	Who/whose/which (Plural Feminine)	It introduces a relative clause modifying a noun	واحدة من أعظم النساء اللاتي عرفهن	wAHidap min >aEZam AlnisA' AllAtiy Earafahun~a	One from the greatest women <u>who</u> knew (I)	One of the strongest and greatest women whom we knew.
اللائي	AllA}iy	Who	It introduces a relative clause modifying a noun				
ال ⁶	Al-	Which	It introduces a relative clause	الاجتماع الهادف إلى إيجاد حلول	Al<ijtimAE AlhAdif <iLY	The meeting <u>which</u> aims at finding	The meeting that aims at finding solutions

⁶ Refer to section 5.2.4 for important annotation issues.

			modifying a noun		<iyjAdi HuluwlK	solutions	
مَنْ - مَنْ	man -m~an (as in Eam~an)	Who	It introduces indirect questions	أَنْتَ مَنْ يُحْزِنِي	>anta man yuHrijuniy	You are <u>who</u> embarrasses me	You are the one who embarrasses me.
مَا - مَّ	mA -m~a (as in Eam~a)	What	It introduces indirect questions	وَقَعَ مَا نَخْشَاهُ	waqaEa mA nax\$Ahu	Happened <u>what</u> we fear	What we fear happened
أَيَّ	>ay~	Who	It introduces indirect questions				
أَيَّا	>ay~AF	Whatever	It introduces a clause modifying a noun or a verb. It is only used in indefinite accusative where it can be replaced mahmA	أَنَّ الْأَرْمَاتِ الدُّوَلِيَّةِ وَ النِّزَاعَاتِ أَيَّا تَكُنْ دَرَجَةُ تَعْقِيدِهَا يُمَكِّنُ أَنْ	>an~a Al>azamAt Alduwaliy~ap wa AlnizAEAt >ay~FA takun darajapu taEqiydihA yumkinu >an	That the crises the national and the conflicts whatever is the degree of its complexity	That the national crises and the conflicts whatever is the degree of its complexity can...
مَهْمَا	mahmA	Whatever	It introduces a clause modifying a noun or a verb	مَهْمَا بَلَغَ حَدَسَ الْإِنْسَانِ وَ قُوَّتِهِ عَلَى التَّنَبُّؤِ تَبْقَى لِحِظَاتٍ فِي دَائِرَةِ الْغَيْبِ	mahmA balaga Hads Al<insAni wa quw~atihi EalY Altanab~u&i tabqY laHaZAtN fiy dA }rapi	<u>Whatever</u> reached intuition the human being and strength his on the prediction remain moments in circle the unknown	Whatever man's intelligence and power of prediction may achieve, some moments will remain in the sphere of the unknown.

					Algaybi		
كم	kam	How much/many	It introduces indirect questions	إِسْطَظَعْتُ أَنْ أَعْرِفَ كَمْ مِنَ الْمَالِ سَأَدْفَعُ	<istaTaEtu >an >aEriFa kam min AlmAli sa>adfaE	Could (I) to know <u>how much</u> of + money will pay (I)	I was able to know how much money I will pay
ماذا	mA*A	What	It introduces indirect questions	لَمْ أَعْرِفْ مَاذَا أَفْعَلُ	lam >Earif mA*A >afEal	Did not know (I) <u>what</u> do (I)	I didn't know what to do

4.3.1.2.6 Exclamative Pronoun: EXCLAM_PRON

Arabic	Buckwalter	Gloss	Syntactic context	Examples	Transliteration	Gloss	Translation
ما	mA	What	Exclamative structure	ما أَحْسَنَ السَّمَاءِ !	mA >aHsana AlsamA'a!	<u>What</u> a beautiful the sky!	What a beautiful sky!
ما [التَّعْجِيبِيَّة]	[mA of Exclamation]	How					

4.3.1.3 Adjectives

4.3.1.3.1 Adjective: ADJ

Qualitative adjectives, adjectives of intensity, and relational adjectives are tagged as ADJ in POS. However, they are still analyzed as adjectives (See Syntactic Annotation Section 3.1.3.1.1 for more details and examples).

4.3.1.3.1.1 Qualitative Adjectives (صِفَةٌ مُشَبَّهَةٌ): ADJ

Qualitative adjectives have the meaning of a stable characteristic and are only derived from intransitive verbs. They have the following patterns:

Pattern	Verb- Adjective Derivation	Example	Transliteration	Gloss	Translation
فَعْلَاءُ - أَفْعَلُ >afEalu - faEIA'u	حمراء - أحمرُ حَمْرٌ - Hamura- >aHmaru - HamrA'	أَحْمَرُ اللّوْنِ	>aHmaru Allawni	<u>Red</u> the color	Red
فاعلة- فاعل fAEilN - fAEilapN	طاهرةٌ - طاهرٌ طَهْرٌ - Tahura- TahirN - TahirapN	ماءٌ طاهرٌ	maA'N TahirN	Water <u>pure</u>	Pure water
فَعْلَةٌ - فَعِلُ faEilN - faEilapN	- فَرِحَ - فَرِحٌ فَرَحٌ fariHa-fariHN- fariHapN	مُحَمَّدٌ فَرِحَ	muHamad fariHN	Muhammad <u>happy</u>	Muhammad is happy
فَعْيِلَةٌ - فَعْيِلُ faEylN - faEylapN	- جَمِيلَةٌ - جَمِيلٌ جَمَلٌ jamula- jamyln - jamylapN	جَمِيلُ الوَجْهِ	jamiylu Alwajhi	<u>Beautiful</u> the face	Of a beautiful face
فعلان - فعلى faEIAnu - faEIY	- جَوَّعَانَةٌ - جَوَّعَانٌ جَاعٌ - جَوْعَى jAEa- jawEAnu - jawEAnapN or jawEY	أَتَى الطِّفْلُ جَوَّعَانٌ	>atY AlTiflu jawEAnN	Came the child <u>hungry</u>	The hungry child came

فَعَالَةٌ - فَعَالٌ fuEAIN - fuEAlapN	شَجَاعَةٌ - شَجَاعٌ شَجَعٌ \$aj~aEa- \$ujAEapN	هَذَا الرَّجُلُ شَجَاعٌ	ha*A Alrajulu \$ujAE	This the man <u>brave</u>	This man is brave
فَعَالٌ faEAIN (used for the mas and fem)	صَنَعَ - صَنَاعٌ SanaEa- SanAEN	إِنَّهُ تَلْمِيزٌ صَنَاعٌ	<in~ahu tilmiy*N SanAE	Indeed he is pupil <u>creative</u>	He is indeed a creative pupil
فَعْلَةٌ - فَعْلٌ faEIN - faElap	سَهْلٌ - سَهْلٌ sahlN - sahlapN	إِمْتِحَانٌ سَهْلٌ	<imtiHAn sahl	Exam <u>easy</u>	An easy exam
فَعِيلٌ fayEalN	سَادٌ - سَيِّدٌ sAda- say~idN	هَذَا الرَّجُلُ سَيِّدٌ	ha*A Alrajulu say~dN	This the man <u>master</u>	This man is a master

4.3.1.3.1.2 Adjectives of Intensity (صيغة المبالغة) : ADJ

Adjectives of intensity are usually derived from the base trilateral verb. They have the following patterns:

Pattern	Verb- Adjective Derivation	Example	Transliteration	Gloss	Translation
faE~AIN	ظَلَمَ – ظَلَامٌ Zalama – Zal~Am	هَذَا الرَّجُلُ ظَلَامٌ	ha*A Alrajulu Zal~Am	This the man <u>unjust</u>	This man is unjust
faEylN	سَمِعَ – سَمِيعٌ samiEa- samyE	رَجُلٌ سَمِيعٌ	rajulN samiyEN	All <u>hearing</u> omniscient	All hearing, Omniscient (God)
mifEAIN	أَعْطَى – مِعْطَاءٌ >aETY- miETA'	رَجُلٌ مِعْطَاءٌ	rajulN miETA'	Man <u>generous</u>	A generous man
faEwlN	زَهَقَ – زَهَاقٌ zahaqa- zahwq	لَقَدْ أَصْبَحَ زَهَاقًا	laqad >aSbaHa zahuwqAF	has become (he) <u>short- tempered</u>	He has become short- tempered
faE~alap	عَلِمَ – عِلْمَةٌ Ealima- Eal~Amap	إِنَّهُ شَخْصٌ عِلْمَةٌ	<in~ahu \$axSN Eal~Amap	Indeed he is man <u>scholar</u>	He is indeed a scholar
faEilN	حَذَرَ – حَذْرٌ Ha*ara – Ha*ir	كُنْ حَذْرًا	kun Ha*irAF	Be (you) <u>cautious</u>	Be cautious

4.3.1.3.1.3 Relational Adjectives (إسم النسبَة): ADJ

Relational adjectives are derived from proper nouns, common nouns, or masdars by adding the suffix –iy~ which denotes possession.

Noun	ADJ Derivation	Example	Transliteration	Gloss	Translation
تونس tuwnis	تونسي tuwnisiy~	الرَّجُلُ التُّونِسِيُّ	Alrajulu Altuwnisiy~u	The man <u>the</u> <u>Tunisian</u>	The Tunisian man
تاريخ tAryx	تاريخي tArixiy~	الْفَتْرَةُ التَّارِيخِيَّةُ	Alfatrap AltArixiy~ap	The era <u>the</u> <u>historic</u>	The historic era
بَصْرَة basrap	بَصْرِي ⁷ basriy~	إِنَّهُ بَصْرِي حَتْمًا	<in~ahu baSriy~ HatmAF	Indeed he is <u>from</u> <u>Basra</u> definitely	He is indeed from Basra
يَهُودِيَّة yahwidiyya	يَهُودِي yahwidiy~	يَهُودِي النَّشَأُ	yahwidiy~u Alna\$>ap	<u>Jewish</u> the bringing up	From a jewish origin
صَحْرَاءُ shahraa	صَحْرَاوِي ⁸ shahrawiy~	مُنَاحُ صَحْرَاوِي	munAxN SaHrAwiy~N	Climate <u>of desert</u>	Desert climate

⁷ If the noun has the "t" as marker of the gender (F). The "p" is omitted in the <ism Alnisbap.

⁸ If the noun has the "w" as marker of the gender (F). The "w" become "w" in the <ism Alnisbap.

SaHrA'	SaHrAwiy~				
اقتصاد <iqtiSAd	اقتصادي <iqtiSAdiy~	الأجواء الاقتصادية	Al>ajwA'u Al<iqtiSAdiy~ap	The economic conditions	The economic conditions

4.3.1.3.2 Ordinal Numbers: ADJ_NUM

When numbers are interpreted as ordinals, they are adjectives and should be tagged as ADJ_NUM. Ordinal numbers, unlike cardinal numbers (See Section 4.3.1.1.5) order and rank rather than quantify. They can occur in pre-nominal and postnominal positions.

Please refer to section 4.6, for tests to distinguish cardinal numbers from ordinal numbers.

Arabic	Buckwalter	Gloss	Example	Transliteration	Gloss	Translation
ثالث	vAliv	Third	ثالث الكتب	vAlivu Al kutubi	<u>Third</u> book (s)	The third book
خامس	xAmis	Fifth	خامس الغرف	xAmisu Algurafi	<u>Fifth</u> room (s)	The fifth room
أول	>aw~al	First	أول يوم	>aw~alu yawmK	<u>First</u> day	The first day
سابع	sAbiE	Seventh	سابع باب	sAbiEu bAbK	<u>Seventh</u> door	The seventh door

4.3.1.3.3 Comparative Adjective: ADJ_COMP

The comparative adjective is a form derived from verbs according to their inflectional category (see Section 4.2.2). In Arabic, no morphological distinction is made between the comparative and the superlative meanings. The distinction is made based on context. The derivational patterns below provide a morphological basis for the assignment of the ADJ_COMP tag to a particular form.

The pattern faEIA' is also a comparative adjective pattern that is the feminine equivalent of أفعل (>afEal). For example, أصغر (>aSgar/small [masc]) → صغرى (SugrY/smallest [fem])

Three comparative adjectives do not have the pattern >afEal. These are xayor/best instead of >axoyaro and \$ar~/most malicious instead of >a\$ar. Both >aHab~/dearest and Hab~ can co-occur as adjective comparative forms (Qabbech, p.337).

Please refer to Section 5.2.2 for important information on the annotation of comparative adjectives أَكْثَر (>akvar) and أَغْلَب (>aglab).

NB: Please note that entries with the pattern أَفْعَل (>afEal) in exclamative constructions are annotated as verbs with the tag (VERB) and should not be confused with comparative adjectives.

(Refer to syntactic annotation sections 4.1.3 and 4.3.1.5 on exclamative constructions)

4.3.1.3.1 From the Base Triliteral Verbs (الفعل الثلاثي المجرد)

The comparative adjective has the pattern >afEal (أفعل).

Verb	Derived Comparative Adjective	Gloss	Example	Transliteration	Gloss	Translation
Hasuna	أَحْسَن >aHsan	Best Better	أَحْسَنُ الْأَعْمَالِ	>aHsanu >al>aEmAli	<u>Best</u> the deeds	The best deeds
kabura _	أَكْبَر >akba كُبْرَى kubrY	Old/Big Oldest/Biggest	أَكْبَرُ الْأَطْفَالِ كُبْرَى الْبَنَاتِ	>akbaru >al>aTfAli kubrY AlbanAt	<u>Biggest/oldest</u> the children Oldest the girls	The biggest/oldest child The oldest of the girls

حَرَصَ	أَحْرَصُ	More cautious	كُنْ أَحْرَصَ مِنْ ذَلِكَ	kun >aHraSa min *alika	Be (you) <u>more</u> <u>careful</u> than that	Be more careful than that
HariSa	>aHraS	Most cautious				
كَثَرَ	أَكْثَرَ	More	زَيْدٌ أَكْثَرُ مَالًا مِنْ عَمْرٍ	zaydN >akvaru mAlAF min EamrK	Zayd <u>more</u> money than Eamr	Zayd has more money than Amr
kavura	>akvar	Most				
قَلَّ	أَقَلُّ	Less	أَقَلُّ مَالًا	>akal~a mAlAF	<u>Less</u> money	Less money
qal~a	>aqal~	Least				

- Note that comparative adjectives cannot be derived from the derived trilateral verbs or the base or derived quadrilateral verbs. For these verb forms, the comparative or the superlative is indicated through the use of the comparative adjectives above followed by the masdar of the verb.

4.3.1.3.3.2 Important Remarks

Cases where some tokens have the pattern >afEal but are not necessarily comparative adjectives are frequent. For example, أَحْمَرُ (>aHmar/red) is an adjective not a comparative adjective. أَيْمَنُ (>ayman) is both a proper noun when it refers to the name of a person and an adjective when it means “right”. The adjective أَيْسَرُ (>aysar) is annotated in POS as both an adjective when it means “left” and a comparative adjective because it is derived from the noun يَسِيرٌ (yasyr) and when it means “easier/easiest”.

4.3.1.3.4 Illustration of the POS Annotation for Adjectives

Arabic	Buckwalter	Gloss	POS Annotation
التونسيُّ	Al+tuwnisiy~+u	<u>The Tunisian</u>	DET+ADJ+NSUFF_MASC_SG+CASE_DEF_NOM
الأولون	Al+>aw~al+uwna	<u>The first</u>	DET+ADJ_NUM+NSUFF_MASC_PL+CASE_DEF_ACC
أَحْرَصُ	>aHraS+u	<u>More cautious</u>	ADJ_COMP+CASE_INDEF_NOM
الأحسن	Al+>aHsan+i	<u>The best</u>	DET+ADJ_COMP+CASE_DEF_GEN
مُثَابِرَاتٌ	muvaBir+At+N	<u>Perseverant</u>	ADJ+NSUFF_FEM_PL+CASE_INDEF_NOM

شَدِيدُو التَّحْفَظِ	\$adiyd+uw AltaHaf~uZi	Of extreme carefulness	ADJ+NSUFF_MASC_PL_NOM_POSS
----------------------	---------------------------	---------------------------	----------------------------

4.3.1.4 Adjectives and Nouns: Principles for Decision Making [Under Experimentation and Review]

This section sets principles for the POS annotation of active and passive participles as well as manufactured gerunds. Refer also to section 7.1 for tests on adjectives and nouns. This section is under continuous experimentation and review.

4.3.1.4.1 Treatment of Active and Passive Participles

Names of activities, occupations, and professions that are active or passive participle at the level of morphology are considered as lexicalized nouns in POS; and therefore, annotated as nouns. Examples include: تاجر (tAjir/retailer), مندوب (manduwb/deputy), معلم (muEal~im/teacher), مُعْتَمِد (muEtamad/commissioner), مُوظَّف (muwaZ~af/employee), مُسْتَشَار (musta\$Ar/consultant), كاتب (kAtib/writer), خادم (xAdim/maid), سارق (sAriq/thief), عامل (EAmil/worker), سائح (sA}Ih/tourist), فنان (fan~An/artist), مُترجم (mutarjim/translator).

4.3.1.4.2 Treatment of the Manufactured Gerund (مصدر صناعي)

The manufactured gerund is a maSdar that is derived from common or derived nouns not verbs. It takes the iy~ap feminine nisbap suffix (Qabb~ish, 1986). It is the instigator of a derivation process which produces nouns and/or adjectives. Ryding (2005) refers to a category of “abstraction nouns” (p. 90) which have the feminine nisba ending –iy~p. Such nouns are derived from nouns, pronouns, participles, and borrowed words. For example, يَهُودِيَّة (yahuwdiy~ap/Judaism) is a manufactured gerund derived from the proper noun هُود (huwd/Hud). The tokens يَهُود (yahuwd/ Jews) and يَهُودِي (yahuwdiy~/ Jew) are derived from.

Ryding (2005, p. 91) calls such gerunds abstraction nouns that can be derived from nouns as in masiyHiy~ap, adjectives as in >afDaliy~ap, pronouns as in huwiy~ap, or participles as in mas}uwliy~ap. Schultz (2004, p.87) argues that the nisba suffix constitutes the basic structure for abstract nouns. Examples include تَعْدُدِيَّة (taEadudiy~ap/pluralism), نَظَرِيَّة (naZariy`ap/theory) أَهْمِيَّة (nujuwmiy~ap/fame), عَمَلِيَّة (Eamaliy~ap/operation), كَمِيَّة (kamiy~ap/quantity), كَيْفِيَّة (kayfiy~ap/manner), أَهْمِيَّة (>aham~iy~ap/importance), أَفْضَلِيَّة (>afDaliy~ap/precedence), فَعَالِيَّة (faE~Aliy~ap/effectiveness), أَقْلِيَّة (>aqaliy~ap/minority), مَسْؤُولِيَّة (mas&uwliy~ap/responsibility), دَيْبُلُومَاسِيَّة (dybluwmasiy~ap/diplomacy), دِيمُوقْرَاطِيَّة (diymuwqrATiy~ap/democracy)

Manufactured gerunds should be tagged as nouns as the feminine nisba suffix –iy~p does not reflect agreement with a noun. The token yahuwdiy~ap in the examples below does not agree with the noun it stand in predication with. Whether that noun is feminine or masculine, the token is invariable.

E.g. 1 اليَهُودِيَّةُ دِيَانَةٌ سَمَاوِيَّةٌ (Alyahuwdiy~apu diyAnapN samAwiyy~apN/Judaism is a divine religion)

E.g. 2 اليَهُودِيَّةُ مَذْهَبٌ دِينِيٌّ (Alyahuwdiy~apu ma*habN diydiy~N/Judaism is a religious affiliation)

4.3.1.5 Verbs

Verbs are divided into three categories: Inflectional verbs, pseudo-verbs, and non-inflectional verbs.

4.3.1.5.1 Inflectional Verbs

Inflectional verbs are verbs that carry inflections for gender, person, number, gender, tense, and/or mood. They are split into three categories based on tense. The perfect verb (فِعْلٌ مَاضِي) (PV) is the verb in the past tense, the imperfect verb (فِعْلٌ مُضَارِع) (IV) is the verb in the present tense, and the command verb (فِعْلٌ أَمْر) (CV) refers to the verb in the imperative.

Verb inflections are represented through the use of prefixes and suffixes (See table below for an illustration of the POS annotation for verbs). Mode is overtly represented when the verb is in the passive and is annotated as IV_PASS. If not mentioned, the active mode is assumed by default; and therefore, is not overtly annotated.

The three types of moods are: The Indicative (I) (المَرْفُوع) as in يَذْهَبُ (ya*habu/go) the Subjunctive (S) (الْمَتَّصُوب) as in لَنْ يَذْهَبَ (lan ya*haba/ will not go) , and the Jussive (J) (الْمَجْرُوم) as in لَمْ يَذْهَبْ (lam ya*habo/did not go) or اذْهَبْ (<i*habo/go in the imperative). The future and the jussive expressed via the morphemes س- (sa-) and ل- (li-) respectively are not cliticized at the TB level but are segmented during POS annotation. They are indicated by the tags FUT and JUS respectively (See table below).

Arabic	Transliteration	Translation	POS Annotation
يَذْهَبُ	ya*habu	He goes	ya/IV3MS+*hab/IV+u/IV_SUFFIX: I
لَنْ يَذْهَبَ	(lan) ya*haba	He will not go	ya/IV3MS+*ahab/IV+a/IV_SUFFIX: :S
لَمْ يَذْهَبْ	(lam) ya*hab	He did not go	ya/IV3MS+*ab/IV+o/IV_SUFFIX:J
ذَهَبَ	*ahaba	He went	*ahab/PV+a/PVSUFFIX_SUBJ: 3MS

قُبِلَ	qubila	It was accepted	qubil/PV_PASS+a/PVSUFF_SUB J: 3MS
تُقْبَلُ	tuqbalu	It is accepted	tu/IV3FS+qobal/IV_PASS+u/IVSUFF_MOOD:I

4.3.1.5.2 Pseudo-Verbs: PSEUDO_VERB (الحُرُوفُ الْمُشَبَّهَةُ بِالْأَفْعَالِ)

Pseudo-verbs are particles that have the same syntactic behavior as verbs in that they take a subject and a predicate. Unlike regular verbs, their subject is in the accusative and their predicate is in the nominative. Pseudo-verbs are <in~a and its sisters as well as EasY which is a pseudo verb in particular contexts (See Section 6.1.8 for important information on annotation decisions). Please refer to ATB Section 4.1.4 for examples on the syntactic structures where pseudo_verbs occur.

Arabic	Buckwalter	Gloss	Syntactic/Semantic Context	Examples	Transliteration	Gloss	Translation
إِنَّ	<in~a	(To be) (Indeed)	It reinforces the meaning of the sentence that follows it	إِنَّهُنَّ نَجْمَاتَانِ لَامِعَاتَانِ	<in~ahun~a najmatAni lAmiEatAni	Indeed they are two stars twinkling	Indeed they are two shining stars
لَيْتَ	layta	(If only) (Wish)	It expresses wish	لَيْتَ الْأَمْرَ تَوَقَّفَ هُنَا	layta Al>amra tawaq~afa hunA	If only the matter stopped here	If only the matter would have stopped here
لَعَلَّ	laEal~a	(May be)	It expresses doubt	لَعَلَّ أَحَمَّهُمَا خَاصِيَّتَانِ	lEal~a >aham~ahA	Perhaps most important of it	Perhaps the most important
عَلَّ	Eal~a	(Perhaps)		أَسَاسِيَّتَانِ	xAS~iyatAni >asAsiy~atAni	two featutes basic	of them are two basic features.
كَأَنَّ	ka>an~a	(As if)	It expresses similitude	كَأَنَّهُ شَاهِدُهُ وَكَانَ حَاضِرًا	ka>an~ahu \$Ahiduhu wa kAna HADirAF	As if he saw it and was (he) present	As if he had seen it and had been present
لَكِنَّ	l'kin~a	(However)	It introduces a contrast between two events	تَبَدُّوا هَذِهِ الْأَسْئَلَةَ بَسِيْطَةً جِدًّا وَ لَكِنَّهَا فِي الْوَاقِعِ شَدِيدَةُ التَّعْقُدِ	tabdwA ha*ihi Al>as;ilapu basiyTapF jid~FA wa l'kin~ahA fiy AlwAqiEi	Seems these questions simple very however in reality extreme	These questions look simple; in reality however, they

					\$adiydap AltaEaq~udi	the complication	are extremely complicated.
لا [النافية للجنس]	lA	(No)	It negates the statement	لا فائدة كبيرة	lA fA}idapa kabiyrap	No benefit great	There is no great use in it
عسى ⁹	EasY	May be/wish	It has a subject in the accusative				

4.3.1.5.3 Non-Inflectional Verbs: VERB (الأفعال الجامدة)

The non-inflectional verbs that do not inflect for any morphological feature such as gender, person, or number. They should be annotated as verbs in POS with the tag (VERB). The following are the most common entries in this category. Please note that verbs of surprise such as >aHsana, >ajmala are annotated as verbs with the tag VERB when they occur in exclamative constructions.

Arabic	Buckwalter	Gloss	Example	Transliteration	Gloss	Translation
حَيَّ	Hay~a	Come to				
ناهيك	nAhiyka	(How excellent)				
هَلُمَّ	hal~uma	Come				
هَيَّا	hay~A	Come on				
حَذَار	Ha*Ari (+min)	(Beware (of))				
أَمِين	miyn	Amen				
تَعَال	taEAl	Come				
هَبَّ ¹⁰	hab	Assume that				
هَيَّهَات	hayhAt	(How impossible)	هَيَّهَاتَ أَنْ يَدُومَ الْبَاطِلُ	hayhAta >an yaduwma AlbATilu		

⁹ EasY is by default tagged as VERB unless there is a graphemic indication of the accusative (in the case of the dual) that leads to assign EasY the tag PSEUDO_VERB.

¹⁰ Do not confuse with the imperative form of the verb (wahaba/donate).

شَتَان	\$at~Ana	(What a difference)		\$at~Ana AlEilmu wa Aljahlu		
سرعان	surEAna ¹¹	(Soon/how quickly)				
سرعان	sarEAna	(Soon/how quickly)				
حَبْدًا	Hab~a*A	(How nice)				
أَفْ	>uf~ (+ min)	(ugh)	أَفْ مِنْكَ	>uf~K minka	<u>Ugh</u> from you	
أَهْ	h ¹² (+min)	(Ah)	أَهْ مِنْكَ	hK minka	<u>Ah</u> from you	
بئس	bi}sa	(How bad)	بئسَ الرَّجُلُ	bi}sa Alrajulu	<u>How bad</u> is the man	
نِعَمَ	niEma	(How good)				
عَسَى	EasY	Wish/may be	عَسَى الْكَرْبُ الَّذِي أَمْسَيْتُ فِيهِ يَكُونُ وَرَاءَهُ فَرَجٌ قَرِيبٌ	EasY Alkarbu Al~a*y >amsaytu fiyhi yakuwnu warA'ahu farajN qariybN		
يَبْغِي	yanbagy	Must				
مَا [المُسْتَبْهَةِ بَلَيْسَ]	mA [the mA acting like laysa]	Not	مَا الْحُسْنُ فِي وَجْهِ الْفَتَى شَرَفًا لَهُ	mA AlHusnu fiy wajhi AlfatY \$arafAF lahu		
تُرَى	turY ¹³	Followed by a pronoun	؟! تَرَاهُ يَنْجَحُ	>a turAhu yanjaHu?		
خِلا	xalA	Except				
حاشا	HA\$A	Except				
عَدَا	EadA	Except				
هَا	hA	Take				

¹¹ surEAna and sarEAna are two phonological variants.

¹² Do not confuse with the Interjection >ah (أه) that does not inflect for case.

¹³ When turY comes by itself it is an interjection.

هات	hAt	Give				
-----	-----	------	--	--	--	--

4.3.1.6 Adverbs: ADV

4.3.1.6.1 Regular Adverbs: ADV

Adverbs are invariable words that give information about the time, location, manner, cause, purpose, or any other adverbial function of the verb or sentence they are modifying.

The list below is meant to be an exhaustive list of adverbs. Entries that are not in this list are not adverbs and should be annotated according to their morphological category. For example, “bayda” used to be annotated as an adverb in previous releases. The POS tag for this entry is now NOUN.

Arabic	Buckwalter	Gloss	Syntactic/Semantic Context	Examples	Transliteration	Gloss	Translation
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 Then	It provides the time of an event	كُنْتُ غَائِبًا عِنْدَذَاكَ	kuntu gA}bAF	Was (I) absent <u>then</u>	I was absent then
وقْتذاك	waqta*Aka	At that time Then	It provides the time of an event	كُنْتُ غَائِبًا وَقْتذاك	kuntu gA}bAF	Was (I) absent <u>then</u>	I was absent then
آنذاك	na*Aka	At that time Then	It provides the time of an event	كَانَ الْقَلَمُ آنذاك كَالْمُسْتَحِيلِ	kAna Alqalamu na*Aka kAlmustaHiyli	Was the pen <u>at that time</u> like the impossible	In those times (the use of) the pen was nearly impossible
آنئذٍ	na}i*K	At that instant	It provides the time of an event				

		Then					
حِينَئِذٍ	Hiyna}i*K	At that moment	It provides the time of an event	حِينَئِذٍ سَيَجْلَى الْحَقُّ	Hiyna}*K sayatajal~Y AlHaq~u	At that time right will be revealed	And then the truth will be revealed
		Then					
بَعْدَئِذٍ	baEda}i*K	After that time	It provides the time of an event				
عِنْدَئِذٍ	Einda}i*K	At that time	It provides the time of an event	عِنْدَئِذٍ سَيَجْلَى الْحَقُّ	Einda}*K sayatajal~Y AlHaq~u	At that time the truth will be revealed	And then the truth will be revealed
		Then					
وَقْتًا	waqta}i*K	At that time	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 at that time	He used to go regularly to that hospital then
		Then					
عَمَّا	Eam~a}i*K	At that moment	It provides the time of an event				
عَامَئِذٍ	EAmA*Aka	In that year	It provides the time of an event				
سَنَئِذٍ	sanata*Aka	In that year	It provides the time of an event				
يَوْمَئِذٍ	yawma}i*K	On that day	It provides the time of an event				
يَوْمَئِذٍ	yawma*Aka	On that day	It provides the time of an event				
مِنْئِذٍ	mu*~Aka	From then on	It provides the time of an event				

مُدْذَاكَ	mud*Aka	From then on	It provides the time of an event				
ثَمَّةٌ تَمَّ	vam~apa vam~a	There	It specifies the location of an event	ثَمَّةٌ وَقَتُّ فَاصِلٌ بَيْنَ الْوَصُولِ وَ الْإِقْلَاعِ الثَّانِي	vam~apa waqtN fASiIN bayna AlwSwli wa Al<iqlAEi AlvAniy	There 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	Here freedom and there slavery	Here 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	سَأَسَافِرُ هُنَاكَ	sa>usAfiru hunAka	Will travel there	I will travel there
هِنَاكَ	hunAlika	There					
أَيْضًا	>ayDAF	Also Too	It has the meaning of addition	زَارَهُ أَيْضًا	zArahu >ayDaA	Visited (he) him also	He visited him also
إِذَا إِذْنٌ	<i*AF <i*an	Hence So	It expresses the consequence of an event	إِذَا وَ بِكُلِّ الْوُضُوحِ وَ الصَّرَاحَةِ نَجَحَ الْإِرْهَابُ فِي قَتْلِ مَوْسِمِ سِيَّاحِي	<i*AF wa bikul~I AlwuDuwHi wa 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.
رُبَّمَا	rub~amA	Perhaps	It expresses doubt	رُبَّمَا يَبْدُو الْمَوْقِفُ 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 وَاقِفًا هَكَذَا	kAna wAQifAF haka*aA		

فَقَطْ	faqaT	Only That is all	It adds the nmeaning of restriction	فَقَطْ أَرَدْتُ أَنْ أُنَاقِشَ	faqaT >aradtu >an >unAqi\$a	Only wanted I to discuss	I only wanted to discuss
فَحَسَبُ	faHasbu	Only	It adds the meaning of restriction	عَلَيْهِ أَنْ يَفِي بِوَعْدِهِ فَحَسَبُ	Ealyhi >an yafiy biwaEdihi faHasbu	Has he to keep his promise <u>only</u>	He has to keep his promise that is all
قَطْ قَطُّ	qaT qaT~u	At all Ever	It reinforces the meaning of negation	مَا سَمِعْنَا بِمِثْلِهِ قَطُّ	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					
طَالَمَا ¹⁵	TAlamA	So often					
إِذَاكَ	<j*~Aka	Then					
سَاعَتَيْكَ	sAEata}i*K	That hour/time					
عَامَيْكَ	EAmA}i*K	That year					
قَبْلَيْكَ	qabola}i*K	Before then					
لِحُظَّتَيْكَ	laHoZata}i*K	That moment					
لَيْلَتَيْكَ	layolata}i*K	That night					

4.3.1.6.2 Relative Adverb: REL_ADV

Relative adverbs are invariable words that introduce adverbial clauses that give specific information about time, location, manner, or purpose. They have a specific adverbial function like time, location, manner, and purpose. The first seven entries in the table below should have the POS tag INTERROG_ADV (See Section 4.3.1.4) when used in direct questions. The remaining relative adverbs¹⁶ are always tagged as REL_ADV since they are not used in direct questions.

Arabic	Buckwalter	Gloss	Syntactic/Semantic Context	Examples	Transliteration	Gloss	Translation
مَتَى	mataY	When	It introduces indirect	لَمْ أَعْرِفْ مَتَى أَسْتَطِيعُ الذَّهَابَ	lam >aErif matY	Did not know (I)	I didn't know when

¹⁴ If it occurs as IA siy~amA, IA has to be annotated as negative particle.

¹⁵ Often occurs as laTAlamA. As such it should be cliticized into la- (EMPHATIC_PART)+-TAlamA (ADV)

¹⁶ The only exception is kam.

			questions		>astaTiyE Al*ahAb	when can (I) leaving	I can leave
كَيْفَ	kayfa	How	It introduces indirect questions	إِسْطَظَعْتُ أَنْ أَعْرِفَ كَيْفَ أَهْرُبُ	<istaTaEtu >an >Erifa kayfa >ahrubu	Could (I) to know how flee (I)	I knew how I will flee
لِمَاذَا	lima*A	Why	It introduces indirect questions	عَرَفْتُ لِمَاذَا أَكُنْتُ خَائِفًا	Earaftu ima*A kuntu xA}ifAF	Knew (I) why was (I) afraid	I knew why I was afraid
لِمَا/لِمَ	lima/limA	Why	It introduces indirect questions	عَرَفْتُ لِمَ كُنْتُ خَائِفًا	Earaftu lima kuntu xA}ifAF	Knew (I) why was (I) afraid	I knew why I was afraid
أَيْنَ	>ayna	Where	It introduces indirect questions	ذَهَبْتُ أَيْنَ كَانَتِ الْبِنْيَابَةُ	*ahabtu >ayna kAnat AlbinAyap	Went (I) where was the building	I went where the building was
أَتَى	>an~Y	Wherever	It introduces indirect questions	يُلاحِقُهَا العَارُ أَتَى ذَهَبَتْ	yulAHiquhA AlEAru >an~Y *ahabat	Follows her the disgrace wherever she went	Her disgrace follows her wherever she goes.
أَيَّانَ	>ay~Ana	Wherever	It introduces indirect questions				
حَيْثُ	Hayvu	Where	It introduces an adverbial clause modifying the verb in the main clause e	سَأَجِدُكَ حَيْثُ كُنْتَ	sa>ajiduka Hayvu kunta	Will find (I) you where were (you)	I will find you where you were
حَيْثُمَا	HayvumA	Wherever	It introduces an adverbial clause modifying the verb in the main clause	سَأَجِدُكَ حَيْثُمَا كُنْتَ	sa>ajiduka HayvumA kunta	Will find (I) you wherever were (you)	I will find you wherever you were
كَيْفَمَا	kayfamA	However	It introduces an adverbial clause modifying the verb in the main clause	هُوَ القَادِرُ عَلَى بَلْوَرَةِ هَذِهِ الدُّنْيَا كَيْفَمَا يَشَاءُ	huwa AlqAdiru EalY balwarapi ha*ihi AldunyA kayfamA ya\$A`u		
أَيْنَمَا	>aynamA	Wherever	It introduces an adverbial clause modifying the verb in the main clause	أَيْنَمَا ذَهَبَ وَ جَالَ فِي عَوَا صِمِ العالم . وَ أَنْخَاءِ الأَرْضِ يَحْمِلُ مَعَهُ شَيْئًا مِنَ الجَنُوبِ	>aynamA *ahaba wa jAla fiyand Eawa Simi AlEAlemi wa >anHaA`I Al>arDi yaHmilu maEahu \$ay}FA mina Aljanuwbi	Wherever goes he and wanders in the in capital cities the world and parts of the globe carry something from the South	Wherever goes (he) and wanders in capital cities the world and parts the globe carry with him something from the south

كُلَّمَا	kul~amA		It introduces an adverbial clause modifying the verb in the main clause	كُلَّمَا مَرَّتِ السَّنَوَاتُ فَقَدْنَا الْكَثِيرَ	kul~amA mar~at AlsanawAtu faqadnA Alkaviyra	Whenever passed the years missed (we) the much	Whenever years pass we lose a lot
كَمْ	kam	Whenever	It introduces exclamative constructions	كَمْ أَنْتَ جَمِيلَةٌ	kam >anti jamylapN!	How beautiful you are	How beautiful you are!
		How....!					

4.3.1.7 Prepositions: PREP

Prepositions are particles which assign case to their dependent constituent. The list of prepositions below is exhaustive and includes both the classical Arabic prepositions (حُرُوفَ جَرَ) that assign the genitive case to their dependent constituent as well as particles (Huruwf Eamilap) that are also considered prepositions in ATB but which assign accusative case to their dependent constituent unlike regular particles.

Classical prepositions in Arabic relate the complement they head to the verb. They assign the genitive case to their complement.

The other prepositions notably <i~A (حَرْفَ إِسْتِثْنَاءٍ) and wa- (وَآوِ الْمَعِيَّةِ) also assign case to the complement they head though accusative not genitive.

Entries that are not mentioned here should not be annotated as prepositions. A list of entries that were annotated as prepositions in previous corpora releases is provided in Section 5.1.1.

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 <iY AlmasA'i	From the morning <u>until</u> the evening	From morning until evening.
		At/To	It indicates the point of arrival	الوصولُ إِلَى إِتْفَاقِيَّةٍ	AlwSwlu <iY 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 <iY *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 عِلَاقَةِ الشَّرْطَةِ بِالنَّاسِ	ha*A Ean EalAqapi AlSurTapi biAlnAsi	This <u>about</u> relation the police with the people	This was about the relation between police and citizens
		From		يَجِبُ أَنْ يَمْتَنِعَ عَنْ التَّدْخِينِ	yajibu >an yamtaniEa Eani Altadxyni	Must to he depart <u>from</u> the smoking	He must refrain from smoking

في	fiy	In	It indicates location	في إستفتاء أجرتُهُ الجمعية الأمريكية	fiy <istiftA'K >ajrathu AljamEay~pu Al>amariykiy~apy	In poll carried out the university the American American	In a poll carried out by the American Society
		By	It is used in multiplication phrases	ثلاثة في ثلاثة تساوي تسعة	valAvapN fiy valAvapN tusAwiy tisEapN	Three multiplied <u>by</u> three equals nine	Three times three equals nine.
من	min	From	It indicates the point of origin be it physical or temporal	من لندن إلى دبي	min landun <iIY dubay	<u>From</u> London to Dubai	From London to Dubai
م-	mi- (as in mim~an)	Of	It selects a proportion or part from a whole.	ثمانون ملياراً من الدولارات	vamAnwna milyArAF mina AldwlArAt	Eighty billions from the dollars	Eighty billion of Dollars

لـ -	li-	Because of	It indicates causality	لِعَبِيْبِ فِيهَا	liEaybK fiyhA.	Because of a defect in it	because of some fault in herself.
		In order to	It introduces the purpose of an event	لِيَحْتِ الْمَوْقِفِ	libaHvi Almawqifi.	In order to examine the situation	In order to examine the situation.
		To/At	It describes the point of arrival	الْوَصُولُ لِإِتِّفَاقِيَّةِ	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 AlDiy AlEaraby	A book by our master MuHyi AlDiy AlEaraby

- بـ	bi-	In/At	It introduces information about location	بِمَدْرَسَةِ الصَّبَاحِيَّةِ الثَّانَوِيَّةِ	bimadraspi AlSabAHiy~ap AlvAnawiy~ap	In school al-Sabahiyya the secondary	At the Sabahia secondary school
		With	Instrument	بِيَدٍ مُرْتَعِشَةٍ	biyadK murtaEi\$apK	With hand shaky	With a shaky hand.

على ¹⁷	EalaY	Over/On	It introduces a sense of burden or obligation	علينا أن نعي ذلك	EalaynA >an naEiya *alika	On us to be aware that	We must be aware of this
		In	It introduces information about the psychological state	أنا على تلك الحالة من الكآبة	>anA EalaY tilka AlHALapi mina Alk bapi	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 مقهى	>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.	هل حدث ذلك كلُّه على مدى خمسة و عشرين عاماً	hal Hadava *alika kul~uhu EalaY madaY xampK 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?

¹⁷ Refer to Section 6.1.13 for important annotation issues.

حَتَّى	Hat~aY	Until/up to	It indicates the end point of an action	أَكَلْتُ السَّمَكَةَ حَتَّى رَأَسِهَا	>akaltu Alsamakapa Hat~aY ra>sihA	Ate (I) the fish <u>up to</u> head her	I ate the fish up to its head
كـ	ka-	As Like	It indicates comparison	لَمْ يَكُنْ مُوفِّقًا 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~api	<u>Since</u> the campaign of the elections	Since the election campaigns.
وَ [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!
تـ [تاء] [القسم]	ta- [the ta- of oath]	By (By God!)	It is used in oaths	تَاللَّهِ	taAll~Ahi	<u>By</u> God	By God!
بـ [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

¹⁸ mun*u has also the POS value of noun. Refer to Section 6.1.10 for important annotation decisions.

حاشا	Ha\$A	Except	It marks an exception	جاء الطلبة حاشا شازيد	jA'a AlTalabapu Ha\$A zaydK.	Came the students <u>except</u> Zaid	The students came <u>except</u> Zayd
خلا	xalA	Except	It marks an exception	جاء الطلبة خلا زيد	jA'a AlTalabapu xalA zaydK.	Came the students <u>except</u> Zaid	Everyone came <u>except</u> Zayd
و [أو المعية]	wa- [the wa- of accompaniment]	Together with Along with	It expresses the meaning of accompaniment	جاء الطلبة و زيدا	jA'a AlTalabapu wa zaydAF	Came the students <u>with</u> Zaid	The students came along with Zaid
إلا [حرف استثناء]	<iI~A [The exceptive particle]	Except	It marks an exception	جاء الطلبة إلا زيدا	jA'a AlTalabapu <iI~A zaydAF	Came the students <u>except</u> Zaid	The students came <u>except</u> Zayd

4.3.1.8 Interjections: INTERJ

They are words that express the speaker's reaction to a particular proposition or sentence. They are termed response particles in Arabic.

Arabic	Buckwalter	Gloss	Syntactic/Semantic Context	Examples	Transliteration	Gloss	Translation
بلى	baIY	Indeed!					

أَجَلْ	>ajal	OK!					
كَلَّا	kal~A	No!					
نَعَمْ	naEam	Yes!					
لَا	lA	No!					
أَهْ	h	Ah!					
تُرَى	turY	I wonder					
لَعْمَرِي	laEamry						
أُم	>um	um					
أَوْهْ	>aw~ah	Oh					
أَوَاهْ	>uw~Ah	Oh					
سَهْبِيَّكْ	\$ub~ayok	[magical encanta tion]					
أَلُو	>aluw	Hello					

4.3.1.9 Particles

Particles are divided based on semantic criteria to aid POS and TB annotation an interpretation. All are tagged, however, as PART. Particles do not assign case and they can be omitted without affecting or altering meaning and/or structure.

4.3.1.9.1 Vocative Particles: VOC_PART

Arabic	Buckwalter	Gloss	Syntactic/Semantic Context	Examples	Transliteration	Gloss	Translation
أ	>a	Hey	Vocative	أَمَحَمَّدُ تَعَالَ	>amuHam~aduN taEAla	Hey Muhammad come	Hey Mohamed, come here!
أَيُّ	>ay	Hey	Vocative	أَيُّ مُحَمَّدْ	>ay muHamad!	Hey Muhammad	Hey Mohamed!
يَا	yA	Hey	Vocative	يَا رَجُلُ تَعَالَ	yA rajulu taEAla	Hey man come	Hey Man, come!

		Oh	Exclamation	يَا لَهَا مِنْ هَزِيمَةٍ	yaA lahA min haziymapK	Oh what a defeat	What a defeat!
أَيُّهَا	>ay~uhA	Hey	Vocative	أَيُّهَا السَيِّدَاتُ	>ay~uhaA Alsy~idaAtu	Hey Ladies	Hey Ladies!
أَيُّهَا	>ay~atuhA	Hey [fem]	Vocative	أَيُّهَا الْفَتَاةُ	>ay~atuhA AlfatApu	Hey girl	Hey girl!
أَيَّا	>ayA	Hey	Vocative	أَيَّا رَجُلًا	>ayA rajulu	Hey Man	Hey Man!

4.3.1.9.2 Emphatic Particle: EMPHATIC_PART

Arabic	Buckwalter	Gloss	Syntactic/Semantic Context	Example	Transliteration	Gloss	Translation
لَا	la-	Definitely Certainly	It is used to emphasize or reinforce the meaning of the verb	لَأَضْرِبَنَّكَ	la>aDraban~aka	most certainly will beat (I) you	I most certainly will beat you

4.3.1.9.3 Restrictive Particles: RESTRIC_PART

Arabic	Buckwalter	Gloss	Syntactic/Semantic Context	Examples	Transliteration	Gloss	Translation
إِلَّا [أداة حَصْر]	<il~A [Restrictive particle]	Except for	It is used in a negative construction marking a restriction.	مَا جَاءَ إِلَّا مُحَمَّدٌ	mA jA' <il~A muHam~dN	Not arrived but muHam~ad	Nobody arrived but Muhamed
إِنَّمَا [أداة قَصْر]	<in~amA [Restrictive particle]	But indeed	It is used in a negative construction marking a restriction.	وَهُمْ لَيْسُوا كَتَابًا وَإِنَّمَا مُجَرَّدُ مُوظَّفِينَ	wa hum layswA kut~AbFA wa <in~amA mujar~adu muwaZ~afiyna	And they are not writers but mere bureaucrats	They are not writers but mere bureaucrats

سوى	siwaY	RESTRIC_PART	It marks a restriction and is always followed by a prepositional phrase.	It always occur before PP and can be substituted by <il~A	لم يحصل المؤتمر سوى على القليل	lam yaHsul Almu&tamar siwY EIY Alqalyl	
-----	-------	--------------	--	---	--------------------------------	---	--

4.3.1.9.4 Negative Particles: NEG_PART

Arabic	Buckwalter	Gloss	Syntactic/Semantic Context	Examples	Transliteration	Gloss	Translation
لا	lA	Not	It negates the meaning expressed by a verb in the present	هذا النوع من الحشائش لا ينمو في المناطق الحارة	ha*A AlnawEu mina AlHa\$A}iSi lA yanmwa fiy AlmanATiqi AlHaA r~api	This the type of the herbs <u>not</u> grow in the regions the hot	This type of grass does not grow in hot regions
لا...لا	lA...lA...	Neither...nor...	It negates persons or things. There must be two negated elements for "lA" to have the value of a negative particle in this context.	رأيت لا زيدا ولا عمرا	ra>aytu lA zaydAF wa lA EamrAF	I saw <u>neither</u> zayd <u>nor</u> Eamr	I saw neither zayd nor Eamr
لم	lam	Not	It negates the meaning expressed by a verb in the past	لم أشتأ السفر	lam >a\$a> Alsafara	Did <u>not</u> want (I) the travel	I did not want to travel

لَمَّا	lam~A	Not yet	It is used before imperfective verbs	هُوَ لَمَّا يَزَلْ حَدِيثَ السِّنِّ	huwa lam~A yazal Hadiyya Alsin~i	He not end modern the age	He has not yet grown up
مَا	mA	Not No	It negates the meaning of the constituent it precedes	مَا مِنْ شَيْءٍ أَنْ الْكَثِيرَ مِنَ الدُّوَلِ تُعَانِي مِنَ أَرْمَاتِ إِقْتِصَادِيَّةٍ	mA min \$ak~K >an~a Alkaviyra mina Alduwali tuEAniy min >azamAtK <iqtiSAdiy~apK	Not from doubt that the many from the countries suffer from crises economic	There is no doubt that many world states are suffering from economic crises.
لَنْ	lan	Never	It denies the occurrence of an action in the future. It always precedes imperfective verbs	لَنْ تَكُنْ هُنَاكَ مَقَاوِمَةً	lan takun hunAka muqaAwamapN	Never be there resistance	There will never be resistance there
لَيْسَ ¹⁹	laysa	Not	It reflects disjunction (between two identical grammatical categories)	تَعُودُ أَدْرَاجُهَا إِلَى الْأَطْفَالِ وَ لَيْسَ الشَّبَابِ	taEuwdu >adrAjuhA <iIY Al>aTfAli wa laysa Al\$abAbi	Due its causes to the children and not the adults	Its causes are described back to children not to young people.

4.3.1.9.5 Verb Particles: VERB_PART

They are regularly used to reinforce the past/perfect aspect of verbs (Badawi, 2004).

Arabic	Buckwalter	Gloss	Syntactic/Semantic Context	Examples	Transliteration	Gloss	Translation
قَدْ	qad	Actually Indeed	It reinforces the accomplishment of an action in the past	قَدْ حَقَّقْتُ بِالْفِعْلِ	qad Haq~aqtu biAlfiEli	Actually achieved (I) with actualization	I actually achieved
لَقَدْ	laqad	Already	It reinforces the accomplishment of	لَقَدْ انْتَهَتْ دُمُوعِي	laqad >intahat dumwEiy	Already finished tears my	My tears already stopped

¹⁹ When laysa inflects for gender, person, and number, it is annotated as a perfective verb with the tag PV.

		an action in the past	
--	--	-----------------------	--

4.3.1.9.6 Future Particle: FUT_PART

Arabic	Buckwalter	Gloss	Syntactic/Semantic Context	Examples	Transliteration	Gloss	Translation
سَوْفَ	sawfa	{Future particle}	It marks the future before verbs	سَوْفَ يَتَأَخَّرُ كَثِيرًا	sawfa yata>ax~aru kaviyrAF	Will be late (he) much	He will be very late
سَـ	sa-	{Future particle}	It cliticizes to verbs to mark the future	سَيَتَأَخَّرُ كَثِيرًا	sayata>ax~aru kaviyrAF	Will be late (he) much	He will be very late

4.3.1.9.7 Interrogative Particles: INTERROG_PART

Arabic	Buckwalter	Gloss	Syntactic/Semantic Context	Examples	Transliteration	Gloss	Translation
هَلْ	hal	{Interrogative particle}	It is used to form questions	هَلْ السَّعَادَةُ تَأْتِي أَمْ نَصْنَعُهَا؟	Hal AlsaEAdaputa>atiy >amnaSnaEuhA	Does happiness come or make we it	Does happiness simply come or do we make it?
أَ	>a	{Interrogative particle}	It is used to form questions	الْمُنِيرُ مَوْجُودٌ؟	>almudiyru mawjuwdN?	The director available	Is the director present?

4.3.1.9.8 Focus Particle: FOCUS_PART

Arabic	Buckwalter	Gloss	Syntactic/Semantic Context	Examples	Transliteration	Gloss	Translation
أَمَّا (أَمَّا ... ف)	>am~A (>am~A...fa...)	As for.....	It highlights the topic of the sentence	أَمَّا الْعَرِيسُ فَهُوَ سَلِيلُ الْحَسَبِ وَالتَّنَسُّبِ	>am~A AlEarysu fahuwa salylu AlHasabi w Alnasabi	As for the groom so he is pure the repute and the lineage	As for the bridegroom, he is pure of repute and lineage.

--	--	--	--	--	--	--	--

4.3.1.9.9 Response Conditional Particles : RC_PART

Arabic	Buckwalter	Gloss	Syntactic Context	Examples	Transliteration	Gloss	Translation
ل -	la-	So Then	It is used in conditional sentences introducing the apodosis sentence/main clause	إِن أُعْطِيتَ لَأُكْرِمَنَّكَ	<in >aETayta la>akramtuka	If give you <u>then</u> honor I you	If you give, I would honour you
ف - [الجزاءف]	fa- [The fa- of reward]	So Then	It is used in conditional sentences introducing the apodosis sentence/main clause which can be imperative, interrogative, non verbal, or verbal	إِن كَانَ لَكَ سِرٌّ فاحفظه إِن لَمْ تَكُنْ لَكَ رُؤْيَا فَلِمَنْ تَكُونُ إِن جِئْتَنِي فَأَنْتَ مُكْرَمٌ	<in kAna laka sir~N fAHfaDhu <in lam takun laka ru&yA faliman takwnu? <in jY'tany faAnta mukramN	If was for you secret <u>so</u> keep (I) it If did not was it for you vision <u>then</u> for whom be it If visit (you) me <u>then</u> you are treated with generosity	If you have a secret so don't reveal it If the vision wasn't yours (then so) for whom it is? If you visit me (then/so) I would be generous with you

				<p>إن قَبِلْتَنِي الْيَوْمَ فَقَدْ قَبِلْتَكِ أَمْسَ</p>	<p><in qabiltany Alyawma fqad qabiltuka Ams</p>	<p>If accepted (you) me today <u>then</u> accepted (I) you yesterday</p>	<p>If you accept me today (then so) I accepted you yesterday</p>
--	--	--	--	--	--	--	--

4.3.1.9.10 Connective Particle: CONNEC_PART

Arabic	Buckwalter	Gloss	Syntactic/Semantic Context	Examples	Transliteration	Gloss	Translation
--------	------------	-------	-------------------------------	----------	-----------------	-------	-------------

[فاء -فـ الربط]	fa- [The fa- of connection]	{Discourse connective}	It connects two clauses. It is most commonly used to introduce a comment clause after a clause starting with >am~A. It also connects a fronted modifier to the rest of the clause it modifies.	أَمَّا الْعَرِيسُ فَهُوَ سَلِيلُ الْحَسَبِ وَ النَّسَبِ وَ يَحْسَبُ بِالتِّيَانَسْكِ فَلَيْسَ هُنَاكَ أَيُّ مُشْكَلَةٍ لِتَمْدِيدِ المُبَاحَثَاتِ	>am~A AlEarysu fahuwa salylu AlHasabi w Alnasabi wa biHasabi baltiyAnskiy, falaysa hunAka >ay~u mu\$kilapK litamdiydi AlmubAHavAti	As for the groom <u>well</u> <u>then</u> he is good the reputation and the family And according to bAlityAnskiy, then there is not any problem in extending the talks	As for the bridegroom, he is of good reputation and family And according to bAlityAnskiy, then there is not any problem in extending the talks
--------------------	---------------------------------------	---------------------------	--	--	---	---	--

4.3.1.9.13 Jussive Particle: JUS_PART

Arabic	Buckwalter	Gloss	Syntactic/Semantic Context	Examples	Transliteration	Gloss	Translation
لـ	li-	Lets	It expresses a meaning of request or command. It assigns the jussive to the verb it cliticizes to.	لِنَذْهَبِ	lina*hab	Lets go	Lets go

4.3.1.9.12 Other Particles: PART

Arabic	Buckwalter	Gloss	Syntactic/Semantic Context	Examples	Transliteration	Gloss	Translation
أَيُّ	>ay	{Particle of explanation}	It provides further explanation or reformulation of a previous statement	أَيُّ أَنْ هَذِهِ الدُّوَلُ	>ay >an~a ha*ihi Alduwal	<u>That is to say</u> that these countries	That is to say that these countries
إِنِّي	<iy~A	{Bound pronoun marker}	It hosts the object pronoun when separated from the entity it normally cliticizes to	إِنِّي أَعْنِي	<iy~Aka >aEniy	You I mean	I mean you.
أَلَا حَرْفِ [تَحْضِيضِ]	>alA	{Imperative/warning particle}	It expresses warning or incitation	أَلَا إِنَّهْضُ وَ سِرْ	>alA <inhaD wa sir	Go get up and walk	Go get up and walk
حَتَّى	Hat~Y	Even	It has an emphatic meaning of inclusion	لَمْ يَمْلِكْ حَتَّى فَدَانًا وَاحِدًا	lam yamluk Hat~Y fad~AnAF wAHidAF	Not own <u>even</u> acre single	He does not own even one single acre
إِذَا [الْفَجَائِيَّةِ]	<i* [<i*A of surprise]	Suddenly	It introduces an unexpected event	كَانَ يَلْعَبُ إِذْ سَقَطَ	kAna yalEabu <i* saqaTa		As he was playing, he fell.
إِذَا	<i*A	Suddenly					
مِنْ	min		It premodifies a noun in a subject position	مَا مِنْ أَحَدٍ جَاءَ	mA min >aHadK jA'a	Nobody came	Nobody came
رُبُّ	rub~a	How	It premodifies a noun				

		much...!	in a subject position				
ما	mA	{Particle}	It comes after an adjective with an adverbial meaning	السُّلْطَاتُ التَّقْدِيَّةُ غَالِبًا مَا تَعْتَمِدُ أَسْعَارَ فَايْذَةَ مُرْتَفِعَةً	AlsuluTAt Alnaqdiy~ap gAlibAF mA taEtamidu >asEAr fa}idap murtafiEap	The monetary authorities relies mostly on prices that have high interests	The monetary authorities relies mostly on prices that have high interests
كأنما	ka>an~amA	as if	It is a sentence particle.	كأنما الأرضُ كلهاملكُ لهُ	ka>an~amA Al>arDu kul~uhAmilkN lahu.	As if the world, all of it, is his possession.	As if the world, all of it, is his possession.
وَوَاو الزائدة	wa [the extra wa-]		It is a constituent particle that can be omitted without affecting the meaning of the sentence.	سَبَقَ وَرَأَيْتُ ذَلِكَ مِنْ قَبْلُ	sabaqa wa ra>ayotu *`lika min qablu		I saw that before
بِ	bi-		It is a constituent particle premodifying a noun in a subject position	هُم عَادُوا عَنِ الْمُقَاطَعَةِ فَيَا يَا قِبَالَهُمْ يَخْلُطُ الْأُورَاقَ	hum EAduwA Ean AlmuqATaEap fa<i*A bi<iqobAlihim yaxoliTu Al>aworAqa	They backed away from the boycott so suddenly their participation mixed the cards	They backed away from the boycott so suddenly their participation redistributed the game cards

4.3.1.10 Coordinating Conjunctions: CONJ

Conjunctions are used to coordinate and link independent constituents with each other.

Arabic	Buckwalter	Gloss	Syntactic/Semantic Context	Examples	Transliteration	Gloss	Translation
--------	------------	-------	----------------------------	----------	-----------------	-------	-------------

- فـ [العطف فاء]	fa- [The fa-of coordination]	And then	It marks a temporal sequence between two events	نظرة فابتسامته هما بداية القصة خَرَجَ مِنْ بَيْتِهِ فَرَأَى صَدِيقَهُ عَلِيًّا فِي الشَّارِعِ فَتَادَاهُ وَ دَهَبَ مَعَهُ إِلَى الْعَمَلِ	naZrapN faAibtisAmapN humA bidAyapu AlqiS~api xaraja min baytihi fara>aY Sadyqahu Ealy~A fy Al\$AriEi fanAdAhu wa *ahaba maEahu <ilaY AlEamali	A look <u>then</u> a smile they are the start the story Went out (he) from house his <u>then</u> saw friend his Ali in the street	A look then a smile were the beginning of the story He went out; then he saw his friend Ali in the street; then he shouted to him and went with him to work.
- و [الربط واو]	wa- [The wa-of coordination]	And	It coordinates two constituents	الضوءُ ناصعٌ و الجوُّ حارٌّ	AlDaw'u nASiEN wa Aljaw~u Har~N	The light bright <u>and</u> the weather hot	The light is bright and the weather is hot
ثم	vum~a	And then	It describes a temporal sequence	يَعْمُرُ الرَّجُلُ ثُمَّ يُغْلِقُ سِتَارَةَ النَّافِذَةِ	yaquwmu Alrajulu vum~a yugliqu sitArapa AlnaAfi*pi	Gets up the man <u>then</u> closes curtains the window	The man gets up, then closes the window curtain
أو	>aw	Or	It expresses disjunction (a choice between two alternatives)	رُبَّمَا لِأَنَّهَا لَنْ تَجِدَ جَوَابًا أَوْ لِأَنَّهَا أَلْفَتْ هَذِهِ الْحَيَاةَ	rub~amA li>an~ahaA lan tajida jawAbAF >aw li>an~ahA >alifat ha*ihi AlHayAp	Perhaps because she will not find answer <u>or</u> because she got accustomed this the life	Maybe because she would never find an answer or because she liked this way of life
أم	>am	Or	It implies exclusive disjunction	فَهَلْ نَمَلًا فَرَاغْنَا بِالْعَزْفِ عَلَى الْبِيَانُو؟ أَمْ بِالْعَزْفِ عَلَى الْعُودِ؟	Fahal naml>u farAganA biAlEazfi EalY AlbiyAnuw? >am biAlEazfi EalY AlEuwdi?	So do we fill leisure our with the playing on the piano <u>or</u> with the playing on the lute	Shall we fill our spare time with playing the piano or playing the lute?
إمَّا	<im~A	(Either...or...)	It sets a choice between two alternatives	إِمَّا أَنْ تَزِيدَ فِي الْإِنْتِاجِ وَ إِمَّا أَنْ تُقَلِّلَ فِي الْمَصْرُوفَاتِ	<im~A >an naziyda fiy Al<intAj wa <im~A >an nuqal~ila fiy AlmaSrouwfAti	Whether that increase (we) in the production or decrease in the spendings	Whether we increase production or that we reduce costs
إمَّا	<im~a	(Either...or...)	It sets a choice between two alternatives				

بَلْ	bal	But Rather	It introduces an assertion that is in opposition with a previous statement	لَمْ تَكُنْ هُنَاكَ مُصَادِرَاتٍ بَلْ كَانَتْ هُنَاكَ حُرِّيَّةٌ lam takun hunAka muSAdrAtN bal kAnat hunAka Hur~iyap	Did not be there confiscations <u>rather</u> was there freedom	There were no confiscations rather there was freedom
حَتَّى	Hat~Y	And	It coordinates two constituents with the same function	أَكَلْتُ السَّمَكَةَ حَتَّى رَأْسَهَا >akaltu Alsamakap Hat~Y ra>sahA	Ate I the fish <u>and</u> head its	I ate the fish along with/and its head.
لَكِنْ	l'kin	But	It introduces a contrast.	كَثِيرُونَ حَاولُوا وَلَكِنْ بِأَوْسَاةٍ بَقِيَتْ مُشِينَةً kaviyrwna HawalwA wa l'kin bAwwA bifa\$alK mu\$iyNk	Many tried <u>and but</u> ended up with failure shameful	Many tried but came back with shameful failure.
لَا [العطف]	IA [The IA of coordination]	And not	It corroborates an assertion that has been previously made	رَأَيْتُ زَيْدًا لَا عَمْرًا ra>aytu zaydAF IA EamrAF	Saw (I) Zaid <u>not</u> Amr	I saw Zayd not Amr
كَمَا	kamA	And	It links two matrix clauses	نُجَدِّدُ التَّحِيَّةَ لِلْمُقَاوَمَةِ كَمَا نُحَيِّي الْإِنْتِظَاظَةَ فِي فَلسطِينِ nujad~idu AltaHiy~apa lilmuqAwamap kamA nuHay~iy Al<intifADapa fiy filsTiyna		We renew the greeting to the resistance and we respect the Palestenian resistance

--	--	--	--	--	--	--	--

4.3.1.11 Subordinating Conjunctions: SUB_CONJ

A subordinating conjunction marks a sentence as dependent to another sentence that is independent and called the main clause.

Arabic	Buckwalter	Gloss	Syntactic/Semantic Context	Examples	Transliteration	Gloss	Translation
إن حَرَفُ فِ شَرَطٍ	<in [The conditional <in]	If	It introduces a conditional clause	إن قَبِلُوا مِثْلَ هَذِهِ الْإِتْفَاقِيَّةِ مَرَقَّتْهُمُ الْجَمَاهِيرُ	<in qabilwA mivla ha*ihi Al<itfAqiy~p maz~aqathum AljamAhiyr	If agree (they) such us this the contract tear (a part) them the public	If they accept such agreements the crowds in the streets will tear them apart
أن [المصدرية]	>an [The gerundive >an]	That	It introduces a clause that can be replaced by the masdar of the verb	فَبِمَا كَانُوا أَنْ يُقَرَّرَ بِنَاءٍ عَلَى مَا تَقَدَّمَ	fabi<mkAninA >an nuqar~ira binA'F EalY mA taqad~ama	So in capacity our <u>that</u> to decide depending on what preceded	So we can decide based on what preceded
				فَبِمَا كَانُوا التَّقْرِيرُ بِنَاءٍ عَلَى مَا تَقَدَّمَ	fabi<mkAninA Altaqriyr binA'F EalY mA taqad~ama	So in capacity our the decision depending on what preceded	So it is in our capacity to decide on what has been proposed.
أنَّ	>an~a	That	It introduces a clause that	قَالَ أَنَّهُ مَرِيضٌ	qAla >an~ahu mariyDN	Said <u>that</u> he is	He said that

			indicates reported/indirect speech			sick	he is sick
حتى	Hat~Y	So that	It introduces a clause indicating the cause of an event in the main clause	إِظْطَرَّرْتُ لِأَكْلِ السَّمَكَةِ حَتَّى لَا أَمُوتَ مِنَ الْجُوعِ	<iZTurirtu li>akli Alsamakapi Hat~Y IA >amuwta mina AljuwEi	Compelled to eating the fish so <u>that</u> not to die out of hunger	I had to eat the fish so that I don't die out of hunger
إذ إذا	<i* <i*A	If/whether	It introduces a clause indicating the cause of an event in the main clause	لا ندري اذا في الإمكان تحسين الجباية.	IA nadoriy <i*A kAna fiy Al<imokAni taHosiyynu AljibAyapi	Not know <u>if</u> possible improving tax.	We don't know if it is possible to better the tax.
إذ	<i*	In view of the fact that/since	It introduces a clause that provides an explanation of an event in the main clause	ذَلِكَ الْمَشْرُوعُ لَمْ يَسْتَمِرَّ إِذْ يَبْدُو أَنَّ الْحُكُومَةَ الْبَرِيطَانِيَّةَ قَدْ صَرَفَتْ النِّظْرَ عَنْهُ	*alika Alma\$rwEa lam yastamir~a <i* yabduw >an~ AlHukwmapa AlbriyTAniy~ap qad Sarafat AlnaZara Eanhu	That the project did not continue <u>as</u> seems that the government the British has removed the view on it	That project did not continue as it seems that the British government turned its attention away from it.
		While		وَ إِذْ أَكَّدَ أَنَّ الْمَصَارِفَ الْلَبْنَانِيَّةَ تَتَمَتَّعُ بِنِسْبَةٍ مَلَاءَةٍ عَالِيَةٍ قَالَ إِنَّ اللَّجْنَةَ تَقْرَضُ مَلَاءَةً عَلَى مَخَاطِرِ التَّشْغِيلِ وَ عَلَى الْمَصَارِفِ أَنْ تَتَّقِيَدَ بِهَا	wa <i* >ak~ada >an~a AlmaSArifa AllubnAniy~ap tatamat~aEu binisbapi malA'pK EAliyapK, qAla <in~a Allajnapa tafriDu malA'ap EalY maxATiri Alta\$gyli wa EalY AlmaSArif >an tataqay~ada bihA	And while he stressed that the Lebanese banks enjoy a high percentage of solvency, he said that the committee forces another solvency on job risks that banks	And while he stressed that the Lebanese banks enjoy a high percentage of solvency, he said that the committee forces another

						have to stick to.	solvency on job risks that banks have to stick to.
بَيْنَمَا	baynamA	Whereas	It contrasts between two ideas	بَيْنَمَا تُشِيرُ الأرقامُ لإنخفاضِ مُعدَّلِ الإِدخارِ فإنَّ أرقامَ البَنكِ المَركَزِي أيضًا يَشِيرُ لِتراجُعِ في نِسبَةِ الإِسْتِمارِ	baynamA tu\$yru Al>arqAmu l<nxifADi muEad~ali Al<id~ixAri fa<in~a >rqAma Albank Almarkaziy >ayDFA y\$yru litrAjuEK fiy nisbapi Al<istivmAri	While point the digits to decrease average the saving so indeed digits the bank the central also point to decrease in rate the investment	While the figures point to a crash in average saving, the central Bank figures also point to a regression in the percentage of investment.
طالما	TalamA	(As long as)	It introduces a conditional clause	طالما أن ذلك لا يضرُّ بالآخرين سنُباشِرُ بالعملِ	TalamA >an~a *alika la yaDur~u biAl xariyna, sanubA\$iru biAlEamali	As long as that not harm with the others, will we start work	As long as that does not harm others, we will start working
حالما	HalamA	As soon as		هذا أمرٌ تَتَّضِحُ صِحَّتُهُ حالما تَتَّقَصَى السِيرَةَ الذاتِيَّةَ	ha*A >amrN tat~aDiHu SiH~atuhu HalamA tataqaD~Y Alsiyrapu Al*Atiy~ap	This is a matter whose truth becomes obvious as soon as you follow the autobiography	This is a matter whose truth becomes obvious as soon as you follow the autobiography
لَمَّا	lam~A	As	It provides information about	لَمَّا هَرَّهُ وَجَدَ الرَّجُلُ مَيْتًا	lam~A haz~ahu wajada	As he lifted him	As he lifted

			the time of an event		Alrajulu maytAF	he found the man dead	him he found the man dead
فيما	fiymA	While	It describes simultaneity between two events	فيما كان أخي نائماً خَرَجْتُ مِنَ الْمَنْزِلِ	fiymA kAna >axiy nA}imAF, xarajtu mina Almanzili	<u>While</u> was my brother sleeping, left I from the house	While my brother was sleeping, I left the house
رَبِّمَا	rayvamA	Until		إِتَّخَذَ هَذَا الْإِجْرَاءَ رَبِّمَا يَجِدُ مَكَانًا لَائِقًا	<it~axa*a ha*A Al<ijrA' rayvamA yajidu makAnAF lA}iqAF	He took this decision <u>until</u> he would find r place suitable	He took this decision until he would find another suitable place
كيما	kaymA	So as/in order to	It introduces a clause expressing purpose	ذَهَبَ إِلَى جَدَّتِهِ كَيْمَا يَجِدَ عِنْدَهَا بَعْضَ الْحُبِّ	*ahaba <iLY jad~atihi kaymA yajida EindahA baEDa AlHub~i	He went to his grandmother <u>in order to</u> find in her some love	He went to his grandmother in order to find in her some love
كي	kay	So as	It introduces a clause expressing the reason of an already undertaken action	كِي تَلْعَبُ نَوْرًا أَسَاسِيًّا	kay talEaba dawrAF >asAsiy~AF	<u>In order to</u> play (she) role basic	In order to play a fundamental role
لو	law	If	It introduces a conditional clause	لَوْ سَأَلْتَنِي لَقُلْتُ لَهُ	law sa>alany laqultu lahu	<u>If</u> ask (he) me would say (I) to him	If he had asked me, then I would have said to him
لولا	lawlA	If not for		كَانُوا قَدْ اتَّفَقُوا عَلَى التَّنْفِيزِ لَوْلَا أَنَّ الْحَرْبَ قَدْ قَامَتْ	kAnwA qad <itafaquwA EalY Altanfiy*I lawlA >an~a AlHarba qad qAmat	Were (they) has agreed on carrying out <u>if not for</u> that the was has wage	They would have agreed to the carrying out [of it] were it not for the

							fact that the war
و- [واو الحال]	wa- [The wa- of manner]	With/While	It introduces a clause describing the manner with which an event took place.	نَحَوَ 400 مِكْسِيكِي تُوُقُوا عَرَقًا وَ هُمْ يُحَاوِلُونَ عُبُورَ الْحُدُودِ الطَّوِيلَةِ بَيْنَ الْبَلَدَيْنِ	naHwa 400 miksyky~ tuwuf~wA garaqAF wa hum yuHAWiluwna Eubuwra AlHuduwdi AlTawiylapi bayna Albaladyni	Approximately, 400 Mexicans died by drowning while they were trying crossing the long borders between the two countries	Approximately, 400 Mexicans died by drowning while they were trying crossing the long borders between the two countries
ف- [فاء السببية]	fa- [The fa- of causality]	So that	It expresses the result or the cause of the main clause which must be imperative, interrogative, wish, or negation	اِنْتِنَا فَنُكْرِمُكَ أ هَذَا زَيْدٌ فَأَكَلَمَهُ لَيْتَ الشَّبَابَ يَعُودُ يَوْمًا بِمَا فَأَخِيرَهُ فَعَلَ الْمَشْيَبُ لَا أَسْتَطِيعُ رُؤْيَكَ يَوْضُوحَ	A'tinA fanukrimaka >a ha*A zaydN fa>ukal~imuhu layta AlSabAba yaEwdu yawmAF fa>uxbirahu bimA faEala Alma\$iybu lA >astaTyEu	Visit (you) us <u>so that</u> be generous (we) you Is this Zaid <u>so that</u> speak I him Wish the youth return someday <u>so that</u> tell(I) him with what did old age Not can (I) see you <u>since</u>	Visit us so that we will be generous with you Is this Zayd so that I will talk to him I wish that the youth returns one day so that I will tell it what did to me the old

		Since/Because		فَالضَّلَامُ دَامِسٌ	ru&yataka biwuDwHK faAlDalAmu dAmisN	darkness deep black	I can't see you clearly because it is too dark
ما [المصدرية ما]	mA [The gerundive mA]	That	It introduces a clause that can be replaced by the masdar of the verb that follows it.	بَعْدَمَا يَأْتِي الصَّبَاحُ نَذْهَبُ إِلَى السُّوقِ	baEdamA ya>tiy AlSabAhu, na*habu <iY Alsuwqi	After <u>that</u> come the morning go (we) to the market	After the sun rises, we will go to the market.
				بَعْدَ إِتْيَانِ الصَّبَاحِ نَذْهَبُ إِلَى السُّوقِ	baEda <ityAni AlSabAHi, na*habu <iY Alsuwqi	After coming the morning go (we) to the market	After sunrise, we will go to the market.
ما [ما الشرطية]	mA [The conditional mA]	Unless	It introduces a conditional sentence	لَنْ أَكُلَ مَا لَمْ تَأْتِ	lan kula mA lam ta>tiy	I won't eat <u>unless</u> you come	I won't eat unless you come
		As soon as/as long		الهِندُ مُسْتَعِدَّةٌ لِجَوَارِ تَنْزِيلِيَّ مَا أَنْ يَنْتَهِيَ الْإِرْهَابُ عَبْرَ			India is ready

ما		as		الْحُدُودِ	Alhindu mustaEid~apN liHiwArK vunA}iy~K mA >an yantahiy Al<irhAbu Eabra AlHuduwdi	India is ready for a bilateral dialogue <u>as</u> <u>soon as</u> terrorism will end up on the frontiers	for a bilateral dialogue as soon as terrorism will end up on the frontiers
أثما	>ana~mA	That	It always occurs in the middle of the sentence (unlike <in~amA)	تَعْلَمُونَ أَثْمَا الْأَعْمَالُ بِالنِّيَّاتِ	taElamuwna >an~amA Al>aEmAlu biAlniy~Ati	You know <u>that</u> deeds are with intentions.	You know that deeds are based on intentions.

4.3.2 POS Summary Chart

The following chart represents a listing of all the entries under the closed categories of noun quantifiers, pronouns, pseudo verbs, non-inflectional verbs, adverbs, prepositions, particles, conjunctions and subordinating conjunctions.

Arabic Entry	Buckwalter	POS Tag (s)						
كُلٌّ	kul~	NOUN_QUANT						
كُلًّا	kilA	NOUN_QUANT						
كُلَّتَا	kiltA	NOUN_QUANT						
جَمِيعٌ	jamiyE	NOUN_QUANT						
بَعْضٌ	baED	NOUN_QUANT						
بَعْضٌ	biDE	NOUN_QUANT						
جُلٌّ	jul~	NOUN_QUANT						
كُلٌّ	kul~	NOUN_QUANT						
كُلًّا	kilA	NOUN_QUANT						
كُلِّفَةٌ	kAf~ap	NOUN_QUANT	ADJ					
مَعْظَمٌ	muEDam	NOUN_QUANT						
أَغْلَبٌ	>aglab	NOUN_QUANT	ADJ_COMP					
أَكْثَرٌ	>akvar	NOUN_QUANT	ADJ_COMP					

نصف	niSf	NOUN_QUANT						
شطر	\$aTr	NOUN_QUANT						
ضعف	DiEf	NOUN_QUANT						
ربع	rubuE	NOUN_QUANT						
ثلث	vuluv	NOUN_QUANT						
خمس	xumus	NOUN_QUANT						
سدس	sudus	NOUN_QUANT						
سبع	subE	NOUN_QUANT						
ثمان	vumn	NOUN_QUANT						
تسع	tusuE	NOUN_QUANT						
عشر	Eu\$ur	NOUN_QUANT						
حوالي	HawAlaY	NOUN_QUANT						
زهاء	zuhA'	NOUN_QUANT						
قراءة	qurAbap	NOUN_QUANT						
قالب	qAb	NOUN_QUANT						
كامل	kAmil	NOUN_QUANT						
أنا	>anA	PRON_1S						
نحن	naHnu	PRON_1P						
أنت	>anta	PRON_2MS						
أنت	>anti	PRON_2FS						
أنتم	>antumA	PRON_2D						
أنتم	>antum	PRON_2MP						
أنتم	>antun~a	PRON_2FP						
هو	huwa	PRON_3MS						
هي	hiya	PRON_3FS						
هما	humA	PRON_3D						
هم	hum	PRON_3MP						
هن	hun~a	PRON_3FP						

ني-	-niy	PRON_1S	POSS_PRON_1S					
ي-	-y	PRON_1S	POSS_PRON_1S					
يَ-	-ya	PRON_1S	POSS_PRON_1S					
نا-	-nA	PRON_1P	POSS_PRON_1P					
ك-	-ka	PRON_2MS	POSS_PRON_2MS					
كِ-	-ki	PRON_2FS	POSS_PRON_2FS					
كُما-	-kumA	PRON_2D	POSS_PRON_2D					
كُم-	-kum	PRON_2MP	POSS_PRON_2MP					
كُنْ-	-kun-a	PRON_2FP	POSS_PRON_2FP					
هُ-	-hu	PRON_3MS	POSS_PRON_3MS					
ها-	-hA	PRON_3FS	POSS_PRON_3FS					
هما-	-humA	PRON_3D	POSS_PRON_3D					
هم-	-hum	PRON_3MP	POSS_PRON_3MP					
هنْ-	-hun-a	PRON_3FP	POSS_PRON_3FP					
هذا	ha*A	DEM_PRON_MS						
ذا	*aA	DEM_PRON_MS						
ذاك	*Aka	DEM_PRON_MS						
ذلك	*alika	DEM_PRON_MS						
تلك	tilka	DEM_PRON_F						
هذه	ha*ihi	DEM_PRON_F						
هذي	ha*iy	DEM_PRON_F						
هاتي	hAtihi	DEM_PRON_F						
هات	hAti	DEM_PRON_F	VERB					
ها	hA	DEM_PRON	VERB					
تيك	tiyka	DEM_PRON_FS						
تاك	tAka	DEM_PRON_FS						

هاتان / هاتين	hAtAni/hAtayn	DEM_PRON_FD						
تَيْن / تان	tAni/tayni	DEM_PRON_FD						
تَانِك / تَيْنِك	tAnika	DEM_PRON_FD						
ذَان / ذَيْن	*aAn/*ayn	DEM_PRON_MD						
ذَانِك / ذَيْنِك	*aAnika/*aynaka	DEM_PRON_MD						
هَذَان / هَذَيْن	ha*Ani/ha*ayni	DEM_PRON_MD						
ذَلِكَمَا	*alikumA	DEM_PRON_D						
ذَلِكُمْ	*alikum	DEM_PRON_MP						
ذَلِكُنَّ	*alikul~a	DEM_PRON_FP						
أُولَئِكُمْ	>uwla}ikum	DEM_PRON_MP						
أُولَئِكَ	>uwla}ka	DEM_PRON_P						
أُولَءِ/هُؤُلَءِ	>uwla'h'ulA'i	DEM_PRON_P						
تِلْكُمْ	tilkum	DEM_PRON_FP						
أَيِّن	>ayona	INTERROG_ADV	REL_ADV					
أَي	>ay~	INTERROG_PRON	REL_PRON	NOUN_QUANT				
أَيَّة	>ay~ap	INTERROG_PRON	REL_PRON	NOUN_QUANT				
أَيَّا	>ay~AF	REL_PRON						
كَمْ	kam	INTERROG_PRON	REL_PRON	REL_ADV				
مَهْمَا	mahomA	REL_PRON						
كَيْفَ	kayofa	INTERROG_ADV	REL_ADV					
لِمَ	lima	INTERROG_ADV	REL_ADV					
لِمَا	limA	INTERROG_ADV	REL_ADV	Tokenization may apply				
لِمَاذَا	lima*A	INTERROG_ADV	REL_ADV					
مَا	mA	INTERROG_PRON	REL_PRON	NEG_PART	SUB_CONJ	NOUN	EXCLAM_P RON	VERB
مَاذَا	mA*A	INTERROG_PRON	REL_PRON					
مَنْ	man	INTERROG_PRON	REL_PRON					
مَتَى	mataY	INTERROG_ADV	REL_ADV					

أَيُّ	>an~Y	INTERROG_ADV	REL_ADV					
أَيَّانَ	>ay~Ana	INTERROG_ADV	REL_ADV					
الـ	Al-	REL_PRON						
الَّذِي	Ala*y	REL_PRON						
الَّذِينَ	Ala*iyna	REL_PRON						
الَّتِي	Alatiy	REL_PRON						
الَّذَانِ	Alla*Ani	REL_PRON						
الَّذَيْنِ	Alla*ayni	REL_PRON						
اللَّائِي	Alla}iy	REL_PRON						
اللَّتَانِ	AllatAni	REL_PRON						
اللَّتَيْنِ	Allatayni	REL_PRON						
اللَّائِي	AllAty	REL_PRON						
اللَّوَاتِي	AllawAty	REL_PRON						
حَيْثُ	Hayovu	REL_ADV						
كَيْفَمَا	kayofamA	REL_ADV						
حَيْثُمَا	HayovumA	REL_ADV						
أَيْنَمَا	>ayonamA	REL_ADV						
كُلَّمَا	kul~amA	REL_ADV						
أَنَّ	>an~a	SUB_CONJ						
إِذْ	<i*	SUB_CONJ	PART					
فِيمَا	fiymA	SUB_CONJ	Tokenization may apply					
إِذَا	<i*A	SUB_CONJ	PART					
إِنْ	<in	SUB_CONJ						
أَنَّ	>an	SUB_CONJ						
أَمَّا	>an~amA	SUB_CONJ						
بَيْنَمَا	bayonamA	SUB_CONJ						

فـ	fa-	SUB_CONJ	CONJ	CONNEC_PART	RC_PART			
حالما	HalamA	SUB_CONJ						
كي	kay	SUB_CONJ						
كيما	kayomA	SUB_CONJ						
لو	law	SUB_CONJ						
لولا	lawolA	SUB_CONJ						
رَيْمًا	rayovamA	SUB_CONJ						
طالما	TAlamA	SUB_CONJ	ADV					
و-	wa-	SUB_CONJ	CONJ	PREP	PART			
إنْ	<in~a	PSEUDO_VERB						
كَانَ	ka>an~a	PSEUDO_VERB						
لا	lA	PSEUDO_VERB	NEG_PART	CONJ	INTERJ			
لَعَلَّ	laEal~a	PSEUDO_VERB						
عَلَّ	Eal~a	PSEUDO_VERB						
لَكِنْ	l'kin~a	PSEUDO_VERB						
لَيْتَ	layota	PSEUDO_VERB						
أَنْذَاكَ	na*Aka	ADV						
أَنْذَيْكَ	na}i*K	ADV						
إِذَا	<i*AF	ADV						
إِذَنْ	<i*an	ADV						
أَيْضًا	>ayoDAF	ADV						
عَمَّيْكَ	Eam~a}i*K	ADV						
عَامَّذَاكَ	Eama*Aka	ADV						
سَمَّذَاكَ	sanata*Aka	ADV						
عِنْدَذَاكَ	Einoda*Aka	ADV						
عِنْدَيْكَ	Einoda}i*K	ADV						

حَتَّى	Hat~Y	PREP	SUB_CONJ	CONJ	PART			
لَا	<il~A	PREP	RESTRIC_PART					
إِلَى	<ilaY	PREP						
بِ-	bi-	PREP	PART					
عَلَى	EalY	PREP	NOUN					
عَنْ	Ean	PREP						
خِلا	xalA	PREP	VERB					
حَاشَا	HA\$A	PREP	VERB					
عِدا	EadA	PREP	VERB					
فِي	fiy	PREP						
كَ-	ka-	PREP						
لِ-	li-	PREP	JUS_PART					
لِ-	la-	EMPHATIC_PART	RC_PART	PREP				
مِنْ	min	PREP	PART					
مُنْذُ	mu*	PREP	NOUN					
مُنْذُ	mun*u	PREP	NOUN					
ت-	ta-	PREP						
بَلَى	balY	INTERJ						
أَجَلْ	>ajal	INTERJ						
كَلَّا	kal~A	INTERJ						
نَعَمْ	naEam	INTERJ						
أَهْ	>ah	INTERJ						
لَعْمَرِي	laEamory	INTERJ						
أُم	>um	INTERJ						
أَوْهْ	>aw~ah	INTERJ						
أُوَاهْ	>uw~Ah	INTERJ						
سُبَيْكْ	\$ub~ayok	INTERJ						
وَاهْ	wAh	INTERJ						

ألو	>aluw	INTERJ						
إيّا	<iy~A	PART						
رُبّ	rub~a	PART						
إِنَّمَا	<in~amA	RESTRIC_PART	Tokenization may apply					
أيا	>ayA	VOC_PART						
أ	>a	INTERROG_PART	VOC_PART					
هل	hal	INTERROG_PART						
أَيُّهَا	>ay~uhA	VOC_PART	Tokenization may apply					
أَيُّهَا	>ay~atuhA	VOC_PART	Tokenization may apply					
يا	yA	VOC_PART						
كأَيُّهَا	ka>an~amA	PART						
وا-	wA-	VOC_PART						
اه-	-Ah	VOC PART						
هيهات	hayohAt	VERB						
هَلَمْ	halum~a	VERB						
هَيَّا	hay~A	VERB						
تَعَالَى	taEAla	VERB						
هَبْ	hab	VERB	CV					
حَذَارِي	Ha*Ari	VERB						
حَبِّدَا	Hab~a*A	VERB						
أه	lh	VERB						
أَفْ	>uf~	VERB						
نَاهِيكَ	nAhiyka	VERB						
بِنَسْ	bi} sa	VERB						
سَاتُنْ	\$at~Ana	VERB						

سرعان	surEAna	VERB						
أمين	miyn	VERB						
نعم	niEma	VERB						
حَيًّا	Hay~A	VERB						
يَنْبَغِي	yanbagy	VERB						
عَسَى	EasY	VERB	PSEUDO_VERB					
تُرَى	turY	VERB	INTERJ					
لَمْ	lam	NEG_PART						
لَمَّا	lam~A	SUB_CONJ	NEG_PART					
لَنْ	lan	NEG_PART						
لَيْسَ	layosa	PV	NEG_PART					
أَمَّا	>am~A	FOCUS_PART						
أَلَا	>alA	PART	Tokenization may apply					
لَقَدْ	laqad	VERB_PART						
قَدْ	qad	VERB_PART						
سَوْفَ	sawofa	FUT_PART						
إِمَّا	<im~A	CONJ						
أَمْ	>am	CONJ						
كَمَا	kamA	CONJ	Tokenization may apply					
أَوْ	>aw	CONJ						
بَلْ	bal	CONJ						
لَكِنْ	l'kin	CONJ						
أَيُّ	>ay	PART	VOC_PART					
ثُمَّ	vum~a	CONJ						

5.0 ADDITIONAL INFORMATION ON ANNOTATION & POLICIES

5.1 POS Annotation

A change in POS annotation policy has already taken place with regard to certain entries, particularly prepositions and nouns. This section has been set for annotators with a POS annotation experience based on the old guidelines.

5.1.1 List of Former Prepositions and/or Adverbs **deleted entries that were not in BAMA 2.0. They were here probably because of the data from the corpus.**

The following entries represent entries previously treated as prepositions and/or adverbs. In the present guidelines, all are nouns²⁰ which act as adverbial modifiers expressing temporal, locative or any other adverbial relationships. In most cases, these prepositionals will be in the accusative unless they are preceded by a preposition in which case they take the genitive case. The accusative case but also the nominative and to a lesser extent the genitive (see >amsi) is also used in Arabic to indicate an adverbial function.

Arabic	Buckwalter	New POS Tag	Gloss
أمام	>amAma	NOUN + CASE_DEF_A CC	The front +[def. acc.]
إثر	<ivora	NOUN + CASE_DEF_A CC	The aftermath/something after+[def. acc.]
إزاء	<izA'a	NOUN + CASE_DEF_A CC	The position towards +[def. acc.]
بعد	baEoda	NOUN + CASE_DEF_A CC	The aftermath/something after +[def. acc.]
بعد	baEodu	NOUN +	The

²⁰ Refer to Badawi (2004), Qabbish, A. (1984) and Sayyid Ahmed, A., and Eabd Alganiy, I. (?)

		CASE_DEF_NOM	aftermath/something after +[def. nom.]
بين	bayona	NOUN + CASE_DEF_ACC	The position between +[def. acc.]
تجاه	tujAha	NOUN + CASE_DEF_ACC	The position towards +[def. acc.]
تحت	taHota	NOUN + CASE_DEF_ACC	The position under +[def. acc.]
تلو	tilowa	NOUN + CASE_DEF_ACC	Something after +[def. acc.]
حذو	Ha*owa	NOUN + CASE_DEF_ACC	The position near +[def. acc.]
حول	Hawola	NOUN + CASE_DEF_ACC	The area around +[def. acc.]
حين	Hiyna	NOUN + CASE_DEF_ACC	The time/period when +[def. acc.]
خلف	xalofa	NOUN + CASE_DEF_ACC	The position behind +[def. acc.]
ضمن	Dimona	NOUN + CASE_DEF_ACC	The inside +[def. acc.]
عقب	Eaqiba	NOUN + CASE_DEF_ACC	The aftermath +[def. acc.]
عبر	Eabora	NOUN + CASE_DEF_ACC	The place across +[def. acc.]
عند	Einoda	NOUN +	The time when +[def. acc.]

		CASE_DEF_A CC	acc.]
فَور	fawora	NOUN + CASE_DEF_A CC	The time immediately after +[def. acc.]
فَوق	fawoqa	NOUN + CASE_DEF_A CC	The position above +[def. acc.]
قَبْل	qabola	NOUN + CASE_DEF_A CC	The time before/ the time prior to +[def. acc.]
قُبَيْل	qubayola	NOUN + CASE_DEF_A CC	The time before +[def. acc.]
قُبَالَة	qubAlapa	NOUN + NSUFF_FEM_ SG+CASE_DE F_ACC	The front +[def. acc.]
قُرْب	quroba	NOUN + CASE_DEF_A CC	The position near +[def. acc.]
مَعَ	maEa	NOUN + CASE_DEF_A CC	Something accompanied by/together with +[def. acc.]
	maEAF	NOUN +CASE_INDE F_ACC	Something accompanied by/together with +[indef. acc.]
أثناء	>avonA`a	NOUN + CASE_DEF_A CC	The meantime +[def. acc.]
طَوال	TiwAla	NOUN + CASE_DEF_A CC	The time during+[def. acc.]
عَوَض	EiwaDa	NOUN + CASE_DEF_A	The alternative +[def. acc.]

		CC	
حَسَبَ	Hasoba	NOUN + CASE_DEF_A ACC	The criterion depended on +[def. acc.]
حَسْب	Hasobu	NOUN + CASE_DEF_N OM	The criterion depended on +[def. nom.]
وَفَقَ	wafiq	NOUN	Conformity/An issue in conformity with
وَفَقَ	wifoqa ²¹	NOUN + CASE_DEF_A CC	Conformity/An issue in conformity with +[def. acc.]
أَمْثَال	>amovAl	NOUN	The same
ضَيْدٌ	Did~a	NOUN + CASE_DEF_A CC	Opposition/ Adversity +[def. acc.]
مِثْل	mivola	NOUN + CASE_DEF_A CC	Something like +[def. acc.]
نَحْو	naHowa	NOUN + CASE_DEF_A CC	Something towards/ Something approximately +[def. acc.]
دُون	duwna	NOUN + CASE_DEF_A CC	Something without/ Something beneath +[def. acc.]
لَدَى	ladaY	NOUN	Something with
خِلَال	xilAla	NOUN + CASE_DEF_A CC	Something during +[def. acc.]
وَرَاء	warA'a	NOUN + CASE_DEF_A CC	The position behind +[def. acc.]
حَيْثُ	HiyAla	NOUN + CASE_DEF_A CC	The standpoint concerning +[def. acc.]

جَرَاء	jar~A'a	NOUN + CASE_DEF_A CC	Something because of+[def. acc.]
وَسَط	wasata	NOUN + CASE_DEF_A CC	The position in the middle+[def. acc.]
رَغْم	ragoma	NOUN + CASE_DEF_A CC	Something despite+[def. acc.]
داخل ²²	dAxil	NOUN	Something inside+[def. acc.]
خارج ¹⁸	xArij	NOUN	Something outside+[def. acc.]
رَهْن	rahona	NOUN + CASE_DEF_A CC	Something depending on+[def. acc.]
أمس	>amosi	NOUN + CASE_DEF_G EN	Yesterday+[def. gen.]
بُعِيد	buEayoda	NOUN + CASE_DEF_A CC	A distance of +[def. acc.]
نُصَب	nuSoba	NOUN + CASE_DEF_A CC	Something in front of+[def. acc.]
قَبْد	qayoda	NOUN + CASE_DEF_A CC	Something depending on+[def. acc.]
طَيَّلَة	Tiylapa	NOUN + CASE_DEF_A CC	Something during +[def. acc.]

5.1.2 List of Former Pronouns

The entries below were previously treated as pronouns.

Arabic	Buckwalter	New POS	Gloss
--------	------------	---------	-------

		Tag	
دُو	*uw	NOUN	Possessor (Masculine)
دُوا	*awA	NOUN	Possessor
ذات	*At	NOUN	Possessor (Feminine)
ذَوِي	*away	NOUN	Possessors
ذَوِي	*awiy	NOUN	Possessors
ذاتا	*At	NOUN	Possessors
ذوات	*awAt	NOUN	Possessors
أولو/أولي	>uwluw/>uwly	NOUN	Possessors

5.1.3 List of Former Particles

The entries in this table were previously treated as particles.

Arabic	Buckwalter	New POS Tag	Gloss	Syntactic/Semantic context	Examples	Transliteration	Gloss	Translation
غير	gayor	NOUN	The other	It is used as a regular noun	جاءَ الْغَيْرُ	jA'a Algayru		
			Exception	It is used in a construct state or in an appositive.	مُحَمَّدٌ غَيْرٌ مَوْجُودٌ	muHam~adN gayru mawjuwdK		
				With an SBAR complement	غَيْرَ أَنَّ الرَّجُلَ لَمْ يَأْتِ	...gayra >an~a Alrajula lam ya'tiy		
سوى	siwaY	NOUN	Exception	It always occur with a noun or SBAR complement	جاءَ الطَّلَبُ سِوَى طَالِبٍ	jA'a AlTul~Abu siwaY TalibK		
سوى	siwaY	RESTRIC_PART	Exception	It always occur before PP	لَمْ يَحْصُلْ	lam yaHsul		

					المؤتمر سوى على القليل	Almu&tamar siwY EIY Alqalyl		
بيد	bayoda	NOUN	Exception	It has always an SBAR complement	بيد أن الرجل لم يأت	bayda >an~a Alrajula lam ya>ti		

5.2 POS Policies

5.2.1 Cardinal Numbers vs. Ordinal Numbers

Policies as well as tests with regard to distinguishing cardinal numbers from ordinal numbers are provided below.

- Digits are treated as they are pronounced. If pronounced as cardinals, they should be tagged as NOUN_NUM. If pronounced as ordinals, they should be tagged as ADJ_NUM.

Example	Transliteration	Gloss	Translation	POS Annotation
النتيجة 10-12	Alnatiyjapu 10-12	The result was 10-12	The score was 10-12	NOUN_NUM
القرن 19	Alqarn AltAsiE Ea\$ar	The century 19	The nineteenth century	ADJ_NUM

- The 10's , 100's and 1000's ... postominals are tagged as NOUN_NUM when they are preceded by noun plural.

Arabic	Buckwalter	POS Tag	Example	Transliteration	Gloss	Translation
عشرون	Ei\$rwn	NOUN_NUM	الصفحات العشرون	AlSafHAtu AlEi\$rwna	The pages the twentieth	The twentieth page
مائة	mA{ap	NOUN_NUM	السنوات المائة والخمسون	AlsanawAtu AlmA{apu wa Alxamsuwna	The year the hundredth and the fiftieth.	The hundred and fiftieth year

- Compound numerals (from the number 21 to the infinite) follow the POS tag of the first element. In case the first element is an ordinal, i.e. tagged as ADJ_NUM, whatever follows should also be tagged as ADJ_NUM.

Arabic	Buckwalter	POS Tag	Example	Transliteration	Gloss	Translation
الحادي وَ العِشْرُونَ	AlHAdiy wa AlEi\$ruwn	ADJ_NUM	الصفحة الحادية وَ العِشْرُونَ	AlSaffHapu AlHAdiyapu wa AlEi\$ruwna	The page the twenty first	The twenty first page
السادس وَ الخَمْسُونَ	AlsAdis wa Alxamsuwn	ADJ_NUM	العام السادس وَ الخَمْسُونَ	AlEAmu AlsAdisu wa Alxamsuwna	The year the sixty fifth	The sixty fifth year
الرابع وَ الثلاثون	AlrAbiE wa AlvalAvuwn	ADJ_NUM	الليلة الرابعة وَ الثلاثون	Allaylapu AlrAbiEapu wa AlvalAvuwna	The night the thirty fourth	The thirty fourth night

- When digits are interpreted and produced as ordinals, they are tagged as ADJ_NUM. Otherwise, they are tagged as NOUN_NUM.

Example	POS Tag	Transliteration	Gloss	Translation
أَنْظِرِ الصَّفحةَ 130	NOUN_NUM	>unZur AlSaffHapa 130 (AlmA{ap wa AlvalAvuwn)	See (you) the page 130	See page 130
عَلَى السَّاعةِ 18:45	ADJ_NUM	EalY AlsAEapi 18:45 AlsAdisap wa xamsap wa >arbaEiyina daqiyqap (ordinal)	At the hour 18:45 At + the hour sixth and forty five minutes	At 18:45 At six forty five o'clock
قَرَأْتُ 26 كِتَابًا	NOUN_NUM	qara>t 26/ sit~ap wa E\$riyna kitAbAF	Read (I) 26/twenty six book	I read twenty six books
فِي المَرحلةِ 16/ السَّادِسةِ عَشْرَةَ	ADJ_NUM	fīy AlmarHalap 16/ AlsAdisap Ea\$arap	In stage the sixteenth	In the sixteenth stage

5.2.2 The Pseudo-Verb <in~a (اِنْ) & the Subordinating Conjunction >an~a (اَنَّ)

The occurrence of <in~a and >an~a in the corpus presents some issues for both POS and TB annotators. The following constructions are common in the ATB corpus and should be dealt with as stated below.

5.2.2.1 In Direct Speech

When there is a clear indication of direct speech as represented by punctuation, the following constructions are annotated in the following way:

Arabic	Transliteration	Translation	Action Needed
قَالَ : " اِنْ ... "	qAla : "<in~a...".	He said : "..."	NONE
قَالَ : " اَنَّ ... "	qAla : ">an~a..."	He said : "..."	NONE

5.2.2.2 In Indirect Speech

Arabic	Transliteration	Translation	Action Needed
... قَالَ اِنَّ	qAla >an~a...	He said that...	NONE
... قَالَ اِنْ	qAla An~a...	He said that...	NONE

5.2.2.3 Default Policy

A default policy has been adopted with respect to constructions where there is no punctuation that indicates whether direct or indirect speech has been used.

In such constructions as qAla An~a.... (قَالَ اِنَّ), the annotator must annotate An~ as a subordinating conjunction by default.

5.2.3 Al- : The Determiner and Relative Pronoun

Al- has two POS values: It is a determiner and a relative pronoun. Al- is a determiner when it attaches to common nouns, proper nouns, almost all noun quantifiers expressing quantity, and adjectives. The relative pronoun Al- is a variant of the relative pronouns Al~a*y and its feminine and plural counterparts. Its use is restricted to deverbals; that is to say, maSdars and participles when these have a verbal reading. A special annotation policy has been adopted relating to the non tokenization of the relative pronoun. Al- will always be annotated as a determiner in the ATB (Refer to Section 5 on deverbals in the Syntactic Annotation Part 2 for the proposed syntactic annotation)

5.4 Annotation of laysa

The POS annotation of laysa is as follows:

lastu: las/PV+tu/PVSUFF_SUBJ:1S	not be + I
lasnA: las/PV+nA/PVSUFF_SUBJ:1P	not be + we
layosa: lays/PV+a/PVSUFF_SUBJ:3MS	not be + he/it/they
layosat: lays/PV+at/PVSUFF_SUBJ: 3FS	not be + she/it/they
lasna: las/PV+na/PVSUFF_SUBJ:3FP	not be + they [fem. pl.]
layosuwA: lays/PV+uwA: PVSUFF_SUBJ: 3MP	not be + they [masc. pl.]

5.5 Miscellenous POS values [Under Review]

This section lists tokens that are particularly confusing for POS annotators. Such tokens are part of an open-class category like noun, adjective or inflectional verbs.

Arabic	Buckwalter	POS Tag	Gloss	Examples	Transliteration	Gloss	Translation
كَيْتَ	kayota	NOUN	such/so				
أَيْمًا	>ay~mA	NOUN	such/every				
دَوَالِيكَ	dawAlayoka	dawAlayo/NOUN+ka/POSS PRON 2MS					
لَيْيِكَ	lab~ayoka	lab~ayo/NOUN+ka/POSS PRON 2MS					
وَيْكَ	wayoka	wayo/NOUN+ka/POSS PRON 2MS					
وَيْحَكَ	wayoHaka	wayoH/NOUN+ka/POSS PRON 2MS					

هداياك	hadAdayoka	hadAdayo/NOUN+ka/POSS PRON 2MS					
والسفاه	wA>asafAh ²³	wA/VOC PART+>asaf/NOUN+Ah/VOC PART					
أكمل	>akomal	NOUN	entire	عقلية جيل بأكملها	Eaqoliy~apu jiylK bi>akomalihi		
أكمل	>akomal	ADJ_COMP	Most perfect/most complete	أكمل الرجال هو علي	>akomalu AlrijAli huwa Ealiy~		
مزيد	maziyd	NOUN					
عديد	Eadiyd	ADJ	numerous				

6.0 PARTS OF SPEECH WITH MORE THAN ONE VALUE

Entries, their tags, and disambiguating context and examples resulting from different POS interpretation and annotation are provided.

Parts of speech that have more than 2 tags are represented here in separate summary charts where the different POS values for a single entry along with disambiguating contexts are provided for easy access and assessment.

6.1. >ay/>ay~ (أي/أي)

Arabic	Buckwalter	POS Tag	Gloss	Syntactic /Semantic Context	Examples	Transliteration	Gloss	Translation
أي	>ay	VOC_PART	Hey	It is used in vocative constructions	أي مُحَمَّد	>ay muHamad!	<u>Hey</u> Muhamm ad	Hey Mohamed!
أي	>ay	PART	That is to say	It provides further explanation or reformulation of a previous statement	أي أَنَّ هَذِهِ الدُول	>ay >an~a ha*ihi Alduwal	<u>That is to say</u> that these countries	That is to say that these countries
أي	>ay~	INTERROG_PR ON	What Which	In direct questions	مِنْ أَي بِلْد أَنْتَ	min >ay~ baladK >anta?	From <u>which</u> country you	From which country are you?

أَيُّ	>ay~	NOUN_QUANT	Any	It is used with one dependent complement.	لَيْسَ لَهُمَا أَيُّ إِسْتِخْدَامَاتٍ عَسْكَرِيَّةٍ وَ فِي أَيِّ مِنْ هَذِهِ الْخِلَافَاتِ لَمْ يَحْتَكِمِ أَيُّ مِنْ لُحُودٍ وَ الْحَرِيرِيِّ إِلَى مَجْلِسِ الْوُزَرَاءِ	laysa lahumA >ay~u <istixdAmAtK Easkariy~apK wa fiy >ay~K min ha*ihi AlxilAfAt lam yaHtakim >ay~N min laHud wa AlHariyriy <iLY majlisi AlwuzarA'	Not for them <u>any</u> usages military And in <u>any</u> of these disputes did not appeal any of Lahood and al-Hariri to the cabinet	They don't have any military usages And in any of these disputes neither Lahood nor al-Hariri appealed to the cabinet
أَيِّ	>ay~	REL_PRON	Whoever	It introduces indirect questions	أَيُّهُمْ قَابَلْتُ أَوْلَا أَكَلَّمُهُ فِي هَذَا الْأَمْرِ	>ay~uhum qAbaltu >aw~alAF >ukalo~imuhu fiy ha*A Al>amr	<u>Whoever</u> of them I meet first I talk him in this matter	Whoever I meet first, I will talk to him about this matter.
أَيَّا	>ay~AF (Invariable form)	REL_PRON	Whatever	Only in indefinite accusative where it can be replaced by nouns of conditions such as mahmA	أَنَّ الْأَزْمَاتِ الدُّوَلِيَّةِ وَ النِّزَاعَاتِ أَيَّا تَكُنْ دَرَجَةٌ تَعْقِيدُهَا يُمَكِّنُ أَنَّ	>an~a Al>azamAt Alduwaliy~ap wa AlnizAEAt >ay~FA takun darajapu taEqiydihA yumkinu >an		

6.2 xalA, EadA, and HASA (حاشا, عدا, حاشا)

Arabic	Buckwalter	POS Tag	Syntactic Context	Examples	Transliteration	Gloss	Translation
حاشا	HASA	VERB	S-ADV	أتى الطلبة <u>حاشا</u> زيدا	>atY AlTalabapu HASA zaydAF		
		VERB	SBAR	أتى الطلبة <u>ما حاشا</u> زيدا	>atY AlTalabapu mA HASA zaydAF		
		PREP	PP	أتى الطلبة <u>حاشا</u> زيدا	>atY AlTalabapu HASA zaydK		
خلا	xalA	VERB	S-ADV	أتى الطلبة <u>خلا</u> زيدا	>atY AlTalabapu xalA zaydAF		
		VERB	SBAR	أتى الطلبة <u>ما خلا</u> زيدا	>atY AlTalabapu mA xalA zaydAF		

		PREP	PP	أتى الطلبة <u>خلا</u> زيد	>atY AlTalabapu xalA zaydK		
عَدَا	EadA	VERB	S-ADV	أتى الطلبة <u>عدا</u> زيداً	>atY AlTalabapu EadA zaydAF		
		VERB	SBAR	أتى الطلبة <u>ما</u> <u>عدا</u> زيداً	>atY AlTalabapu mA EadA zaydAF		
		PREP	PP	أتى الطلبة <u>عدا</u> زيد	>atY AlTalabapu EadA zaydK		

N.B: خلا (xalA/become empty), عدا (EadA/run) are also regular inflectional verbs and they should be distinguished from other POS values mentioned above.

6.3 fa- (فَا-)

Arabic	Buckwalter	POS Tag	Gloss	Syntactic /Semantic Context	Examples	Transliteration	Gloss	Translation
<p>- فـ</p> <p>فاء [الجزء]</p>	<p>fa-</p> <p>[The fa- of reward]</p>	RC_PART	<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>إِنْ كَانَ لَكَ سِرٌّ فَاحْفَظْهُ</p> <p>إِنْ لَمْ تَكُنْ لَكَ رُؤْيَا فَلِمَنْ تَكُونُ</p> <p>إِنْ جِئْتَنِي فَأَنْتَ مُكْرَمٌ</p> <p>إِنْ قَبِلْتَنِي الْيَوْمَ فَقَدْ</p>	<p><in kAna laka sir~N fAHfaDhu</p> <p><in lam takun laka ru&yA faliman takwnu?</p> <p><in jY'tany faAnta mukramN</p> <p><in qabiltany Alyawma fqad qabiltuka Ams</p>	<p>If was for you secret <u>so</u> keep (I) it</p> <p>If did not was it for you vision <u>then</u> for whom be it</p> <p>If visit (you) me <u>then</u> you are treated with generosity</p> <p>If accepted (you) me today <u>then</u> accepted (I) you yesterday</p>	<p>If you have a secret so don't reveal it</p> <p>If the vision wasn't yours (then so) for whom it is?</p> <p>If you visit me (then/so) I would be generous with you</p> <p>If you accept me today (then</p>

					قَبْلَكَ أَمْسَ			so) I accepted you yesterday
[فاء فـ الربط]	fa- [The fa- of connection]	CONNEC_PART	{Discourse connective}	It connects two clauses. It is most commonly used to introduce a comment clause after a clause starting with >am~A. It also connects a fronted modifier to the rest of the clause it modifies.	أَمَّا الْعَرِيسُ فَهُوَ سَلِيلُ الْحَسَبِ وَ النَّسَبِ وَ بِحَسَبِ الْبَلْتِيَانَسْكِى، فَلَيْسَ هُنَاكَ أَيُّ مُشْكَلَةٍ لِتَمْدِيدِ الْمُبَاحَثَاتِ	>am~A AlEarysu fahuwa salylu AlHasabi w Alnasabi wa biHasabi baltiyAnskiy, falaysa hunAka >ay~u mu\$kilapK litamdiydi AlmubAHavAti	As for the groom <u>well then</u> he is good the reputation and the family And according to bAlityAnskiy, then there is not any problem in extending the talks	As for the bridegroom, he is of good reputation and family And according to bAlityAnskiy, then there is not any problem in extending the talks

<p>- فـ</p> <p>[فاء العطف]</p>	<p>fa-</p> <p>[The fa- of coordination]</p>	<p>CONJ</p>	<p>And then</p>	<p>It marks a temporal sequence between two events</p>	<p>نظرة قَابِتْسَامَةٌ هُمَا بِدَايَةُ الْقِصَّةِ</p> <p>خَرَجَ مِنْ بَيْتِهِ فَرَأَى صَدِيقَهُ عَلِيًّا فِي الشَّارِعِ فَنَادَاهُ وَدَهَبَ مَعَهُ إِلَى الْعَمَلِ</p>	<p>naZrapN faAibtisAmapN humA bidAyapu AlqiS~api</p> <p>xaraja min baytihi fara>aY Sadyqahu Ealy~A fy Al\$AriEi fanAdAhu wa *ahaba maEahu <ilaY AlEamali</p>	<p>A look <u>then</u> a smile they are the start the story</p> <p>Went out (he) from house his <u>then</u> saw friend his Ali in the street</p>	<p>A look then a smile were the beginning of the story</p> <p>He went out; then he saw his friend Ali in the street; then he shouted to him and went with him to work.</p>
<p>- فـ</p> <p>[فاء السببية]</p>	<p>fa-</p> <p>[The fa- of causality]</p>	<p>SUB_CONJ</p>	<p>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>اِنْتِنَا فَنُكْرِمُكَ</p> <p>أ هَذَا زَيْدٌ فَأَكَلَمَهُ</p> <p>لَبِيتَ الشَّبَابَ يَعُودُ يَوْمًا فَأَخْبِرُهُ بِمَا فَعَلَ الْمَشْيِيبُ</p>	<p>A'tinA fanukrimaka</p> <p>>a ha*A zaydN fa>ukal~imuhu</p> <p>layta Al\$abAba yaEwdu yawmAF fa>uxbirahu bimA faEala Alma\$iybu</p>	<p>Visit (you) us <u>so that</u> be generous (we) you</p> <p>Is this Zaid <u>so that</u> speak I him</p> <p>Wish the youth return someday <u>so that</u> tell(I) him with what did old age</p> <p>Not can (I) see you <u>since</u></p>	<p>Visit us so that we will be generous with you</p> <p>Is this Zayd so that I will talk to him</p> <p>I wish that the youth returns one day so that I will tell it what did to me the old age</p>

			Since/Because	لا أستطيع رؤيتك بوضوح فالضلام دامس	IA >astaTyEu ru&yataka biwuDwHK faAIDaIAmu dAmisN	darkness deep black	I can't see you clearly because it is too dark
--	--	--	---------------	--	---	------------------------	--

6.4 mA (ما)

Arabic	Buckwalter	POS Tag	Gloss	Syntactic /Semantic Context	Examples	Transliteration	Gloss	Translation
ما	mA	NEG_PART	Not No	It negates the meaning of the constituent it precedes.	ما من شك أن الكثير من الدول تُعاني من أزماتٍ إقتصاديةٍ	mA min \$ak~K >an~a Alkaviyra mina Alduwali tuEAniy min >azamAtK <iqtiSAdiy~apK	Not from doubt that the many from the countries suffer from crises economic	There is no doubt that many world states are suffering from economic crises.
ما [ما المصدرية]	[The gerundi ve mA]	SUB_CONJ	That	It introduces a clause that can be replaced by the masdar of the verb that follows it.	بعدما يأتي الصباح، نذهب إلى السوق بعد إتيان الصباح، نذهب إلى السوق	baEdamA ya>tiy AlSabAhu, na*habu <iLY Alsuwqi baEda <ityAni AlSabAHi, na*habu <iLY Alsuwqi	After <u>that</u> come the morning go (we) to the market After coming the morning go (we) to the market	After the sun rises, we will go to the market. After sunrise, we will go to the market.

ما [الشرطيّة]	mA [The conditio nal mA]		Unless		لن أكل ما لم تأتي	lan kula mA lam ta>tiy	I won't eat <u>unless</u> you come	I won't eat unless you come
ما	mA		As soon as/as long as	It introduces a conditional sentence	الهندُ مُسَعِدَةٌ لجوارِ ثنائيِّ ما أن يَنْتَهِى الإرهابُ عَبْرَ الحدودِ	Alhindu mustaEid~apN liHiwArK vunA}iy~K mA >an yantahiy Al<irhAbu Eabra AlHuduwdi	India is ready for a bilateral dialogue <u>as soon</u> <u>as</u> terrorism will end up on the frontiers	India is ready for a bilateral dialogue as soon as terrorism will end up on the frontiers
ما	mA	INTERROG_ PRON	What	In direct questions	ما الذي أوقَعها في ذلك المصيرِ المُخيفِ	mA Ala*iy >awqaEahA fiy *alika AlmaSiyr Almuxiyf ?	<u>What</u> which plunged her in that the fate the scary	What was it that plunged her into that fearful fate?
ما	mA	REL_PRON	What	It introduces indirect questions	وقَع ما نخشاهُ	waqaEa mA nax\$Ahu	Happened <u>what</u> we fear	What we fear happened

ما [ما التعجبية]	mA [The mA of Exclamation]	EXCLAM_PR ON	What How	Exclamative structure	! ما أحسن السماء	mA >aHsana Alsama'a!	What a beautiful the sky!	What a beautiful sky!
ما	mA	NOUN	Some		سنلتقي يوماً ما	snaltaqiy yawmAF mA		We will meet one day
ما المُسَبِّحَةُ [ليس]	mA [mA acting lika laysa]	VERB	Not		ما الحسنُ في وجه القتي شرفاً له	mA AlHusnu fiy wajhi AlfatY \$arafAF lahu		Beauty for a man is not an honor
ما	mA	PART	{Particle}	It comes after an adjective with an adverbial meaning	السُّلطاتُ التَّقْدِيَّةُ غَالِباً ما تَعْتَمِدُ أَسْعَارَ فَائِدَةٍ مُرْتَفِعَةٍ	AlsuluTAt Alnaqdiy~ap gAlibAF mA taEtamidu >asEAra fA}idap murtafiEap	The monetary authorities relies mostly on prices that have high interests	The monetary authorities relies mostly on prices that have high interests

6.5 wa- (وَ)

Arabic	Buckwalter	POS Tag	Gloss	Syntactic /Semantic Context	Examples	Transliteration	Gloss	Translation
وَ [واو القسم]	wa- [The wa- of oath]	PREP	By <By God!>	It is used in oaths	وَ اللّٰهُ	wa All~ahi	By God	By God!

وـ	wa-	PREP	Together with Along with	It expresses the meaning of accompaniment	جاء الطلبة و زيدا	jA'a AlTalabapu wa zaydAF	Came the students <u>with</u> Zaid	The students came along with Zaid
[واو المعية]	[The wa- of accompaniment]							
و-	wa-	CONJ	And	It coordinates two constituents	الضوء ناصع و الجو حار	AlDaw'u nASiEN wa Aljaw~u Har~N	The light bright <u>and</u> the weather hot	The light is bright and the weather is hot
[الربط]	[The wa- of connection]							
و-	wa-	SUB_CONJ	{Subordinating conjunction}	It introduces a clause describing the manner with which an event took place	و لكني رغم ذلك احييها و انا واثق انه لن يغضب	wa lakin~iy ragma *alika >aHkiyhA wa >anA wAviqN >an~ahu lan yagDaba	And but I am despite that tell it <u>with</u> confidence that he is will not be angry	But in spite of that I told [the joke] being confident that he would not get angry this time
[واو الحال]	[The wa- of manner]							
و (واو الزائدة)	wa- [The extra wa]	PART		It is a constituent particle.	سبق و رأيت ذلك من قبل	sabaqa wa ra>ayotu *'lika min qablu		I saw that before

6.6 Hat~Y (حتى)

Arabic	Buckwalter	POS Tag	Gloss	Syntactic /Semantic Context	Examples	Transliteration	Gloss	Translation
حتى	Hat~Y	PART	Even	It has an emphatic meaning of inclusion	لم يملك حتى فداناً واحداً	lam yamluk Hat~Y fad~AnAF wAHidAF	Not own <u>even</u> acre single	He does not own even one single acre
حتى	Hat~Y	CONJ	And	It coordinates two constituents with the same function	أكلت السمكة حتى رأسها	>akaltu Alsamakap Hat~Y ra>sahA	Ate I the fish <u>and</u> head its	I ate the fish along with/and its head.
حتى	Hat~aY	PREP	Until/up to	It indicates the end point of an action or	أكلت السمكة حتى رأسها	>akaltu Alsamakapa Hat~aY ra>sihA	Ate (I) the fish <u>up to</u> head her	I ate the fish up to its head

				its extent.				
--	--	--	--	-------------	--	--	--	--

6.7 IA (لا)

Arabic	Buckwalter	POS Tag	Gloss	Syntactic /Semantic Context	Examples	Transliteration	Gloss	Translation
لا [لا النافية للجنس]	IA	PSEUDO_VERB	{PSEUDO_VERB}	It negates an equational sentence.	لا فائدة كبيرة	IA fA}idapa kabiyrap	No benefit great	There is no great use in it
لا لا...لا	IA IA...IA ...	NEG_PART	Not Neither..nor...	It negates the meaning expressed by a verb in the present It negates persons or things. There must be two negated elements for IA to have the value of a negative particle in this context.	هَذَا النَّوْعُ مِنَ الحشائش لا يَنمو في المناطق الحارة رَأَيْتُ لا زَيْداً وَلا عَمراً	ha*A AlnawEu mina AlHa\$A}iSi IA yanmwA fiy AlmanATiqi AlHaA r~api ra>aytu IA zaydAF wa IA EamrAF	This the type of the herbs <u>not</u> grow in the regions the hot I saw <u>neither</u> zayd <u>nor</u> Eamr	This type of grass does not grow in hot regions I saw neither zayd nor Eamr
لا [العطف]	IA [The IA of coordinat ion]	CONJ	And not	It corroborates an assertion that has been previously made	رَأَيْتُ زَيْداً لا عَمراً	ra>aytu zaydAF IA EamrAF	Saw (I) Zaid <u>not</u> Amr	I saw Zayd not Amr
لا	IA	INTERJ	No!					

6.8 The Particle and the Preposition لا <il~A

6.8.1 <il~A as a Restrictive Particle

<il~A as a restrictive particle in a noun phrase occurs in negative constructions. In other words, the clause where <il~A occurs in is usually negated by means of a negative particle. It is interpreted as an apposition to the noun it follows. So the phrase starting with <il~A can substitute and be substituted by the noun it is in apposition with. The case of the noun in the <il~A phrase agrees with the case of the noun it is in apposition with.

In all the examples below, one can substitute أَحَدٌ (>aHad/nobody) with the phrase إِلَّا (<il~A zayd/except for zayd) and vice versa.

Examples	Transliteration	Gloss
مَا رَأَيْتُ أَحَدًا إِلَّا زَيْدًا	mA ra>aytu >aHadAF <il~A zaydAF	Not I saw nobody except for zayd
مَا مَرَرْتُ بِأَحَدٍ إِلَّا زَيْدٍ	mA marartu bi>aHadK <il~A zaydK	Not I passed by nobody except for zayd
مَا جَاءَنِي أَحَدٌ إِلَّا زَيْدٌ	mA jA'aniy >aHadN <il~A zaydN	Not came nobody except for zayd

6.8.2 <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.

Example	Transliteration	Gloss
جَاءَنِي قَوْمُكَ إِلَّا قَلِيلًا	jA'aniy qawmuka <il~A qaliylAF	Came your citizens except a few

Note here that the noun phrase constituent that starts with <il~A cannot occur in apposition with the other NP; that is to say, the phrase starting with <il~A cannot substitute for the noun phrase qawmuka (your citizens).

6.9 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 (يَوْمَان). 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.

6.10 The preposition EalY

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.

6.11 The Particle and Preposition la-

Arabic	Buckwalter	POS Tag	Gloss	Syntactic context	Examples	Transliteration	Gloss	Translation
لـ -	li- la- (with bound pronouns)	PREP	To For	Always governs a nominal constituent (noun, for example) or an SBAR.	لِيَحْدِثِ الْمَوْقِفَ	libaHvi Almawqifi.	To examine the situation	In order to examine 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.	لِنَذْهَبْ	lina*hab	Lets go	Lets go

ا -	la- (invariable)	EMPHATIC_PART	Definetly Certainly	Always preceding a verb and indicates reinforcement of the action the verb is describing	لأَصْرَبْتُكَ	la>aDraban~a ka	most certainly will beat (I) you	I most certainly will beat you
ا -	la- (invariable)	RC_PART	So	It is used in conditional sentences introducing the apodosis sentence/main clause	إِنْ أَعْطَيْتَ لَأَكْرَمْتُكَ	<in >aETayta la>akramtuka	If give (you) for sure honor (I) you	If you give, I would honour you

7.0 INTERACTION BETWEEN POS AND TREEBANK

7.1 Adjectives and Nouns [The following tests are still under experimentation and will be reviewed on a regular basis with more information from new annotation production]

Due to misconceptions and issues related to the annotation of nouns and adjectives in PATB, this section provides tests to check whether a particular token is a noun or an adjective. The issue of homographs, entries that have the same spelling but have distinct meanings, will also be covered.

7.1.1 Insertion of a Missing Lexical Head

The test is relevant for substantive adjectives; that is to say, adjectives that substitute nouns in context. Nominal heads of syntactic positions such as subject, object, or complement can be missing and adjectives will function as heads in such null-headed NPs. The semantic head must be recoverable from the context of annotation.

7.1.1.1 The Case of Relational Adjectives

The token “عَرَبٌ /Earab” is the plural form of the adjective عَرَبِيٌّ/Earabiy~. It is annotated as an adjective in POS. Given the POS value, the constructions below will be analyzed without any reconsideration of the POS value of the token “عَرَبٌ /Earab” due to syntactic function.

1. جاءَ عَرَبُ الجَنُوبِ (jA’a Earabu Aljanuwbi/the Arabs of the south)
2. جاءَ الكُتَّابُ العَرَبُ (jA’a Alkut~Abu AlEarabu/the Arab writers)
3. جاءَ العَرَبُ المَسِيحِيُّونَ (jA’a AlEarabu AlmasiyHiy~uwna/the Christian Arabs came)
4. لِسَانُ العَرَبِ (lisAnu AlEarabi/The language of the Arabs)

In example 1, Earab is an adjective, head of a non-true construct state (إضافة غير حَقِيقِيَّة) /<iDAfap gayr Haqyqiy~ap) construction, عَرَبُ الجَنُوبِ /Earabu Aljanuwbi a type of complementation structure headed by an adjective.

(S (VP jA’a (NP-SBJ Earabu (NP Aljanuwbi)))

In example 2, Earab is an adjective; therefore, it modifies the noun كُتَّابُ /kut~Ab

(S (VP jA’a (NP-SBJ Alkut~Abu AlEarabu)))

In example 3, Earab is an adjective that modifies, along with the adjective masyHiy~uwn, a missing head noun.

(S (VP jA’a (NP-SBJ AlEarabu AlmasiyHiy~yuwna))).

A test that can be used to recover the missing noun is to insert a lexical head before the adjectives. So we can say جاءَ الرجالُ العَرَبُ المَسِيحِيُّونَ (jA’a AlrijAlu AlEarabu AlmasiyHiy~uwna/Came the men the Arabs the Christians).

In example 4, Earab is an adjective that is a complement of a noun which is the head of a true construct-state construction (إضافة حَقِيقِيَّة) /<iDAfap Haqyqiy~ap).

(NP lisAnu (NP AlEarabi))

Here again AlEarabi is used to modify a missing head noun that can be inserted before the adjective. So we can recover the missing head noun in the following sentence by saying لسانُ القومِ العَرَبِ /lisAnu Alqawmi AlEarabi.

7.1.1.2 The Case of Active and Passive Participles

القاطِنُ (AlqATinu/ resident) and المَقْهُورُ (Almaqhuwru/the oppressed) are adjectives in POS even though in TB, they will be under an NP node in both 1 (a) and 2 (a).

1. (a) جاءَ القاطِنُ في القاهرةِ (jA'a AlqATinu fiy AlqAhirapi/came the resident in Cairo)
(S (VP jA'a (NP-SBJ (NP adj AlqATinu (PP fiy NP AlqAhirapi))))

1. (b) جاءَ الرَّجُلُ القاطِنُ في القاهرةِ (jA'a Alrajulu AlqATinu fiy AlqAhirapi/came the resident in Cairo)
(S (VP jA'a (NP-SBJ (NP Alrajulu (ADJP AlqATinu (PP fiy (NP AlqAhirapi))))))

2. (a) جاءَ المَقْهُورُ (jA'a Almaqhuwru/came the oppressed)
(S (VP jA'a (NP-SBJ adj Almaqhuwru)))

2. (b) جاءَ الرَّجُلُ المَقْهُورُ (jA'a Alrajulu Almaqhuwru/came the oppressed man)
(S (VP jA'a (NP-SBJ Alrajulu Almaqhuwru)))

7.1.1.3 The Case of إسم الصِّفَة (<ism AlSifap)

الْفَرِحُ (AlfariHu/the happy) and الصَّعْبُ (AlSaEbu/the difficult) are adjectives in POS even though in TB, they will be under an NP node in both 1 (a) and 2 (a).

1. (a) جاءَ الفَرِحُ (jA'a AlfariHu/ came the happy)
(S (VP jA'a (NP-SBJ adj AlfariHu)))

1. (b) جاءَ الرَّجُلُ الفَرِحُ (jA'a Alrajulu AlfariHu/came the happy man)

(S (VP jA'a (NP-SBJ Alrajulu AlfariHu)))

2. (a) الصَّعْبُ هُوَ أَن تُبَاشِرَ بِالْعَمَلِ بَعْدَ هَذَا الحَاطِثِ (AlSaEbu huwa >an tubA\$ira biAlEamali baEda ha*A AlHadivi/the difficult is that you resume work after this accident)

(S (NP-SBJ (NP adj AlSaEbu) (NP huwa) (SBAR >an (S (VP tubA\$ira (NP-SBJ *)...))

2. (b) الشَّيْءُ الصَّعْبُ هُوَ أَن تُبَاشِرَ بِالْعَمَلِ بَعْدَ هَذَا الحَاطِثِ (Al\$ay}u AlSaEbu huwa >an tuba\$ira biAlEamali baEda ha*A AlHadivi/the thing difficult is that you resume work after this accident)

(S (NP-SBJ Al\$ay{u AlSaEbu) (NP huwa) (SBAR >an (S (VP tubA\$ira (NP-SBJ *)...))

7.1.2 Homographs: Semantic Tests

Reference to physical entities, concepts, occupations is a property of nouns. Adjectives refer to attributes and qualities that nouns have. They are used to describe an aspect, part, or characteristic of the noun. Semantic tests are particularly useful in the case of homographs where two identical morphological forms have distinct meanings. Such forms should be located under two different entries.

7.1.2.1 The Case of Manufactured Gerunds

يَهُودِيَّة (yahuwdiy~ap/Judaism) will be taken as an example to illustrate the role of semantic tests when it comes to certain POS decisions.

If the token refers to religious affiliation as in its English counterpart “Judaism”, then the token should be tagged as a noun as in example 1. If it is used as a modifier of a noun, then the token should be annotated as an adjective.

E.g.1 اليَهُودِيَّةُ سَمَويَّةٌ (Alyahuwdiy~apu diyAnapN samAwiy~apN/Judaism is a religion divine) **and** اليَهُودِيَّةُ مَذَهَبٌ دِينِيٌّ (Alyahuwdiy~apu ma*habN diyniy~N/Judaism is a religious affiliation)

(S (NP-SBJ noun Alyahuwdiy~ap) (NP-PRD diyAnapN samAwiy~apN))

E.g. 2 المرأة اليهودية (Almar>apu Alyahuwdiy~apu/the woman the Jew) vs. الرجل اليهودي (Alrajulu Alyahuwdiy/ The man the Jew)

(NP-SBJ Almar>ap Alyahuwdiy~ap adj)

In example 2, the second token agrees in number and gender with the noun it modifies. It is therefore, an adjective.

Remark:

Annotators should pay attention when using morphological operations in the derivation of إسم النسبة (<ism Alnisbap) (refer to Section 4.3.1.3.1.3). An issue with the token إقتصاديات /<iqtiSAdiy~At was raised in one of the experimental tests we have been assigning for ATB annotators. The word was mistaken to be the plural form of the adjective إقتصادي (<iqtiSAdiy/economic) whereas it is the plural form of the noun إقتصاد (<iqtiSAd/economy). The fact that إقتصاديات (<iqtiSAdiy~At/economies) in its context referred to the concept of economies was a useful test to determine its POS value.

7.1.2.2 The Case of صفة /Sifap

In 1 (a) “karym” is an adjective whereas in 1 (b) it is a proper noun. Both will be annotated under an NP node.

1. (a) جاء الكريم (jA’a Alkarym/came the generous)
1. (b) جاء كريم (jA’a Karym/came Karym)

7.1.2.3 The Case of Active and Passive Participles

كافة /kAf~ap has two POS values: Noun quantifier in example 1, though under an NP-ADV node in TB, and an active participle ; i.e. adjective in example 2 and 3.

- 1: الناس كافة مجتمعون (Aln~Asu kAf~apF mujtamiEuna/the people all are meeting)
2. الرجل كفاف عن العمل (Alrajulu kAf~N Ean AlEamali /the man stopping from work)
3. الما الكافة عن العمل (AlmA AlkAf~ap Ean AlEamal /The mA that ceases to have a syntactic function)

Difference in meaning is one test to differentiate between the two forms. As a noun quantifier, كَافَّةٌ /kAf~ap has the meaning of “all”. As an active participle, the meaning is “ceasing to do something”. Morphologically, كَافَّةٌ /kAf~ap as a noun quantifier is one morpheme annotated as NOUN_QUANT. كَافَّةٌ (kAf~ap/ceasing) as an active participle is annotated as ADJ+NSUFF_MASC_SG in example 2 and ADJ+NSUFF_FEM_SG in example 3. In both examples 2 and 3, the adjective agrees with the noun in gender and number. In other words, the noun that precedes the adjectives controls agreement on its modifier.

7.1.3 Morphological Tests

Adjectives in Arabic are inflected for case, number, and gender to agree with the noun they modify or that they stand in a predicate relationship with them.

A test annotators can use is to alter a morphological feature and see its impact on the relevant token.

E.g 1. الرِّجَالُ العَرَبُ /AlrijAlu AlEarabu /The men the Arabs

E.g. 2. الرَّجُلُ العَرَبِيُّ /Alrajulu AlEarabiy~u / The man the Arab

Another test will be to place the ambiguous token in an equational sentence. If a feature in the token changes to agree with a parallel change in the same feature in the noun, then that token is an adjective.

E.g. 3. المَرَأَةُ المُسْتَحِيلَةُ /Almar>apu AlmustaHylapu/the woman the impossible (singular)

E.g. 4. النِّسَاءُ المُسْتَحِيلَاتُ /AlnisA'u AlmustaHiylAtu/the women the impossible (plural)

7.1.4 Adjectives in Adverbial Positions

Both adjectives and nouns can function as adverbials in context. The list below represents the most commonly mistaken tokens in terms of their POS core value, noun or adjective. Note that this not an exhaustive list. Most of those tokens are annotated under an NP-ADV node in TB. Others were just noticed to be problematic for annotators. For many of the entries below, the determination of the POS can be achieved by moving back to the form from which the token was “derived”. E.g. تَدْرِيجِيًّا (tadriyjii~AF/gradually) is the adverbial form of the adjective تَدْرِيجِي (tadriyjii/gradual).

7.1.4.1 Nouns

جَدًّا (jid~AF/much), جَمًّا (jam~AF/much), حَقًّا (Haq~AF/really), خُصُوصًا (xuSuwSAF/particularly), أَحَدٌ (>aHad/one person), أَسَاسًا (>asAsAF/basically), تَمَامًا (tamAmAF/exactly), حَتْمًا (HatmAF/sure), أَكْثَرِيَّةً (>akvariyy~ap/majority), أَغْلَبِيَّةً (>aglabiyy~ap/majority), مَزِيدٍ (maziyd/addition, excess)

7.1.4.2 Adjectives

خَاصَّةً (xAS~apF/especially), كَثِيرًا (kaviyrAF/much), قَلِيلًا (qaliylAF/few), آخِرَ (|xir/last), آخِرًا (>axiyrAF/lastly), حَالِيًّا (Haliy~AF/currently), دَائِمًا (dA}imAF/always), عَامَّةً (EAm~apF/generally), مُخْتَلِفٌ (muxtalif/various), تَدْرِيجِيًّا (tadriyjiy~AF/gradually), شَتَّى (Šat~Y/various), عِدَّةً (Eid~ap/many), عَدِيدٍ (Eadiyd/many) (when there is an agreement with the noun it modifies. E.g. أَجْوَاءٌ عَدِيدَةٌ />ajwA' Eadiydap/many atmospheres).

7.1.5 Adjectival vs. Nominal Modification

Both nouns and adjectives are used as modifiers of nouns. Modification as a syntactic function cannot; therefore, be used to differentiate both POS. Noun modification; however, is restricted to particular cases that can be pinpointed; and therefore, distinguished from cases of adjectival modification.

Nominal modification is usually an apposition type of construction and is represented syntactically as adjunction. Nouns that can act as modifiers of other nouns are noun quantifiers, numerical nouns, unbound personal pronouns, nouns of possession or one of their synonyms, and maSdars.

7.1.5.1 Verbal Corroboration/Lexical Corroboration (توكيد لفظي)

“The NP adjunct in this construction should be an emphatic repetition of the NP head or an unbound pronoun if the head NP is a clitic pronoun. In the case of verbal corroboration, or it can be a repetition of the word along with different or additional modification.” (PATB, p.30)

E.g. 1 : كَانَتْ تِلْكَ مُفَاجَأَةً رَهِيْبَةً مُفَاجَأَةً هَزَّتَ الْجَمِيْعَ : (kAnat tilka mufAja>apA rahybapA mufAja>apA haz~at AljamyE/that was a horrible surprise a surprise that shook all)

E.g. 2 أَدْعُوكَ أَنْتَ (>adEuka >anta/I summon you: you this is more of a gloss than a free translation)

7.1.5.2 Paraphrased/Periphrastic Corroboration توكيد معنوي

The same principle as verbal corroboration applies here except that other nouns not a repetition of the head itself are used as emphasers. “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” (PATB, p. 31). A noun without a pronominal complement can also act as an emphaser.

E.g. 1. الْأَسُّ كُلُّهُم (Aln~Asu kul~uhum/the people all of them)

E.g. 2. إِنَّ الشَّعْبَ اللَّيْبِيَّ وَالْعَالَمَ الْعَرَبِيَّ أَجْمَعُ مُسْتَأْسَاءٌ (<in~a Al\$aEba Allybiy~a wa AlEAlama AlEarabiy~a >ajmaEa mustA’N/The Libyans and all the Arab world are upset)

Noun emphasers are usually noun quantifiers, nouns of possession such as ذات (*At/self) or one of their synonyms such as نفس (nafs/self) (PATB, p.32).

7.1.5.3 Modification with Numerical Nouns

The numerical noun phrase that is modifying a head noun phrase is adjoined to it.

E.g. 1. قَرَأْتُ الْكِتَابَ عَشْرِينَ صَفْحَةً مِنْهُ (qara>tu AlkitAba Ei\$riyna SaffHapAF minhu/I read the book, twenty pages of it)

Refer also to Section 3.3.2.3.4 of the syntactic annotation guidelines for more details on the syntactic annotation of such tokens.

7.1.5.4 Modification with maSdars

E.g. الرَّجُلُ الشَّهَامَةُ (Alrajulu Al\$ahAmapu/the man the courage)

7.1.5.5 Special cases

- Titles are interpreted as part of the proper nouns. They are annotated flat. E.g. الْأُسْتَاذُ مُحَمَّدٌ (Al>ustA* muHam~ad/the teacher muHam~d). Note that مُحَمَّدُ الْأُسْتَاذُ (muHam~ad Al>ustA*/muHam~ad the teacher) is annotated as apposition.

- Cases like **الفلسطينيُّ مُحَمَّدٌ** (AlfilasTyniy~ muHam~ad/the palestenian muHam~ad) will be annotated as apposition even though the first token is an adjective. We assume that the latter modifies a missing head noun.

7.2 Trees for POS with more than one POS value

7.2.1 fa- /فـ ‘and, so, then. so that’

‘fa-’ has 4 different POS values and 3 different TB annotations based on these POS values:

7.2.1.1 The conjunction fa- (فَاء العطف / fA’ AIEatf) ‘and/so’ (CONJ)

fa- as a conjunction is almost an equivalent of the conjunction wa- with the difference that ‘fa’ carries an additional meaning of temporality, simultaneity or explanation.

Below are the syntactic contexts where fa- is annotated as a conjunction:

- Starting a sentence where fa- is in an initial position just like wa-

```
(S (NP-SBJ تلك::tiloka::that/those)
  (NP-PRD ظاهرة::ZAhirapN::phenomenon+[fem.sg.]
   طبيعية::TabiyEiy~apN::natural/normal)
. )

(S ف::fa-::and/so
  (NP-TPC-1 اليأس::-Al+ya>osu::the+despair/hopelessness
   الكبير::Al+kabiyru::the+large/great/important)
  (VP يفضي::yu+foDiy+null::he/it+inform/notify/lead/take
   (NP-SBJ-1 *T*)
   (PP-CLR الى::<ilaY::to/towards
    (NP اليأس::Al+ya>osi::the+despair/hopelessness
     المطلق::Al+muTolaqi::the+absolute/unlimited))))
```

تِلْكَ ظَاهِرَةٌ طَبِيعِيَّةٌ فَالْيَأْسُ الْكَبِيرُ يُفْضِي إِلَى الْيَأْسِ الْمَطْلَقِ

tilka + ZAhirapN + TabyEiy~apN. fa+Alya>osu + Alkabiyr + yufDiy + <iIY + Alya>si AlmuTlaqi
That is a natural phenomenon. And big desperation leads to absolute desperation.

- Coordinating 2 constituents with the same syntactic function and category

(S (S (VP كان::kAn+a::be/was/were+he/it
(NP-1 البطريرك::Al+baToriyaroku::the+Patriarch)
(VP استقبل::isotaqobal::receive/greet/welcome+he/it
(NP-SBJ-1 *T*)
(NP-OBJ (NP (NP (NP البروفسور::Al+bruwfisuw::the+professor)
(PP في::fiy::in
(NP (NP جامعة::jAmiEapi::university)
(NP السوربون::Al+swrbwn))))
(NP دومينيك::duwmiyniyk::Dominik
شوفالييه::\$wfaIyyh::NOT_IN_LEXICON))
ف::fa-::and/so
(NP الشيخ::-Al+\$ayoxa::the+Sheikh/chieftain
وديع::wadiyE::Wadie
الخازن::Al+xAzn::the+NOT_IN_LEXICON)
ف::fa-::and/so
(NP المحامية::-Al+muHamiyapa::the+lawyer/defense_counsel
بهية::bhyp::NOT_IN_LEXICON
أبو::>abuw::Abu
حمد::Hamod::Hamd))))))

وكان البطريرك استقبل البروفسور في جامعة السوربون دومينيك شوفالييه فالشيخ وديع الخازن فالمحامية بهية أبو حمد

wa+ kAna + AlbaTriyrk + <istaqobala + Albruwfiswur + fiy + jAmiEapi + Alsurobwun duwmiyniyk + \$uwfaIyyih +
fa+Al\$ayokh + wadiyE + Alxaz~An + fa+AlmuHamiyapi bahy~ap + Abuw + Hamd
*And the patriarch welcomed the professor in La Sorbonne Dominique Chevalier, and then the sheikh Wadi' Alkhazzen and then
the the lawer bahya Hamd.*

(S (S (VP كان::kAn+a::be/was/were+he/it
(NP-1 الناخبون::Al+nAxibuwna::the+voter/elector)
(VP ينظرون::ya+noZur+uwna::they_[people]+look/observe/see

فور السماح لهم بعبور بدأ الاجئون يعبرون الضفة حشودا كبيرة فبأعداد أقل

Fawora + AlsamAhi + la+hum + bi+Eubuwri + AlHudwdi + bada>a + All~Aji}wna + yaEoburwuna + AlDif~apa + Hu\$wdAF + kabiyrapF + fa + bi+>aEodAdK >aqal~a

Immediately after they were allowed to pass the borders, the refugees started crossing the bank (in) big crowds then in smaller numbers.

7.2.1.2 The particle fa-

There are two POS annotations of fa- as a particle:

- fa- (فاء الرّبط /fa' AlrabT) which is annotated as a connective particle (CONNEC_PART)
- and fa- (فاء الجزاء / fa' AljazA') which is annotated as a response conditional particle (RC_PART).

Both of them are given the node label 'PRT' in the tree, and neither ever occurs between two syntactic elements of the same function and category (compare to fa- as CONJ).

7.2.1.2.1 fa- as a connective particle 'and/so' (CONNEC_PART)

fa- is a connective particle when it plays the role of a discourse connector. This means that it links, not coordinates, two discourse elements or constituents together. These discourse elements do not have the same syntactic function and are not necessarily of the same syntactic category.

fa- as a connective particle connects, or links, a pre-mentioned constituent (NP, PP...) of a given sentence with the rest of the sentence.

- A constituent introduced by the focus particle أمّا (>am~A). Usually, the constituent will be a topicalized NP.

```
(S (PRT اما::>am~A::as_for/concerning)
(NP-TPC-1 السلطة::Al+suloTapu::the+power/authority/rule)
::,
(PRT ف::fa-::and/so)
(VP ليست::-layos+at::not_be+it/they/she_[verb])
```

(NP-SBJ-1 *T*)
 (ADJP-PRD مسألة::msAlmp::nogloss))

أما السُّلطة فليست مُسالمة

>am~A + AlsulTapu + fa+layosat + musAlimapF.
As for the government, then it was not pacifist.

- A fronted constituent, whether a PP or an NP-ADV that functions as a modifier of the verb.

(S (PP-PRP لـ:: li::for
 (NP ذلك::*\`lika::that))

'
 (PRT ف::fa-::and/so)
 (VP ان::-<in~a::indeed
 (NP-SBJ (NP حل::Hal~a::solution/dissolution/cancellation
 (NP القضية::Al+qaDiy~api::the+problem/issue
 الفلسطينية::Al+filasoTiyiny~api::the+Palestinian))
 (NP هو::huwa::it/he))
 (NP-PRD (NP المدخل::Al+madoxalu::the+entrance/introduction)
 (PP لـ::li-::for/to
 (NP حل::-<Hal~i::solution/dissolution/cancellation
 (NP القضايا::Al+qaDayA::the+problems/issues
 الاخرى::Al+>uxoraY::the+other/additional))))))

لِذَلِكَ فَإِنَّ حَلَّ الْقَضِيَّةِ الْفِلِسْطِينِيَّةِ هُوَ الْمَدْخَلُ لِحَلِّ الْقَضَايَا الْآخَرَى

li+*alika + fa+<in~a + Hal~a + AlqaDiy~api + AlfilisTiyiny~api + huwa + Almadxalu li+Hal~i + AlqaDayA + Al>uxraY
For that then solving the Palestenian problem is the introduction for solving other problems.

7.2.1.2.2 fa- as a response conditional particle 'then' (RC_PART)

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.

(S (SBAR-ADV ان::-<in::if/whether
 (S (S (VP (PRT لم::lam::did_not
 يجمع::yu+jomaE+o::he/it+be_agreed_unanimously+[jus.]

(NP-SBJ العرب ::Al+Earabu::the+Arabs)))
 و ::wa-::and
 (S (VP (PRT لم ::-lam::did_not)
 يوحدوا ::yu+waH~id+uwA::they+unite/unify
 (NP-SBJ *)
 (NP-OBJ كلمت ::kalim+at+a-::word/remark/speech
 (NP هم ::-hum::their)))

'
 (PRT ف ::fa-::and/so)
 (NP-TPC-2 القضية ::-Al+qaDiy~apu::the+problem/issue
 الفلسطينية ::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 كَلِمَتَهُمُ فَالْقَضِيَّةُ الْفِلِسْطِينِيَّةُ لَهَا عِدَّةٌ أَوْجُهٍ

<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.

7.2.1.3 The subordinating conjunction fa- (فاء السببية /fa' Alsababy~ap) 'so that' (SUB_CONJ)

fa- is a subordinating conjunction annotated as SUB_CONJ when it introduces a subordinating clause that contains the result or cause of an event stated in the main clause. Below are some additional features that characterize the syntactic context where fa- as a subordinating conjunction occurs:

- The verb of the clause introduced by 'fa' is in the subjunctive.
- This 'fa' could be replaced by likayo, kayo, or li>an~a.

(S (VP بحثنا ::-baHav+nA::discuss/search_[for]+we_[verb]
 (NP-SBJ *)
 (PP-CLR في ::fiy::in
 (NP اختصار ::{ixotiSAr+i::abbreviation/shortening
 (NP (NP الممرات ::Al+mamar~Ati::the+passage/corridor

الجوية::Al+jaw~iy~api::the+air/aerial)
 (NP-LOC بين::bayona::between/among
 (NP (NP بلدان::bulodAn+i-::countries
 (NPنا::-nA::our)
 (ADJP الثلاثة::Al+valAv+ap+i::
 the+three))))))
 (PP-PRP ل::li-::for
 (SBAR 0
 (S (VP تشمل::-ta+\$omal+a::it/them/her_to+comprise
 (NP-SBJ *)
 (NP-OBJ (NP الخليج::Al+xaliyja::the+gulf
 العربي::Al+Earabiy~a::the+Arab/Arabic)
(SBAR-PRP ف::fa-::and/so
 (S (VP تكون::-ta+kuwn+a::it/they/she
 (NP-SBJ أجواؤ::>ajowA&+u-
 ::air/atmosphere
 (NP نا::-nA::our))
 (ADJP-PRD جاذبة::jA*ib+apF::
 attractive/alluring
 (PP ل::li-::for/to
 (NP مختلف::-muxotalafi::
 disputed
 (NP شركات::\$arikAti::
 companies
 (NP الطيران::
 Al+TayarAni::
 the+airline
))))))))))))

بَحَثْنَا فِي إِخْتِصَارِ الْمَمَرَاتِ الْجَوِيَّةِ بَيْنَ بُلْدَانِنَا الثَّلَاثَةِ لِتَشْمَلَ الْخَلِيجَ الْعَرَبِيَّ فَتَكُونُ أَجْوَاؤُنَا جَاذِبَةً لِمُخْتَلَفِ شَرَكَاتِ الطَّيْرَانِ

baHavnA + fiy + <ixtiSArK + Almamar~Ati + Aljawiy~apa + bayna + buldAninA + AlvalAvap + li+ta\$mala + Alxaliyja +
 AlEarabiy~a fa+takuwna + >jawA&unA + jA*ibapF + li+muxotalafi + \$arikAti + AlTayarAni.
*We studied in brief the air passages between our three countries for them to comprise the Arabic Gulf and so our location will
 attract different airline companies.*

7.2.2 <il~A/ لا/’except’, ‘except for’

<il~A has 2 different POS values with two matching syntactic annotations. <il~A is a restrictive particle annotated as RESTRIC_PART and a preposition annotated as PREP.

7.2.2.1 The restrictive particle <il~A (RESTRIC_PART)

<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*.

<il~A as particle assigns no case ending. The constituent modified by <il~A gets its case ending according to its syntactic role and/or position in the sentence:

- when <il~A modifies a constituent that has a specific syntactic function in context and **is not in apposition with a previous one**, that constituent takes its case ending based on its function in the sentence (see all 3 examples below).

with *layosa*:

```
(S (VP ليس::layos+a::not+[it]
  (NP-PRD امام::>amAma::in_front_of/facing/vis--vis
    (NP الاسرائيليين::Al<isorA}iyliy~iyana::the+Israeli))
  (NP-SBJ (NP (PRT لا::<il~A::however/except/other)
    الرضوخ::Al+ruDuwx+u::the+submission/compliance+[def.nom.]
    (PP ل::li-::to/for
      (NP (NP لمعادلات::-Al+muEAdalAti::the+equalizing/balancing)
        (SBAR (WHNP-2 التي::Al~atiy::which/who/whom)
          (S (VP تكرر::tu+kar~is+u-::it/they/she+dedicate/devote+[ind.]
            (NP-OBJ ها::-hA::it/them/her)
            (NP-2 (-NONE- *T*)))
          (NP-SBJ المقاومة::Al+muqAwamapu::the+resistance/opposition)))))))))
```

ليسَ أمامَ الإِسْرَائِيلِيِّينَ إلَّا الرُّضُوخَ لِلمُعَادَلَاتِ الَّتِي تُكْرَسُهَا المُقَاوِمَةُ

layosa + >amAma + Al<isrA}iyly~ina + <il~A + AlruDuwxa + li+AlmuEAdalAti + Allatiy + tukar~isuhA + AlmuqAwamapu
 There is nothing else for the Israeli to do except for the compliance to the equations that the resistance is submitting.

With lan:

(S (NP-TPC-4 سَرَاب::sarAb+a::mirage+[def.acc.]
(NP التَّفَاوُض::Al+tafAwuDi::the+negotiation/consultation))
(VP (PRT لَنْ::lan::[will]_not/never)
يُصْبِح::yu+SobiH+a::he/it+become/begin/start+[sub.]
(NP-SBJ-4 (-NONE- *T*))
(NP-PRD حَقِيقَةٌ::Haqiyq+ap+F::truth/reality+[fem.sg.]+[indef.acc.])
(NP-TMP (PRT إِلَّا::<il~A::however/except/other)
بَعْدَ::baEoda::after
(S-NOM (VP عَوْدَةٌ::Eawodapi::return
(NP-SBJ الْجَوْلَانُ::Al+jawolAn::the+Golan
السُّورِيَّيْنِ::Al+suwriy~::the+Syrian
الْمُحْتَلِّينَ::Al+muHotal~i::the+occupied)
(PP-CLR إِلَى::<ilaY::to/towards
(NP اصْحَابُ::>aSoHAb+i-::owners/originators
هِيَ::-hi::its/his)))

سَرَابُ التَّفَاوُضِ لَنْ يُصْبِحَ حَقِيقَةً إِلَّا بَعْدَ عَوْدَةِ الْجَوْلَانِ السُّورِيِّ الْمُحْتَلِّ

sarAbu + AltafAwuDi + lan + yuSbiHa + HaqiyqapF + <il~A + baEoda + Eawodapi + AljuwlAni + Alsuwriy + AlmuHtal~i
The negotiation mirage will not become real except after the occupied Syrian Golan is back to its owners.

with mA:

(S وَ::wa-::and
(VP (PRT لَمْ::-mA::not)
كَانَ::kAn+a::be/was/were+he/it_[verb]
(VP يَلْعَبُ::ya+loEab+u-::he/it+play+[ind.]
(NP-OBJ هَا::-hA::it/them/her)
(NP-SBJ (NP (PRT إِلَّا::<il~A::except/other)
سِيَاسِيَّيْنِ::siyAsiy~N::politician/statesman)
(PP بِ::bi-::by/with

(NP حجم::-Hajom+i::volume/size+[def.gen.]
 (NP الملك::Al+maliki::the+king
 الراحل::Al+rAHili::the+deceased/late))))))

وَمَا كَانَ يَلْعَبُهَا إِلَّا سِيَّاسِيٌّ بِحَجْمِ الْمَلِكِ الرَّاحِلِ

wa+mA + kAna + yalEabuhA + <il~A + siyAsiy~N + bi+Hajmi + Almaliki + AlrAHili
And no body was playing that role except for a politician like the deceased 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+-LRB-null-RRB-::he/it+remain/last
 (NP-SBJ (NP أحد::AHadN::somebody/someone
 (NP (PRT إلا::-<il~A::except/other)
 الكبار::Al+kibAr+u::the+adults+[def.nom.]
 (PP-TMP في::fiy::in
 (NP نهاية::nihAy+ap+i::end/termination+[fem.sg.]+[def.gen.]
 (NP اللعبة::Al+luEobapi::the+game/toy)))

لَا يَبْقَى أَحَدٌ إِلَّا الْكِبَارُ فِي نِهَآيَةِ اللَّعْبَةِ

lA + yabqaY + >aHadN + <il~A + AlkibAru + fiy + nihAyapi + AlluEbapi
In the end of the game, nobody stays except for the adults

Note: if the two constituents that are supposed to be in an apposition structure are separated by other constituents, the second one modified by <il~A, should be ICHed to the first one.

(S ف::fa-::and/so
 (VP لا::-lA::no/not/non-)

(NP-SBJ (NP منطق::manoTiq+a::logic+[def.acc.]
 (NP-1)
 (NP-PRD عند::Einoda-::with/at
 (NP ها::-hA::it/them/her))
 (NP-1 (PRT لا::<il~A::however/except/other)
 الآلة::Al+|lapa::the+instrument/apparatus/appliance/machine
 العسكرية::Al+Easokariy~apa::the+military/army)))

فَلَا مَنْطِقَ عِنْدَهَا إِلَّا الْآلَةَ الْعَسْكَرِيَّةَ

fa+IA + manTiq + EindahA + <il~A + Al|lapa + AlEaskariy~apa
And there is no logic for her except the military machine.

- <il~A also occurs before prepositions, even when what precedes the prepositional phrase is not negated.

With negation:

(S و::wa-::and
 (VP (PRT لن::-lan::[will]_not/never)
 يسمح::yu+somaH+a::he/it+be_allowed/be_permitted/be_authorized+[sub.]
 (NP-SBJ-1 (-NONE- *))
 (PP-CLR ب::bi-::with/by
 (NP-1 (NP الدخول::-Al+duxuwl+i::the+entrance/penetration+[def.gen.]
 (PP الي::<ilay-::to/towards
 (NP ها::-hA::it/them/her))))
 (PP-MNR (PRT لا::<il~A::however/except/other)
 ب::bi-::by/with
 (NP اذن::-<i*onK::permission/authorization
 رسمي::rasomiy~+K::official+[indef.gen.])))

وَلَنْ يُسْمَحَ بِالذُّخُولِ إِلَيْهَا إِلَّا بِإِذْنِ رَاسِمِيٍّ

wa+lan + yusomaHa + bi+Alduxuwli + <ilay+hA + <il~A + bi+<i*nK + rasmiy~K

And it will not be allowed the entrance to it except with an official authorization.

Without negation:

```
(S (VP تردد:::tarad~ad+a::hesitate+he/it [verb]
  (NP-SBJ (NP المتعاملون:::Al+mutaEAmil+uwna::the+customer+[masc.pl.nom.])
    (PP-CLR في:::fiy::in
      (NP (NP التخلي:::Al+taxal~iy::the+relinquishment/surrender/renunciation)
        (PP عن:::Ean-::from/about/of
          (NP ه:::-hu::it/him))))))
    (PP (PRT الا:::<il~A::except/except for)
      في:::fiy::in
      (SBAR-NOM (WHNP-1 ما:::mA::what)
        (S (VP ندر:::nadar+a::be_rare/be_unusual+he/it_[verb]
          (NP-SBJ-1 (-NONE- *T*))
          (PP من:::min::from
            (NP الحالات:::Al+HALAti::the+cases/situations/instances))))))))))
تردد المتعاملون في التخلي عنه إلا في ما ندر من الحالات.
```

tarad~ada + AlmutaEAmiluwna + fiy + Altaxal~iy + Ean+hu + <il~A + fiy + mA + nadara + min + AlHALAti

The brokers hesitated in giving up on it except in very few cases.

7.2.2.2 The preposition <il~A (PREP)

<il~A as a preposition is used to express exception. Unlike <il~A as a restrictive particle, the preposition is always preceded by an affirmative proposition. <il~A as a preposition assigns the accusative case.

- <il~A + noun in the accusative

```
(SBAR (WHNP-1 الذي:::Al~a*iy::which/who/whom)
  (S (VP يمكن:::yu+mokinu::he/it+be_possible
    (SBAR-SBJ ان:::>an::to
      (S (VP يتناول:::ya+tanAwala::he/it+deal_with/take
        (NP-SBJ-1 (-NONE- *T*))
        (NP-OBJ (NP كل:::kul~a::every/all/each_one
```

(NP شيء::\$ayo'K::something/thing)
 (PP لا::<il~A::except/other)
 (NP موضوع::mawoDuwEa::subject/theme
 (NP (NP الوجود::Al+wujuwdi::the+presence/existence
 الغريب::Al+gariybi::the+strange
 المسلح::Al+musal~aHi::the+armed+[def.gen.]
 (PP-LOC على::EalaY::on/above
 (NP ارض::>aroDi-::earth/territory/land
 (NP نا::~nA::our))))))))))

...الذي يُمكنُ أن يَتَنَاوَلَ كُلَّ شَيْءٍ إِلَّا مَوْضُوعَ الْوُجُودِ الْغَرِيبِ الْمُسَلَّحِ عَلَى أَرْضِنَا

...Ala*iy + yumokinu + >an + yatanAwala + kul~a + \$ay}K + <il~A + mawDuwEa + Alwujuwdi + Algariybi + Almusal~aHi + EalY + >aroDinA

...which can deal with anything except the foreign military presence on our land.

- <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::Al+HiwAri::the+conversation/dialogue/discussion+
 السياسي::Al+siyAsiy~i::the+political))
 (PP لا::<il~A::however/except/other
 (SBAR ان::>an~a-::that
 (S (NP-SBJ ه::~hu::it/he)
 (ADJP-PRD معدوم::maEoduwmN::absent/vanished)
 (NP-TMP اليوم::Al+yawoma::the+today))))

مِنَ الْوَاجِبِ تَوَافُرِ الْحوَارِ السِّيَاسِيِّ إِلَّا أَنَّهُ مَعْدُومٌ الْيَوْمَ

mina + AlwAjibi + tawAfuru AlHiwAri AlsiyAsiy <il~A >an~ahu maEduwmN Alyawma
 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':

لبنان::lubonAni::Lebanon

(NP

الغد::Al+gadi::the+tomorrow))))))))))

(NP

إن الصحافة لا يمكن إلا أن تكون إلى جانب هؤلاء الذين يعملون بجهد من أجل لبنان الغد.

<in~a + AlsaHAfapa + IA + yumokinu + <il~A + >an + takuwna + <ilaY + jAnibi + All~a*iyna + yaEomaluwna + bi+jahodK + min + <ajoli + lubonAni + Algadi

The press can nothing but be on the side of those, who work hard for tomorrow's Lebanon.

7.2.3 EadA عدا / 'except', xalA خلا / 'except', and HA\$A حاشا / 'except'

EadA, xalA and HA\$A: All three sister constructions convey a basic meaning of exception, but have two different POS values with two matching syntactic annotations.

7.2.3.1 The prepositions EadA, xalA, and HA\$A (PREP)

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
(NP-SBJ AljamiyEu الجميع)
(PP EadA عدا
(NP zayodK زيد))))

جاء الجميع عدا زيد

jA'a + AljamiyEu + EadA + zayodK
came + 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
(NP-SBJ AljamiyEu الجميع
(PP EadA عدا
(NP hu هُوَ))))

جاء الجميع عداه

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 المُعَلِّمَات))))

جاء الجميع عدا المعلمات

jA'a + AljamiyEu + EadA + AlmuEal~imAti
came + the all + except + the (female) teachers
Everybody came except the (female) teachers

7.2.3.2 The verbs EadA, xalA, and HASA (VERB)

EadA, xalA, and HASA 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
(NP-SBJ AljamiyEu الجميع)
(S-ADV (VP EadA عدا
(NP-SBJ *)
(NP-OBJ zayodAF زَيْدًا)))))

جاء الجميع عدا زيدًا

jA'a + AljamiyEu + EadA + zayodAF

came + the all + except (he/it) + 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
 (NP-SBJ (NP AljamyEu الجميع)
 (SBAR (WHNP-1 mA ما
 (S (VP EadA عدا
 (NP-SBJ-1 *T*)
 (NP-OBJ zayodAF زَيْدًا))))))

جاء الجميع ما عدا زيداً

jA'a + AljamiyEu + 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
 (NP Al+madoEw~yina المَدْعُوَيْنَ))
 (SBAR-3 *ICH*))
 (VP HaDarwA حَضَرُوا
 (NP-SBJ-1 *T*)
 (SBAR-3 (WHNP-2 mA ما)
 (S (VP EadA عدا
 (NP-SBJ-2 *T*)
 (NP-OBJ (NP Al+muEoTaSimyna الْمُعْتَصِمِينَ
 (PP min مِن
 (NP hum هُمْ))))))))))

كل المدعوين حضروا ما عدا المعتصمين منهم

kul~u + AlmadEuw~iyina + 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

7.2.4 layosa/ليس/ 'not be', 'not'

layosa has 2 POS values and 2 matching syntactic annotations.

7.2.4.1 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.

The TB annotation of layosa is closely related to that of kAna

- 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))
```

لَيْسَتْ الْمَسْأَلَةُ خَطِيرَةً

laysat + Almas>alapu + xaTiyrapF
The issue is not serious

```
(S و::wa-::and
(VP قال::qAl+a::said+he/it_[verb]
(NP-SBJ (-NONE- *))
(S (VP ليس::layos+a::not_be+he/it_[verb]
(PP-PRD في::fiy::in
(NP الامكان::Al+<imokAni::the+possibility/possible))
(NP-SBJ(NP الوقوف::Al+wuquwfu::the+standing/stopping/halting
(NP-ADV طويلاً::TawiylAF::for_a_long_time/at_length)
(NP_ADV عند::Einoda::with/at
(NP(NP المناهج::Al+manAhiji::the+methods/curricula)
(SBAR (WHNP-1 التي::Al~atiy::which/who/whom)
```


(S (VP كانتkAn+at be/was/were+it/they/she
 (NP-SBJ-1 (-NONE- *T*))
 (ADJP-PRD متبعة mut~abaEapFfollowed/observed)
 (ADVP آنذاك::|na*Ak::then/at_that_time))))

وَقَالَ لَيْسَ فِي الْإِمْكَانِ الْوُقُوفُ طَوِيلًا عِنْدَ الْمَنَاهِجِ الَّتِي كَانَتْ أُنْذَاكَ

wa + qAla + laysa + fiy + Al<imkAni + Alwuquwfu + TawiylAF + Einda + AlmanAhiji + Alaty kAnat + |na*Aka
And he said: "It is not possible to ponder a long time at the methods that were in effect then."

- laysa 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))))))

لَيْسَ يُقِيمُ فِي الْقَاهِرَةِ

laysa + yuqiymu + fiy + AlqAhirapi
He does not reside in Cairo

7.2.4.2 The negative particle *laysa* ‘not’ (PART)

laysa is a negative particle annotated as NEG_PART in the following contexts:

- *laysa* is not involved in a subject-predicate relationship.
- *laysa* does not inflect for gender or number to agree with a preceding constituent.
- *laysa* 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 ل::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::the+conversation/discussion)
            و::wa-::and
            (NP PRT ليس::-layos+a::not+[it]
              المواجهة::
              Al+muwAjahapa::the+confrontation))))))))))

```

هناك إهتمامٌ مُتزايدٌ لِحَلِّ النزاعاتِ مِنْ خِلالِ الحوارِ وَ لَيْسَ المُواجَهَة

hunAka + <i>ihtimAmN + mutazAyidN + li+Hal~i + AlnizAEAti + min + xilAli + AlHiwAri + wa laysa + AlmuwAjahap.
 There is an increasing interest in solving the issues through dialogue and not confrontation.

7.2.5 Hat~aY/حتى/‘and even’, ‘so that’, ‘even’, ‘until’

Hat~aY has 4 POS values and 4 matching syntactic representations.

7.2.5.1 The conjunction Hat~aY ‘and even’ (CONJ)

Hat~aY is a conjunction annotated as CONJ when it coordinates two or more constituents with the same syntactic function. As a conjunction, it assigns no case ending. It should be noted that the use of Hat~Y as a conjunction is very rare. The example below illustrates the TB annotation of Hat~Y as a conjunction:

```

(S (VP درشتُ darsoṭu
  (NP-SBJ *))

```

(NP-OBJ (NP kul~a كُـلُّ
 (NP Al<iHtimAlAti الإِـحْتِمَالَاتِ))
 Hat~aY حَتَّى
 (NP(NP AlSagiyrapa الصَّغِيرَةَ
 (PP mino مِنْ
 (NP hA هَا))))))

دَرَسْتُ كُـلَّ الإِـحْتِمَالَاتِ حَتَّى الصَّغِيرَةَ مِنْهَا

darastu + kul~a + Al<iHtimAlAti + Hat~Y + AlSagiyrapa + min+hA
I studied all the possibilities and (even) the insignificant ones too.

7.2.5.2 The subordinating conjunction Hat~aY ‘so that’ (SUB_CONJ)

Hat~aY is a subordinating conjunction when it introduces an adverbial subordinating clause with the meaning of purpose, causality, or condition.

(S (VP (PRT لن:::-lan:::[will]_not/never)
 يحملوا:::ya+Holum+uwA:::the+dream+[masc.pl.]
 (NP-SBJ (-NONE- *))
 (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 (-NONE- *))
 (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 نا:::-nA:::our))))

لَنْ يَحْلُمُوا بِالْأَمْنِ حَتَّى يَرْفَعُوا أَيْدِيهِمْ عَن أُمَّتِنَا

lan + yaHlumuWA + bi+Al>amni + Hat~Y + yarfaEuWA + >aydiy+him + Ean + >um~ati+nA
They will not dream of security unless they take out their rule from our nation.

7.2.5.3 The particle Hat~aY ‘even’ (PART)

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.

```
(S و::wa-::and
  (NP-TPC-1 PRT حتى:::-Hat~aY::even/including/although/even_though)
    واشنطن و::wA$inoTun::Washington
    بوست بوست::buwsot::Post nogloss)
  (VP (PRT لم::lam::did_not)
    تردد::ta+tarad~ad+o::it/they/she+be_hesitant/have_doubts+[jus.]
    (NP-SBJ-1 (-NONE- *T*))
    (PP-CLR في::fiy::in
      (NP إثارة::<ivAr+api::provocation/agitation
        (SBAR-NOM (WHNP-2 ما::mA::what)
          (S (VP يشبه::yu+$obihu::he/it+resemble
            (NP-SBJ-2 (-NONE- *T*))
            (NP-OBJ السخرية::Al+suxoriy~apa::the+sarcasm)
          )
        )
      )
    )
  (NP-TMP (NP يوم::yawoma::day)
    (NP السبت::Al+saboti::the+Saturday
      (NP الماضي::Al+mADiy::the+past/bygone))))))
```

وَ حَتَّى واشنطن بوست لم تتردد في إثارة ما يُشبه السخرية يوم السبت الماضي

wa + Hat~aY + wA\$inTun + buwsot + lam + tatarad~ad + fiy+<ivArapi + mA + yu\$bihu Alsuxriyapa + yawma + Alsabti + AlmADy
And even Washington Post did not hesitate to provokewhat looks like sarcasm last Staurday.

```
(S (VP يجرؤ::ya+joru&a::he/it+dare/risk
  (NP-SBJ-2 (NONE *))
  (PP-CLR (PRT حتى::Hat~aY::even/including/although)
    على::EalaY::on/above
    (NP (NP التفكير::Al+tafokiyyri::the+thinking/meditation/reflection)
      (PP في::fiy::in
        (NP استخدام::isotixodAm::usage/using/utilization+
          (NP السلاح::Al+silAHi::the+weapon
```

النووي::Al+nawawiy~i::the+nuclear/atomic/nucleic)))))))))

يَجْرُؤُ حَتَّى عَلَى التَّفْكِيرِ فِي إِسْتِخْدَامِ السِّلَاحِ النَّوَوِيِّ

yajru&u + Hat~Y + EalY + Altafkiyri fiy <istixdAmi AlsilAHi Alnawawiy~i
He dares even thinking on using nuclear weapons.

7.2.5.4 The preposition Hat~aY ‘until/up to’ (PREP)

Hat~aY is a preposition and is annotated as PREP when it carries a meaning of extent, which, in many cases, can be extent in location or time. In this case, the PP takes an additional –LOC or –TMP dashtags based on the meaning in context. Hat~Y assigns the genitive case to its complement.

(S ::wa-::and
(VP استكملت::-{isotakomal+at::complete/fulfill+it/they/she_[verb]
(NP-SBJ سيول::siyuwl::Seoul)
(NP-OBJ بناء::binA'+a::constructing/building+[def.acc.]
(NP خط::xaT~i::handwriting/script/calligraphy
(NP السكة::Al+sik~api::the+road)))
(PP-LOC حتى::Hat~aY::until/up_to
(NP (NP المنطقة::Al+minoTaqapi::the+area/zone/territory)
(ADJP المنزوعة::Al+manozuwEapi::the+removed/taken_away
(NP السلاح::Al+silAHi::the+weapon)))

وَإِسْتَكْمَلَتْ سِيُولُ بِنَاءَ خَطِّ السِّكَّةِ حَتَّى الْمِنْطَقَةِ الْمَنْزُوعَةِ السِّلَاحِ

wa + <istakmalat + siyuwl + binA'a + xaT~i + Alsik~api + Hat~aY + AlminTaqapi + AlmanzuwEapi + AlsilAHi.
And Seoul finished the construction of a rail line up to the zone of removed weapons.

(S (NP-TPC-1 مسألة::maso>alapu::issue/affair/matter/question
(NP خفض::xafoDi::lowering/decreasing/reduction
(NP (NP الاسلحة::Al+>asoliHapi::the+weapons

(ADJP
 الاستراتيجية::Al+<istrAtiyjiy~api::the+strategic
 النووية::Al+nawawiy~api::the+nuclear/atomic)))
 (VP ستناقش::sa+tu+nAqa\$+u::will+it/they/she+be_debated/be_argued+[ind.]
 (NP-SBJ-1 (-NONE- *T*))
 (NP-OBJ-1 (-NONE_ *))
 (PP على::EalaY::on/above
 (NPمستويات::musotawayAtK::levels/standards
 عدة::Eid~apK::several/numerous/many))
 (PP-TMP من::min::from
 (NPالآن::Al+|n+a::now+[def.acc.])))
 (PP-TMP حتى::Hat~aY::until/up_to
 (NPالصيف::Al+Sayofi::the+summer
 المقبل::Al+muqobili::the+next/coming/approaching))))))

مسألة خفض الأسلحة الاستراتيجية النووية ستناقش على مستويات عدة من الان حتى الصيف المقبل.

maso>alapu + xafDi + Al>asolihapi + Al<istrAtiy~api + Alnawawiy~pi + satunAqa\$u + EalaY + musotawayAtK+ Eid~apK + min
 + Al|ana + Hat~aY + AlSayofi + Almuqobili

The issue of strategic nuclear weapons will be discussed at various levels from now until next summer.

7.2.6 IA/ لا/ 'not', 'no', 'not be'

IA has 4 POS values: a pseudo verb, a conjunction, a negative particle, and an interjection. It has 4 different syntactic representations based on those POS values.

7.2.6.1 The pseudo verb IA 'not be' (PSEUDO_VERB)

IA is a pseudo-verb and annotated PSEUDO_VERB when it involves a relationship of subject and predicate. Like any other pseudo-verb, IA assigns the accusative case to its subject. Typical constructions involving IA as a pseudo-verb are: لا بد (IA bu~da/ no escape/avoiding) لا شكَّ (IA \$ak~a/no doubt), لا مَنَاصَ (IA manASa/no solution), لا حاجة (IA HAjapa/no need), etc.

(S ::wa-::and
 (S (VP لا::lA::no/not/non-)

(NP-SBJ بد::bud~+a::escape/avoiding+[def.acc.])
 (PP-PRD من::min::from
 (NP (NP (NP التوصل::Al+tawaS~uli::the+attainment/arrival/reunion)
 (PP الى::<ilaY::to/towards
 (NP (NP اتفاق::{it~ifAqK::agreement/accord/treaty)
 (NP-ADV مع::maEa::with
 (NP (NP (NP الدول::Al+duwali::the+states/countries
 الاخرى::Al+>uxoraY::the+other/another/additional)
 (NP (NP (NP الاعضاء::Al+>aEoDA'i::the+members)
 (PP في::fiy::in
 (NP (NP (NP الاتحاد::Al+{it~iHAdi::the+union
 الاوروبي::Al+>uwruwb~iy~i::the+European))))))))))

لا بد من التوصل إلى اتفاق مع الدول الأخرى الأعضاء في الاتحاد الأوروبي.

IA + bud~a + min + AltawaS~uli + <ilaY + Hal~K + maEa + Alduwali + Al<uxraY + Al>aEoDA}I + fi Al>it~iHAdi + Al>uruwby~iyi

There must be a solution with the other countries members of the European Union.

(VP لا::-lA::no/not/non-)
 (NP-SBJ شك::\$ak~a::doubt)
 (PP-PRD في::fiy::in
 (SBAR ان::>an~a::that
 (S (NP-TPC-1 اسرائيل::<isorA}iyla::Israel)
 (VP (PRT قد::qad::[has/have])
 تعرضت::taEar~aDat::be_exposed_to/encounter
 (NP-SBJ-1 (-NONE- *T*))
 (PP-CLR ل::li-::for/to
 (NP (NP (NP صفة::-SafoEapK::slap/blow
 قاسية::qAsiy+ap+K::harsh/cruel/severe)
 (PP من::min::from
 (NP (NP (NP المقاومة::Al+muqAwamapi::the+resistance/opposition)
 (PP-LOC في::fiy::in
 (NP (NP (NP لبنان::lubonAna::Lebanon))))))))))

وَلَا شَكَّ أَنَّ إِسْرَائِيلَ قَدْ تَعَرَّضَتْ إِلَى صَفْعَةٍ قَاسِيَةٍ مِنَ الْمُعَارَضَةِ فِي لُبْنَانَ

wa+IA + \$ak~a + >an~a + <isrA}iyla + qad + taEar~aDat + <iIY + SafEapK + qAsiyapK + mina + AlmuEAraDapi + fiy + lubnAna
And there is no doubt that Israel got a severe blow from the part of the resistance in Lebanon.

(S (VP لا:::~lA:::no/not/non-
 (NP-SBJ-1 احد:::>aHada:::one)
 (S (VP يتجرأ:::ya+tajar~a>+u:::he/it+dare/risk+[ind.]
 (NP-SBJ-1 (-NONE- *))
 (PP-CLR على:::EalaY:::on/above
 (NP (NP رفض:::rafoDi:::rejection
 (NP (NP هذا:::h`*A:::this
 (NP العمل:::Al+Eamali:::the+work/action))))))

لا أَحَدَ يَتَجَرَّأُ عَلَى رَفْضِ هَذَا الْعَمَلِ

IA >aHada yatajar~a>u EalY rafDi ha*A AlEamali
Nobody dares to refuse this job.

7.2.6.2 The conjunction IA ‘not’ (CONJ)

IA is a conjunction and annotated as CONJ when it coordinates two adjacent constituents sharing the same syntactic function with the second conjunct being negated. Refer to section 7.2.6.5 for tests distinguishing IA as a conjunction from the IA as a negative particle. As a conjunction, IA does not involve a relationship of subject and predicate.

(S) (S (NP-SBJ المسألة:::Al+maso>alapu:::the+issue/affair/matter/question)
 (ADJP-PRD سياسية:::siyAsiy~apN:::political
 (لا:::IA:::no/not)
 قضائية:::qaDA}iy~+ap+N:::judicial/legal+[fem.sg.]+[indef.nom.]
 (S و:::wa-:::and
 (NP-TPC-1 الموقعون:::~Al+muwaq~iEuwna:::the+signing/signatory)
 (VP يمثلون:::yu+mav~il+uwna:::they_[people]+represent+[masc.pl.]
 (NP-SBJ-1 (-NONE- *T*))
 (NP-OBJ (NP دولاً:::duwal+AF:::states/countries+[acc.indef.]
 (PP-LOC (PP على:::EalaY:::on/above

(NP الارض::Al+>aroDi::the+earth))
 (لا::lA::no/not)
 (PP في::fiy::in
 (NP المريخ::Al+mir~iyxi::the+Mars))))))

المسألة سياسية لا قضائية و الموقعون يمثلون دولاً على الأرض لا في المريخ

Almas>alapu + siyAsiy~apN + lA qaDA}iy~apN + wa+Almuwaq~iEuwna + yumav~iluwna + duwalAF + EalY + Al>arDi + lA + fiy + Almir~yxi

The issue is political not judicial and the signers represent countries on earth not on Mars.

7.2.6.3 The negative particle lA ‘no/not’ (NEG_PART)

lA is a negative particle that is annotated as NEG_PART when it negates a verb, any other non-verbal phrases not preceded by any form of assertion (see second and third examples below for more explanation), or when it negates two entities simultaneously (see fourth example below). It does not involve a relationship of subject and predicate.

- Negation of the verb

(S (NP-TPC-1 (NP التهديدات::Al+tahodiydAtu::the+threats/dangers)
 (SBAR (WHNP-2 التي::Al~atiy::which/who/whom_[fem.sg.]
 (S (VP وجهت::wuj~ih+at::be_sent/be_directed+it/they/she_[verb]
 (NP-SBJ-2 (-NONE- *T*))
 (NP-OBJ-2 (-NONE- *))
 (PP-CLR إلى::<ilaY::to/towards
 (NP العقيد::Al+Eaqiydi::the+colonel
 القذافي::Al+qa*~Afiy~::the+Qaddafi))))))
 (VP (PRT لا::lA::no/not/non-)
 يستهان::yu+sotahAnu::he/it+be_disdained/be_underestimated
 (NP-SBJ-3 (-NONE- *))
 (PP ب::bi::with/by
 (NP-3 (NP ها::-hA::it/them/her)
 (NP-1 (-NONE- *T*))))))

التهديدات التي وُجّهت إلى العقيد القذافي لا يُستهانُ بها

AltahdiydAtu + Alatiy + wuj~ihat + <ilaY + AlEaqiydi + Alqa*Afiy + IA + yustahAnu + bi+hA
The threats addressed to Colonel Qaddafi should not be underestimated.

- Negation of other phrases

```
(S  و::wa-::and
  (PP  ب::-bi-::by/with
    (NP  ذلك::-*`lika::that))
  (VP  يكون::ya+kuwn+u::he/it+be+[ind.]
    (NP-1  الامام::Al+<imAmu::the+leader/imam
      الصدر::Al+Sdr::the+NOT_IN_LEXICON))
    (VP  (PRT  قد::qad::[has/have])
      عاد::EAd+a::return/go_back+he/it_[verb]
      (NP-SBJ-1  (-NONE- *T*))
      (PP-DIR  (PP  (PP  (PRT  لا::lA::no/not/non-)
        من::min::from
          (NP  ليبيا::liybiyA::Libya))
        بل::bal::but_rather/in_fact
          (PP  من::min::from
            (NP  فلسطين::filasoTiyna::Palestine))))))
```

وَبَدَلِكَ يَكُونُ الْإِمَامُ الصَّدْرُ قَدْ عَادَ لَا مِنْ لِيْبِيَا بَلْ مِنْ فِلِسْطِيْنِ

wa+bi+*alika + yakuwnu + Al<imAmu + AlSadru + qad + EAda + IA + min + lybyA + bal + min + filisTyn
And with that the Imam AlSadru would have returned not from Libya but from Palestine

Note: In this example, IA is annotated as PART because the assertion ‘bal min falaSotiyn’ follows the negation. When the assertion precedes the negation, IA is annotated as CONJ as shown in the section 7.2.6.2. above.

```
(S (VP  يغدو::-ya+goduw+-null::he/it+depart/become_current/appear/seem
  (NP-SBJ  الصراع::Al+SirAEu::the+struggle/conflict/fight)
  (PP-PRD  ب::bi-::with
    (NP  (PRT  لا::-lA::no_[without])
      حدود::HuduwdK::border/frontier/limits))
```

yagoduw+ AlSirAEu + bi+IA + HuduwdK
The conflict becomes without borders

Note: In this example, the negation is not accompanied by any form of assertion whether before or after the negated phrase.

7.2.6.4 The interjection IA ‘no’ (INTERJ)

IA is an interjection and annotated as INTERJ when it is not bound syntactically to any other sentence elements most typically as an answer to a yes-no question. In this regard, its most common occurrence is in direct speech.

```
(S wa و
  (VP >ajAba أَجابُ
    (NP-SBJ * )
    (PP-MNR bi ب
      (NP Hid~apK حِدَّةٌ ) )
  : :
  (S "
    (INTJ IA لا )

    (VP (PRT IA لا )
      >uwAfiqū أُوافقُ
      (NP-SBJ * )
      (PP-CLR EalaY عَلَى
        (NP ha*A هَذَا ) ) )
  " ) ) )
```

وَأَجَابَ بِحِدَّةٍ : لا، لا أوافقُ عَلَى هَذَا

wa+>ajAba + bi+Hid~apK: "IA + IA + >uwAfiqū + EalY + ha*A"
He answered severely : « No, I do not agree on that »

The interjection IA 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.

```
(S (NP-TPC-2 جواب::jawAb+u::answer+[def.nom.]
    (NP خبائر::xubarA}+i-::experts/specialists+[def.gen.]
      (NP لنا::-nA::our))
  (VP كان::kAn+a::be/was/were+he/it_[verb]
    (NP-SBJ-2 (-NONE- *T*))
    (NP-PRD لا::lA::no))))))
```

جَوَابُ خُبْرَائِنَا كَانَ لَا

jawAbu xubarA}inA kAna lA
The response of our experts was no

7.2.6.5 Tests to distinguish IA as a conjunction from IA as a negative particle

IA as a negative particle is not ambiguous when it comes before a verb. It becomes ambiguous when it precedes a noun, a prepositional phrase or any constituent other than a verb in which case it has to be clearly separated from its function as a conjunction.

A useful test to differentiate between IA as a conjunction and IA as a negative particle is to look at the syntactic context where both occur.

- When the IA is used before two negated elements; that is to say, when there is a IA before the first element and another IA before the second, then the two IAs are annotated as negative particles. In these contexts, the second IA is preceded by a wa-, a phenomenon that is not possible when IA 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, IA 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

ra>aytu + zaydAF + IA + EamrAF

I saw Zayd not Amr

7.2.7 wa/ و /’and’, ‘while’, ‘by’, ‘with’

wa has 3 POS values: conjunction, subordinating conjunction, and preposition.

7.2.7.1 The conjunction wa ‘and’ (CONJ)

Both the wa of coordination (واو العطف / wAw AlEatf) and the wa of connection (واو الرّبط / wAw AlrabT) are annotated as conjunctions in the following contexts:

- Coordination of two constituents of the same syntactic category and function.

```
(PP-PRP ل::li-::for/to
  (NP لتقويم::-taqowiymi::rating/valuation
    (NP مدى::madaY::extent/range
      (NP (NP جدية::jid~iy~api::seriousness
          (NP الحكومة::Al+Hukuw+api::the+government
            و::wa-::and
            النواب::-Al+vuw~Ari::the+deputies))
        (PP في::fiy::in
          (NP احلال::<iHolAl::achieving/bout
            (NP السلام::Al+salAmi::the+peace))))))
```

... لتقويم مدى جدية الحكومة والنواب في احلال السلام

li+taqowiymi + madaY + jid~yapi + AlHukwumapi + wa+Alnuw~Abi + fiy + <iHolAli + AlsalAmi
...in order to evaluate the level of seriousness of the government and the deputies in establishing peace.

(S (VP حضر::HaDar+a-::attend/appear/be_present+he/it_[verb]
 (NP-OBJ هـ:::-hu::it/him)
 (NP-SBJ (NP العميد::Al+Eamiydu::the+dean/chief
 الدكتور::Al+dukotuwru+u::the+doctor/Doctor/Dr.
 عبد::Eabod::Abd
 الرؤوف::Al+ra&uwf::the+Raouf
 سنو::snw::NOT_IN_LEXICON)
 و::wa-::and
 (NP مديرو::-mudiyruw::director/manager/chief
 (NP الفروع::Al+furuwEi::the+branches/sections))
 و::wa-::and
 (NP الاساتذة:::-Al+>asAti*apu::the+professors))))

حضره العميد الدكتور عبد الرؤوف سنو ومديرو الفروع والأساتذة.

wa+haDara+hu + AlEamiydu + Alduktuwru + Eabd + Alra&uwf + sin~uw + wa+mudiyruw + AlfuruwEi + wa+AlsAti*apu.
Dean Dr. Adb Alraoof Sinu and the directors of the subsections and the teachers attended it

(NP (NP القرار:::-Al+qarAri::the+decision/resolution)
 (ADJP التاريخي::Al+tArixyi~i::the+historical
 و::wa-::and
 الفريد:::-Al+fariydi::the+unique/incomparable/exceptional))

القرار التاريخي والفريد

AlqarAru + AltArixyi~u + wa+Alfariydu
The historic and exceptional decision

- 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.

(S و::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/permitted))
 و::wa-::and

(NP غير:::gayor+a::not/other
 (NP مقبول:::maqobuwlK::accepted/welcome))
 (PP-LOC في:::fiy:::in
 (NP (NP منطقة:::minoTaq+apK::area/zone/territory)
 (UCP (ADJP مكتظة:::munktaD~apK::populated
 (PP ب:::bi-:::with/by
 (NP السكان:::-Al+suk~Ani:::the+residents
 و:::wa-:::and
 المتاجر:::-Al+matAjiri:::the+stores)))
 و:::wa-:::and
 (SBAR (WHNP-1 *0*))
 (S (VP تستحق:::-ta+sotaHiq~u:::it+deserve
 (NP-SBJ-1 *T*)
 (NP-OBJ (NP الاهتمام:::
 Al+{ihotimAma:::
 the+interest

”

)))))))))

وقالوا هذا غير مسموح وغير مقبول في منطقة مكتظة بالسكان والمتاجر وتستحق الإهتمام.

wa + qAluwA + “ + ha*A + gayoru + masmuwHK + wa+gayoru + maqbuwlK + fiy + mintaqapK + mukotaD~apK +
 bi+Alsuk~Ani + wa+AlmatAjiri + wa+tastaHiq~u + Al<ihtimAma ”

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”

- when it occurs at the beginning of the paragraph functioning like a discourse connector.

وَأَشَارَ الْمَوْطِنُ الْفِرَنْسِيُّ وَ بَلْكَنَّةٍ إِنْكَلِيزِيَّةٍ مُكْسَّرَةً إِلَى مَعْلُومَاتٍ سِرِّيَّةٍ

wa >a\$Ara AlmuwATinu Alfiransiy~u wa bilaknapK <inklyziy~ap mukas~arap <iLY maEluwmAtK siriy~apK.

And the French citizen pointed out and with a broken English accent to confidential information.

[The translation here ends up as not great English – perhaps a simpler example would do just as well?]

- 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.

(NP (NP رئيس::ra}iysu::president/head/chairman
 (NP الشرطة::Al+\$uroTapi::the+police+[fem.sg.])))
 (PP-LOC في::fiy::in
 (NP (NP اقليم::<iqoliymi::region/district
 (NP السند::Al+sinod::the+Sind))
 (NAC و::wa-::and
 (SBAR (WHNP-3 *0*)
 (S (NP-SBJ عاصمة::EASimatu-::capital_city
 (NP (NP ه::hu::his/its)
 (NP-3 *T*))
 (NP-PRD كراتشي::karAto\$iy::Karachi)))))))))

رئيس الشرطة في اقليم السند وعاصمته كراتشي.

ra}iysu + Al\$uroTapi + fiy + <iqoliymi + Alsind + wa+EASimatu+hu + KaraAt\$
The police chief in the district of Sind and with Karatshi as its capital.

(S (NP-SBJ هذا::-h`*A::this_[masc.sg.]
 (SBAR-NOM-PRD (WHNP-2 ما::mA::what)
 (S (VP يعطي::yu+EoTiy+null::he/it+give/provide
 (NP-SBJ-2 *T*)
 (NP-DTV الفلسطينيين::Al+filasoTiyiny~iyana::
 the+Palestinian)
 (NP-OBJ حقا::Haq~AF::truth/right
 اكثر::>akovar+a::more/most)
 (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.

ha*A + mA + yuEoTiy + li+AlfalasTinyyna + Haq~AF + <akvara + wa+li + mud~apK + <ATowala
This is what gives the Palestinians more right and for a longer time

- Another type of NAC is the extraposed constituent where the constituent inside the NAC will be ICHed to another constituent it usually coordinates with, because of an intervening constituent between the two conjuncts.

(PP على::EalaY::on/above
 (NP امل::>amali::hope/wish))
 (SBAR ان::>an::to
 (S (VP يفى::ya+fiy+a::he/it+fulfill/carry_out+[sub.]
 (NP-SBJ المسؤولون::Al+maso&wluwna::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 يفى المسؤولون ببقية وعودهم في أقرب فرصة ممكنة وخاصة إعادة تأهيل البنية التحتية.

ElaY + <amali + <an + yafiy + Almaso&wluwna + bi+baqy~api + wuEwdi+him + fiy + <aqorabi 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.

N.B. An NAC of a coordinated NP-OBJ might be a more clear example to have here.

7.2.7.2 The subordinating conjunction wa ‘while’ (SUB_CONJ)

The circumstantial wa (او/ او الحال/ waw AIHA) is a subordinating conjunction annotated as SUB_CONJ when it introduces an adverbial clause that describes the circumstances under which an event took place. As such, the subordinating clause introduced by wa will take its most appropriate adverbial dashtag according to context.

The SBAR introduced by the circumstantial wa can be an equational sentence with a pronoun subject or a verbal sentence with a topicalized NP. In both cases, the circumstantial wa is always followed by a nominal entity.

```
(S wa وَ
  (VP lakin~ لَكِنَّ
    (NP-SBJ iyِي )
    (NP-ADV ragma رَغْمَ
      (NP *alika ذَلِكَ ) )
    (S (VP >aHkiy أَحْكِي
      (NP-SBJ * )
      (NP-OBJ hA هَا )
      (SBAR-MNR wa وَ
        (S (NP-SBJ >anA أَنَا )
          (SBAR-NOM-PRD (WHNP-1 0 )
            (S (VP wAviqN وَائِقُ
              (NP-SBJ-1 *T* )
              (SBAR >an~a أَنَّ
                (S (NP-TPC-2 hu هُوَ )
                  (VP (PRT lan لَنْ )
                    yagDaba يَغْضَبُ
                    (NP-SBJ-2 *T*)
                    ))))))))))))
```

وَلَكِنِّي رَغْمَ ذَلِكَ أَحْكِيهَا وَأَنَا وَائِقٌ أَنَّهُ لَنْ يَغْضَبَ

wa+lakin~iy + ragma + *alika + >aHokiy+hA + wa+>anA + wAviqN + >an~a+hu + lan yagDaba
And despite this I tell it being confident that he will not be angry

```
(S (VP يحتاج:::ya+HotAj+u::he/it+need/want+[ind.]
  (NP-SBJ *)
  (PP-CLR الى:::ilaY::to/towards
    (NP (NP معجزة:::mEjrp::nogloss)
      (PP من:::min::from
        (NP (NP (NP ذلك:::*`lika::that )
          (NP النوع:::Al+nawoEi::the+type/kind)
          (SBAR (WHNP-2 الذي:::Al~a*iy::which/who/whom)
            (S (VP يحيي:::yu+Hoyiy+null-:::he/it+enliven
              (NP-SBJ-2 *T*)
```

(NP-OBJ العظام ::Al+EiZama::the+bones)

(SBAR-ADV و ::wa-::while

(S (NP-SBJ هي ::-hiya::

it/they/she)

(NP-PRD رميم ::ramiym+N::

decayed/rotten

))))))))))

يحتاج إلى معجزة من ذلك النوع الذي يحيي العظام وهي رميم.

yaHotAju + <ilaY + muEojizapK min *alika AlnawoEi Al~a*iy yuHoyiy AlEiDAma wa hya ramiymN

It needs a miracle of the type that resuscitates bones when they become ashes

7.2.7.3 The preposition wa (PREP)

wa as a preposition has two meanings: the wa of oath (واو القَسَم /waw Alqasam) and the wa of accompaniment (واو المَعِيَّة / waw AlmaEiy~ap).

7.2.7.3.1 The wa of oath ‘by’

This wa is used in oaths. It assigns the genitive case to its complement.

(PP و ::wa::by

(NP الله ::Al~ahi::God))

وَ اللهُ

wa+All~ahi

By God

7.2.7.3.2 The wa of accompaniment ‘with’

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.’

(NP ضرورة ::-Daruwrapi::necessity/need/imperative

(NP (NP (NP الحفاظ ::Al+HifaZi::the+preservation/guarding)

(PP علي::Ealay-::on/above
 (NP ه::-hi::it/him)))
 و::wa-::and
 (NP (NP تطوير::-taTowiyri-::development/advancement/promotion
 (NP ه::-hi::its/his))
 (PP ب::bi-::by/with
 (SBAR-NOM (WHNP-2 ما::-mA::what)
 (S (VP يتوافق::ya+tawAfaqu::he/it+agree/concur
 (NP-SBJ-2 *T*)
 (PP-CLR و::wa-::with
 (NP (NP تصنيف::-taSoniyyfa::
 classification
 (NP جبيل::jubayol::
 Jubayl))
 (PP ك::ka-::like/such_as
 (NP (NP موقع::-mawoqiEK::
 location)
 (PP ل::li-::to/for
 (NP لتراث::-Al+turAvi
 ::the+heritage
 العالمي::
 Al+EAlamiy~i::
 the+
 international/
 world
)))))))

ضرورة الحفاظ عليه وتطويره بما يتوافق وتصنيف جبيل كموقع للتراث العالمي.

Daruwrapu + AlHifADi + Ealayhi + wa+taTwiyrhi + bi+mA + yatawAfaqu + wa+tasniyya + jubayol + ka+maowqiEK +
 li+lturAvi + AlEAlamy~i

The necessecity of preserving it and developing it with what complies with the classification of Jubail as a site of world heritage

```

(S (VP (PRT لا::lA::no/not/non-)
      (NP-SBJ *))
  (VP يستطيع::ya+sotaTiyE+u::he/it+be_able/be_capable
    (SBAR (SBAR ان::>an::to
            (S (VP يفهم::ya+foham+a-::he/it+understand
                  (NP-SBJ *)
                  (NP-OBJ ه::-hu::it/him))))
            أو
            (SBAR ان::>an::to
              (S (VP يتفاهم::ya+tafAhama::he+understand_one_another
                    (NP-SBJ *)
                    (PP-CLR و::wa-::with
                      (NP (PRT لا::<iy~A-::to/for)
                          ه::-hu::it/him))))))))))

```

لا يستطيع أن يفهمه أو أن يتفاهم و إياه.

lA + yasotaTiyEu + >an + yafohama+hu + awo + >an + yatafAhama + wa+<iy~Ahu
He can not understand him or agree with him

7.2.7.4 The particle wa- (PART) [Under Construction]

7.2.7.5 Tests distinguishing the wa of coordination, NAC clauses introduced by wa, and the circumstantial wa

The subordinating conjunction wa is particularly problematic to differentiate from the conjunction wa whether the latter serves as a coordinator between two constituents or as a conjunction heading an NAC construction.

7.2.7.5.1 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 bizuhuwrK EalaY \$akoli raqomi 7 wahuwa raqmu Alqamiysi Al~a*iy kAna yarotadiyhi mAviuwz xilAla AlmuArAti.

And the gasket was decorated with flowers in the shape of the number 7 and this was Matthews' shirt's number.

هذه الفئة من العمال تعمل في مجال عمل واحد وهو مجال صناعة الأحذية

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

jA'a wa huwa DAHikN/jA'a wa huwa yaDHaku → jA'a DAHikAF

He came and he was laughing/He came while he was laughing → He came laughing

- 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

mAtat <ibonatuhu wahya talidu bisababi <ihomAli Al>aTib~A'i

She died while she was giving birth because of the doctors' carelessness

7.2.7.5.2 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:

```
(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::-waEoyK::consciousness/awareness
                      کامل::kAmilK::complete/full/integral)))
              (SBAR-NOM-OBJ (WHNP-3 ماذا::mA*A::what)
                  (S (VP حصل::HaSal+a::occur/happen+he/it
                      (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 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 *)
            (PP-MNR ب::-bi-::by/with
                  (NP وعي::-waEoyK::consciousness/awareness
                       کامل::kAmilK::complete/full/integral))
            (SBAR-NOM-OBJ (WHNP-3 ماذا::mA*A::what)
                          (S (VP حصل::HaSal+a::occur/happen+he/it
                                (NP-SBJ-3 *T*))))))))))
```

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

قد بتنا ندرک ووعي کامل ماذا حصل

Example with an SBAR relative clause:

```
(S (VP قاد::qAd+a::lead/guide+he/it_[verb]
  (NP-SBJ *)
  (NP-TMP اخيرا::>axiyraAF::last/latest/recent/latter)
  (NP-OBJ (NP (NP المفاوضات::Al+mufAwaDAti::the+negotiationstalks)
    (NP-ADV بين::bayona::between/among
      (NP (NP الاردن::Al+>urodun~i::the+Jordan)
        و::wa-::and
        (NP (NP الولايات::-Al+wilAyAti::the+States
          المتحدة::Al+mut~aHidapi::the+United))))))
  (NAC و::wa-::and
    (SBAR (WHNP-2 التي::-Al~atiy::which/who/whom_[fem.sg.])
      (S (VP توجت::tuw~ij+at::be_crowned+it
        (NP-SBJ-2 *T*)
        (NP-OBJ-2 *)
        (PP-CLR ب::bi-::by/with
          (NP توقيع::-tawoqiyEi::signing
            (NP اتفاق::{it~ifAqi::treaty
              (NP التجارة::
                Al+tijArapi::
                the+commerce
                الحرة::
                Al+Hur~api::
                the+free
                ))))))))))))
```

→

(S (VP قاد::qAd+a::lead/guide+he/it_[verb]
 (NP-SBJ *)
 (NP-TMP اخيرا::>axiyRAF::last/latest/recent/latter)
 (NP-OBJ (NP (NP المفاوضات::Al+mufAwaDAti::the+negotiationstalks)
 (NP-ADV بين::bayona::between/among
 (NP (NP الاردن::Al+>urodun~i::the+Jordan)
 و::wa-::and
 (NP الولايات::-Al+wilAyAti::the+States
 المتحدة::Al+mut~aHidapi::the+United))))))
 (SBAR (WHNP-2 التي::-Al~atiy::which/who/whom_[fem.sg.]
 (S (VP توجت::tuw~ij+at::be_crowned+it
 (NP-SBJ-2 *T*)
 (NP-OBJ-2 *)
 (PP-CLR ب::bi-::by/with
 (NP توقيع::-tawoqiyEi::signing
 (NP اتفاق::it~ifAqi::/treaty
 (NP التجارة::Al+tijArapi::
 the+commerce
 الحرة::Al+Hur~api::the+free
)))))))

قاد اخيرا المفاوضات بين الاردن و الولايات المتحدة والتي توجت ب توقيع اتفاق التجارة الحرة.

qAda + AxiyrAF + AlmufAwaDAti + bayona + Al<urduni + wa+AlwilAyAti + ALmut~aHidapi + wa+All~atiy + tuw~ijjat +
 bi+<it~ifAqi + AltijArati + AlHur~api

And conducted lately the negotiations between Jordan and the United States that led to signing the treaty of free trade

7.2.8 mA/ ما /'not', 'that', 'as long as', 'as soon as', 'what', 'some', 'not be'

mA has 8 POS values: negative particle, particle, interrogative pronoun, relative pronoun, exclamative pronoun, subordinating conjunction, non-inflectional verb, and noun.

7.2.8.1 The negative particle mA/ 'not' (NEG_PART)

mA is a negative particle annotated as NEG_PART when it occurs only before verbs to negate them. In this context, the mA can be replaced by another negative particle with no change in meaning. mA as a negative particle often occurs in collocation with verbs like زال (zAla/ no longer be), فتي (fati}a/ refrain from), انفاك (Ainfak~a/ cease), etc. See section 4.1.2.2.2 of the syntactic annotation guidelines for further examples.

```
(S (PP-PRD في::fiy::in
  (NP معنى::maEonaY::meaning/sense
    (NP (NP هذا::h`*A::this_[masc.sg.]
      (NP الاسم::Al+{isom+i::the+name+[def.gen.]
        السامي::Al+sAmiy~i::the+elevated))))
    (NP-SBJ بداية::bidAy+ap+u::beginning/start+[fem.sg.]+[def.nom.]
      (NP (NP قداسة::qadAsapK::sanctity/His_Holiness)
        (SBAR (WHNP-1 *0*)
          (S (VP (PRT ما::mA::not)
            عرفت::Earaf+at::know+it/they/she
            (NP-SBJ-1 *T*)
            (NP-OBJ ظللاً::ZilAlAF::auspices/
              shelter))))))))))
```

في معنى هذا الاسم السامي بداية قداسة ما عرفت ظللاً

Fiy + maEonaY + ha*A + Al<ismi + bidAyapu + qadAsapK + mA + Earifat + ZilalAF
In the meaning of this sacred name there is the beginning of a holiness that never knew any flaws

7.2.8.2 The subordinating conjunction mA/ ‘that’, ‘as long as’, ‘as soon as’ (SUB_CONJ)

mA is a subordinating conjunction and annotated as SUB_CONJ in two cases:

- 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 / mA Almasodary~ap) and usually follows words like Einda, baEoda, mivola, etc.

(S (NP-TPC-1 (NP كثير::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 أخطأوا::>axoTa>+uwA::be_wrong/do_incorrectly_[verb]
 (NP-SBJ-2 *T*)
 (NP-TMP عند::Einoda-::when)
 (SBAR لما::-mA::that
 (S (VP تولوا::tawal~+awoA::
 take_charge_of
 (NP-SBJ *)
 (NP-OBJ المناصب::Al+manASiba::
 the+posts/positions
 الرسمية::Al+rasomiy~apa::
 the+official))))))))))

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 + Alrasmy~apa.

And many religious scholars confessed that they were wrong when they took on official (government) positions

- When mA introduces a subordinate clause with a conditional meaning. This is called the conditional mA (ما الشرطيّة / mA Al\$arTiy~ap)

(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/guiding/sending
 (NP المضادات::Al+muDAd~Ati::the+anti-/counter))))

(SBAR-ADV ما::mA::as long as

(S (VP دام::dAm+a::last/persevere/continue+he/it_
(NP-SBJ الانتهاك::Al+{inotihAku::the+violation
الاسرائيلي::Al+{isorA}ilyiy~u::
the+Israeli)
(ADJP-PRD قائما::qA}imAF::ongoing/
existing/present))))))

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 {ilyiy~u +
qA {imAF.

Hizbullah will not accept not aiming anti-missile missiles as long as the Israeli violations are ongoing

N.B. An additional example of mA in the meaning of ‘as soon as’ needs to be added.

7.2.8.3 The relative pronoun mA/ ‘what’ (REL_PRON)

mA can also be a relative pronoun and annotated as 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 in 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.’

[Need another example with the free relative in a regular subject position with a verb]

(S (NP-SBJ هذا::-h`*A::this_[masc.sg.]
(SBAR-NOM-PRD (WHNP-3 ما::mA::what)
(S (VP كان::kAn+a::be/was/were+he/it_[verb]
(NP-SBJ-3 *T*)
(ADJP-PRD حاصل::HASil+AF::
holding+[acc.indef.]))))))

هَذَا مَا كَانَ حَاصِلًا

ha*A + mA + kAna + HASilAN
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 vum~a ثم
  (VP bada>a بدأ
    (NP-SBJ Aljumohuwru الجُمهُورُ )
    (PP-CLR bi ب
      (NP (NP AlhutAfi الهُتافِ )
        wa و
        (NP (NP <iloqA'i إلقاءِ
          (NP AlqAruwrAti القارُوراتِ
            AlfArigapi الفارِغَةِ ) )
          (NP-DIR wasaTa وَسَطَ
            (NP AlmaloEabi المَلْعَبِ ) ) ) ) ) )
    (SBAR-ADV (WHNP-1 mA ما )
      (S (VP ADTar~a اضطرَّ
        (NP-SBJ-1 *T* )
        (NP-OBJ AlHakama الحَكَمَ )
        (PP-CLR <ilaY إلى
          (NP <iyqAfi إيقافِ
            (NP AlmubArAti المباراتِ))))))))
    ثم بدأ الجمهور بالهتاف وإلقاء القارورات الفارعة وسط الملعب ما اضطر الحكم إلى إيقاف المبارات
  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
```

7.2.8.4 The interrogative pronoun mA/ 'what' (INTERROG_PRON)

mA can be an interrogative pronoun and annotated as 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 دفعك::dfEtk::nogloss
    (NP-SBJ-1 *T*)
    (PP-CLR ال::<ilaY::to/towards
```

(NP (NP المشاركة::Al+mu\$Arakapi::the+participation)
 (PP في::fiy::in
 (NP (NP هذه::h`*ihi::this)
 (NP المغامرة:: Al+mugAmarapi::
 the+adventure)))

?)

ما دفعك إلى المشاركة في هذه المغامرة؟

mA + dafaEa+ka + <ilaY + Almu\$Arakapi + fiy + ha*ihi + AlmugAmarapi
What pushed you to take part in this adventure?

7.2.8.5 The exclamative pronoun mA/ 'what' (EXCLAM_PRON)

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 (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::Al+taTar~ufa::the+extremism/radicalism
 اليمني::Al+yamiyniy~a::the+rightist/
 right-wing))))

ما اخطر التطرف اليمني

mA + <axo + Tara + AltaTar~ufa + Alyaniyny~a
How dangerous is the right extremism

7.2.8.6 The noun mA /'some' (NOUN)

mA is a noun (ما المبهمة/ mA Almubohamap ‘mA of indefiniteness’) and is annotated as 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 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
He was waiting for something (some thing)

7.2.8.7 The non-inflectional verb mA/ ‘not be’ (VERB)

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 انعتاق::{inoEitAqi::self-liberation
        (NP الفرد::Al+farodi::the+individual ) )
        (PP من::min::from
          (NP قبيلت::qabiyl+at+i-::tribe
            (NP هـ::-hi::his/its))))))))))
```

ما تلك سوى بذور انعتاق الفرد من قبيلته

mA + tiloka + siwaY + bu*uwru + <inoEitAqi + Alfarodi + min + qabiylati+hi
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:::-Al+taman~iy:::the+wish/hope))))

ما نيل المطالب بالتمني

mA + nayolu + AlmaTAlibi + bi+Altaman~iy
Reaching goals is not (just) by wishing

7.2.8.8 The particle mA (PART)

mA is a particle and annotated as 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.

[Some discussion is still needed about the real relationship of the adjective with the verbal phrase – policy for POS tag and tree structure both in question]

(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 غالباً:::gAlibAF:::generally/commonly/typically
 (PRT ما:::mA)
 (VP تعتمد:::ta+Eotamid+u:::it/they/she+depend/rely_[on]+[ind.]
 (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 لِدفاع:::-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 the case that monetary institutions in a number of countries employ high interest rates to protect stock market prices

7.2.9 <i*A/إذا / 'if', 'whether', 'suddenly'

<i*A has two POS values: subordinating conjunction and particle, and two matching syntactic annotations.

7.2.9.1 The subordinating conjunction <i*A 'if,' 'whether' (SUB_CONJ)

<i*A is a subordinating conjunction annotated as SUB_CONJ when it heads a subordinate clause SBAR. It can occur in the following contexts:

- when it introduces **adverbial** subordinating clauses with a conditional meaning where <i*A can be used by itself or along with the subordinating conjunction mA.

(S (SBAR-ADV إذا::-<i*A::if/whether
 (S (VP كان::kAn+a::be/was/were+he/it_[verb]
 (ADVP-PRD هناك::hunAka::there)
 (NP-SBJ (NP حاجة::HAj+apN::need)
 (PP الى::<ilaY::to/towards
 (NP استخدام::isotixodAmi::usage/utilization
 (NP القوة::Al+quw~api::the+violence))))))
 (PRT ف::fa-::and/so)
 (VP ليكن::-li+ya+kun+o::let+him/it+be+[jus.]
 (NP-SBJ *)))

إذا كان هناك حاجة إلى استخدام القوة فليكن.

<i*A + kAna + hunAka + HAjapN + <ilaY + <istixodAmi + Alquw~api + fa + li+yakuno

If there is a need to use violence, so be it.

(S (VP ان::<in~a::indeed
 (NP-SBJ الشريعة::Al+\$ariyEapa::the+Sharia_[Islamic_law])
 (ADJP-PRD حسنة::Hasan+ap+N::good+[fem.sg.]+[indef.nom.])
 (SBAR-ADV اذا::<i*A::if/whether
 (S (VP استعملت::AusotuEomil+at::nogloss
 (NP-SBJ-1 *)
 (NP-OBJ-1 *)
 (PP-MNR على::EalaY::on/above
 (SBAR ان::>an~a-::that
 (S (NP-SBJ ها::-hA::it/they/she)
 (NP-PRD شريعة::\$ariyEapN::Sharia_
 [Islamic_law]))))))))

إنَّ الشريعةَ حَسَنَةٌ إذا أُسْتُعْمِلتْ عَلَى أَنَّهَا شَرِيعَةٌ

<in~a + Al\$ariyEapa + HasnapN + <i* + >usotuEmilat + EalaY + >an~ah+A + \$ariyEapN
 The Shari'a is good if it is used as such.

(S (SBAR-ADV اذا::<i*A::if/whether ما::mA::that
 (S (VP اراد::>arAd+a::want/desire/intend+he/it_[verb]
 (NP-SBJ البعض::Al+baEoDu::the+some/several)
 (NP-OBJ (NP المجازفة::Al+mujAzafapa::the+recklessness)
 و::wa-::and
 (NP (NP الإقامة::-Al+<iqAmapa::the+residency)
 (PP في::fiy::in
 (NP (NP اي::>ay~i::any+[def.gen.]
 (NP مكان::makAn+K::place))))))
 (PRT ف::fa-::and/so)
 (VP ليتحملوا::-li+ya+taHam~al+uWA::let+to+assume_[reponsibility]+[masc.pl.]
 (NP-SBJ *)
 (NP-OBJ مسؤولية::maso&uwliy~apa::responsibility
 (NP انفس::>anofus+i-::selves+[def.gen.]
 (NP هم::-him::their))))))

إذا ما أراد البعض المجازفة والإقامة في مكان ما فليتحملوا مسؤولية أنفسهم.

<i*A + mA + >arAda + AlbaEoDu + AlmujaZafapa + wa+Al<iqAmapa + fiy + >ay~i + makAnK + fa+li+yataHam~alwA +
 maso&wly~apa + >anfusihim

If some want to take on the risk and reside in any place, then let them assume their own responsibility.

- when it introduces an SBAR **complement** of the verb whether by itself or along with the subordinating conjunction mA.

(S VP (PRT لا:: -lA:: no/not/non-)
ندري:: na+doriy+null:: we+know/be_aware_of/notice
(NP-SBJ *)
(SBAR إذا:: <i*A:: if/whether
(S (PP-PRD في:: fiy:: in
(NP الامكان:: Al+<imokAni:: the+possibility)
(NP-SBJ تحسين:: taHosiynu:: improving/making_better
(NP الجباية:: Al+jibAyapi:: the+tax/duty))

لا ندري اذا في الإمكان تحسين الجباية.

lA + nadoriy + <i*A + kAna + fiy + Al<imokAni + taHosiynu + AljibAyapi
We don't know if an improvement of the tax is possible

(S (VP (PRT لا:::-lA:::no/not/non-)
 يستطيع:::ya+sotaTiyE+u:::he/it+be_able/be_capable+[ind.]
 (NP-SBJ احد:::>aHadN:::one)
 (SBAR ان:::>an:::to
 (S (VP يقول:::ya+quwl+a:::he/it+say+[sub.]
 (NP-SBJ *)
 (SBAR ما:::mA:::that اذا:::<i*A:::if/whether
 (S (S (VP كنا:::kunnA:::be/was/were+we
 (VP سننجح:::sa+na+nojaHu:::will+we+succeed
 (NP-SBJ *)
 او:::>aw:::or
 (S (VP سنفشل:::sa+na+fo\$alu:::will+we+fail
 (NP-SBJ *))))))
 (PP-TMP الى:::<ilaY:::to/towards
 (SBAR ان:::>an:::to
 (S (VP تنظر:::ta+noZur+a:::it/they/she+
 look/observe+[sub.]
 (NP-SBJ (NP محكمة:::maHokam+apu:::
 court/tribunal
 (NP استئناف:::
 {isoti}onAfK:::
 resumption/appeal))
 (PP-LOC في:::fiy:::in
 (NP بروكسيل:::
 bruwkosiyl:::
 Brussels)))
 (PP-CLR في:::fiy:::in
 (SBAR ما:::mA:::what
 اذا:::<i*A:::if/whether
 (S (VP كان:::kAn+a:::
 was +he/it
 (PP-PRD ل:::li-
 :::for/to
 (NP بلجيكا:::
 -bilozjiykA:::
 Belgium))
 (NP-SBJ (NP الحق:::
 Al+Haq~u:::

the+right)
 (PP في::fiy::in
 (NP مقاضات::
 muqADAati-::
 court_trial
 (NP ه::-hi::
 his/its
)))))))

لا يَسْتَطِيعُ أَحَدٌ أَنْ يَقُولَ مَا إِذَا كُنَّا سَنَنْجِحُ أَوْ سَنَفْشَلُ إِلَى أَنْ تَنْظُرَ مَحْكَمَةُ الإِسْتِنَافِ فِي بَرُوكْسَالِ مَا إِذَا كَانَ لِيُجْلِبِكَ الحَقُّ فِي مُقَاضَاتِهِ

lA + yastaTyEu + >aHadN + >an + yaquwla + mA + <i*A + kun~A + sananjaHu + >aw + sanaf\$alu + <ilaY + >an + tanoZura +
 maHokamapu + <istinAfK + fiy + bruwksAl + mA + <i*A + kAna + li+byljyKA + AlHaq~u + fiy + muqADAti+hi

Nobody can tell whether we will succeed or lose until an appeal court in Brussels decides whether Belgium has the right to pursue him.

7.2.9.2 The particle <i*A/ ‘suddenly’ (PART)

<i*A is a particle and annotated as PART when it conveys the meaning of “suddenly.” It heads a main clause. This <i*A is referred to in Arabic as the <i*A of surprise (إذا الفجائية <i*A Alfuja} iy~ap).

wa+fiy + nihA}iy~i + AlsibAqi + <inodafaEat + ka+Alqunbulapi + li+tusaj~ila + fiy + Al>amtAri + Al>uwIY + 28.26 + vAniyapF, fa+<i*A + bi+Almutafar~ijiyina + yanohaDuwna + jamiyEAF + wa+yuwAkibuwna+hA + bi+ta\$jyEK + SAxib
And at the end of the race, she propelled like a bomb to achieve in the first meters 28.26 seconds, then suddenly the spectators all rose and accompanied her sprint with an intense cheer.

N.B. The above annotation needs to be revised to include the new POS value of bi- .

7.2.10 <i*/ذ/ 'because', 'as', 'suddenly'

<i* has 2 POS values and 2 matching syntactic representations.

7.2.10.1 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 ذيك*`lika::that)
  (VP شكا::$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 ي :: iy::
 my))
 و::wa-::and
 (NP زوجت::
 -zawojati-::
 wife
 (NP ه::-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 قيد ا::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)))))))))

كذلك شكَا خالد من أن " الطريقة التي قبض فيها على شقيقي وزوجته كانت وحشية" إذ قيذا 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

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

Two more nuances in meaning can also be added to the bare explanatory one:

- a meaning of simultaneity

(S و::wa-::and
 (SBAR-ADV إذ::-<i*::because/since/in_view_of_the_fact_that
 (S (VP شدد::\$ad~ad+a::intensify/emphasize+he/it_[verb]
 (NP-SBJ *)
 (PP-CLR على::EalaY::on/above
 (NP قيام::qiyAmi::undertaking/carrying_out
 (NP دولة::dawolapi::state/country
 (NP المؤسسات::Al+mu&as~asAti::
 the+institutions+[fem.pl.]
 و::wa-::and
 القانون::-Al+qAnuwni::
 the+law/statutes))))))
 (VP اعلن::>aEolan+a::announce/declare+he/it_[verb]
 (NP-SBJ *)
 (NP-OBJ (NP وضع::waDoEa::laying_down/putting/placing
 (NP نفس::nafos+i-::same/self+[def.gen.]
 (NP ه::-hi::its/his))
 (PP في::fiy::in
 (NP خدمة::xidom+api::service/assistance
 (NP الرابطة::Al+rAbiTapi::the+league/union))))))

وَإِذْ شَدَّدَ عَلَى قِيَامِ دَوْلَةِ الْمُؤَسَّسَاتِ وَالْقَانُونِ، أَعْلَنَ وَضَعَ نَفْسِهِ فِي خِدْمَةِ الرَابِطَةِ

wa <i* \$ad~ada EalY qiyAmi dawolapi Almu&as~asAti wa AlqAnuwni, >aElana waDEa nafsihi fiy xidmapi AlrAbiTapi

And while he stressed the implementation of a state founded on institutions and law, he also declared that he is at the service of the league.

- and a causative meaning:

(S (VP (PRT لم::lam::did_not)
 تستمر::ta+sotamir~+a::it/they/she+continue/last_[time]+[jus.]
 (NP-SBJ *)
 (NP-ADV طويلاً::TawiyLAF::for_a_long_time/at_length/extensively)
 (PP-CLR في::fiy::in
 (NP المدرسة::Al+madorasapi::the+school))
 (SBAR-PRP إذ::<i*::because/since/in_view_of_the_fact_that
 (S (VP تزوجت::tazaw~aj+at::get_married+it/they/she_[verb]
 (NP-SBJ *)
 (PP-CLR من::min::from
 (NP (NP (NP نسيب::nasiybi-::relative/kinsman
 (NP ها::-hA::its/their/her)
 (NP حنا::Han~A::Hannah
 صالح::SaliH::Salih/Saleh
 كرم::karam::Karam))
 (ADJP الآتي::Al+|tiy::the+coming/following
 (PP-LOC من::min::from
 (NP المهجر::Al+mahojari::
 the+overseas/
 in_exile/diaspora))))))
 (NP-TMP عام::EAmA::year
 (NP 1896::1896::nogloss))))))

لم تستمر طويلاً في المدرسة إذ تزوجت من نسيبها حنا صالح كرم الآتي من المهجر عام 1896

lam + tasotamir~a + TawilAF + fiy + Almadorasapi + <i*o + tazaw~ajat + min + nasiybi+hA + Han~A + SaliH + karam +
 Al|Atiy + min + Almahojari + EAmA + 1896

She did not stay long in school because she married in the year 1896 her relative Hana Salih Karam, who returned from overseas

Syntactically, <i* in its general adverbial usage can also be followed by <an~a which is annotated as SUB_CONJ:

(S wa-::and
 (VP اعتبرت::-{iEotabar+at::consider/regard/believe+it/they/she_[verb]
 (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 ه::-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 ه::-hu::it/him))
 (NP-OBJ تسمية::tasomiyapa::
 designation/naming
 (NP القذافي::Al+qa*~Afiy~::
 the+Qaddafi))))))))))

واعتبرت أن رئيس الوزراء ذهب بعيدا 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

And she considered that the prime minister went too far in his official report, since Musa did not ask him to mention Gaddafi.

7.2.10.2 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.’

See section 7.2.9.2 on the particle <i*A ‘suddenly’ (PART) for examples.

7.2.11 turaY/تري / ‘I wonder’

turaY has 2 POS values and 2 corresponding Treebank annotations.

7.2.11.1 The interjection turaY (INTERJ)

turaY is an interjection annotated as INTERJ when it occurs as an exclamation and expresses the meaning of doubt. Syntactically, it is not bound to any other constituent that may follow it. It occupies a first position in affirmative sentences as well as questions.

```
(SQ (INTJ turaY تُرى )
  (PRT hal هل )
  (VP saya>otiy سيأتي
    (NP-SBJ (NP Alyawomu اليوم )
      (SBAR (WHNP-1 Al~*iy الذي )
        (S (VP <arA أرا
          (NP-SBJ * )
          (NP-OBJ ka ك )
          (PP-TMP fi في
            (NP (NP hi ه )
              (NP-1 *T* ) ) )
          (PP min من
            (NP jadiyd جديد ) ) ) ) ) ) ) )
  ? )
```

تري هل سيأتي اليوم الذي سأراك فيه من جديد؟

turaY + hal + saya>otiy + Alyawomu + Al~*iy + <arA ka + fi+hi + min + jadiyd ?
I wonder, will there be the day when I can see you again?

7.2.11.2 The non-inflectional verb turaY ‘I wonder’ (VERB)

turaY is a non-inflectional verb annotated as VERB when it selects a subject and a predicate. It is usually followed by a bound pronoun which is its subject and a non-verbal predicate or an S complement.

```

(S (VP turA تُرا
  (NP-SBJ-1 hu هُ )
  (S (VP y>otiy يأتي
    (NP-SBJ-1 * )
    (NP-TMP Alyawoma اليَوْمَ ) ) ) )
. )

```

تراه يأتي اليوم

turA+hu + ya>otiy + Alyawoma
Maybe he will come today.

7.2.12 Syntactic Annotation of Particles

All particles are annotated under a PRT node in TB. Depending on whether they modify a particular constituent inside the sentence or whether they modify the sentence as a whole, particles are annotated at the sentence level or inside one of its constituents. Based on this criterion, a classification of particles is provided in the sections below.

7.2.12.1 Particles annotated under constituents

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

7.2.12.1.1 Vocative particles (VOC_PART)

They are annotated under an NP-VOC node.

(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 | أَنْ
 (S (NP-TPC-4 honor/distinction |\$arafa | شَرَفَ
 (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 | إِنْ
 on/above |EalaY | عَلَى
 (NP hands |>ayodiy | أَيْدِي
 (NP the+believer |Almu&ominyana |
 الْمُؤْمِنِينَ))))))))))
 اعلم يا ابا عمَّار 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

7.2.12.1.2 Emphatic particle (EMPHATIC_PART)

The emphatic particle لا (la-/definitely) can cliticize to 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.

7.2.12.1.3 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 than a verb.

(S (VP (PRT لن:::-lan:::[will]_not/never)

يَحْلُمُوا:::ya+Holum+uwA:::the+dream+[masc.pl.]

(NP-SBJ (-NONE- *))

(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 (-NONE- *))

(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))

لَنْ يَحْلُمُوا بِالْأَمْنِ حَتَّى يَرْفَعُوا أَيْدِيَهُمْ عَنِ أُمَّتِنَا

lan + yaHlumuwA + bi+Al>amni + Hat~Y + yarfaEuwA + >aydiy+him + Ean + >um~ati+nA

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

7.2.12.1.4 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 و::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

7.2.12.1.5 Other Particles

The restrictive particle إنا (<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 التفاف::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 (-NONE- *T*))
 (NP-PRD حقيقة::Haqiyq+ap+F::truth/reality+[fem.sg.]+[indef.acc.]
 (NP-TMP (PRT إنا::<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 إلى::<ilaY::to/towards
 (NP اصحاب::>aSoHAb+i-::owners/originators
 ه::-hi::its/his)))

سَرَابُ النَّفَاوُضِ لَنْ يُصْبِحَ حَقِيقَةً إِلَّا بَعْدَ عَوْدَةِ الْجَوْلَانِ السُّورِيِّ الْمُحْتَلِّ

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.

The particle Hat~Y and >ay modify any phrase other than VPs.

The particle rub~a is always annotated under an NP node.

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

The particle <iy~A is always attached to object pronouns; and therefore, is always under an NP-OBJ or NP-DTV nodes.

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.

7.2.12.2 Particles Annotated under the Sentence Level

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

7.2.12.2.1 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 | هل |
(NP-SBJ (NP measures/steps |<ijorA'Atu | إجراءات |
(NP the+government/administration |AlHukuwmapi | الحكومة |
(ADJP the+taken/adopted |Almut~axa*api | المتخذة |
(ADJP-PRD guaranteeing |kafiylapN | كفيلاً |
(PP by/with |bi- | بـ |
(NP treatment/therapy/processing |-muEAlajapi | معالجة - |
(NP the+conditions/situation/status |
Al>awoDAEi | الأوضاع |
the+economic/economical |

The particles (PART), 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) (PART) 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) (PART) modifies the proposition it introduces.

```
(S (SBAR-ADV ان ::-<in::if/whether
  (S (S (VP (PRT لم ::lam::did_not)
    جمع ::yu+jomaE+o::he/it+be_agreed_unanimously+[jus.]
    (NP-SBJ العرب ::Al+Earabu::the+Arabs)))
  و ::wa-::and
  (S (VP (PRT لم ::-lam::did_not)
    يوحدوا ::yu+waH~id+uwA::they+unite/unify
    (NP-SBJ *)
    (NP-OBJ كلمت ::kalim+at+a-::word/remark/speech
      (NP هم ::-hum::their))
    )
  )
  )
  )
```

(PRT ف ::fa-::and/so)

```
(NP-TPC-2 القضية ::-Al+qaDiy~apu::the+problem/issue
  الفلسطينية ::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 كَلِمَتَهُمُ فَالْقَضِيَّةُ الْفِلِسْطِينِيَّةُ لَهَا عِدَّةٌ أَوْجُهٍ

<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.

```
(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.

7.2.12.3 Summary Chart

The summary chart below is provided as a quick reference for annotators as per the distribution of particles in TB.

7.2.12.3.1 Particles under Constituents

Particles	Buckwalter	Level of annotation
أ	>a	Inside NP
أي	>ay	Inside NP
أيها	>ay~uhA	Inside NP
أيها	>ay~atuhA	Inside NP
أيا	>ayA	Inside NP

إِذَا	<iy~A	Inside NP
مِنْ	min	Inside NP
رُبَّ	rub~a	Inside NP
يَا	yA	Inside NP or PP
قَدْ	qad	Inside VP
لَقَدْ	laqad	Inside VP
سَوْفَ	sawfa	Inside VP
سَاءَ	sa-	Inside VP
لِ-	li-	Inside VP
لَمْ	lam	Inside VP
لَمَّا	lam~A	Inside VP
مَا	mA [NEG_PART]	Inside VP
لَنْ	lan	Inside VP
أِلَّا	>aIA	Inside VP
لِ-	la- [EMPHATIC PART]	Inside VP,ADVP, NP, PP
لَا	IA	Inside VP, NP, PP,ADVP
إِلَّا	<iI~A	Inside NP, PP, VP, ADJP, SBAR
أَيَّ	>ay	Inside NP, PP, ADJP, SBAR,
حَتَّى	Hat~Y	Inside NP, PP, ADJP, SBAR
إِنَّمَا	<in~amA	Inside NP, PP, ADJP, SBAR
سِوَى	siwY	Inside PP
وَ-	wa-	Inside SBAR, S
بِ-	bi-	Inside NP
وَأَ-	wA-	Inside NP
أَ-	-Ah	Inside NP

7.2.12.3.2 Particles under the Sentence

Particles	Buckwalter	Level of annotation
إِنَّمَا	<in~amA	Inside S
لِ-	la- [RC PART]	Inside S
أَمَّا	>am~A	Inside S
فَ-	fa- [RC PART]	Inside S
فَ-	fa- [CONNEC_PART]	Inside S
مَا	mA [PART]	Inside S
كَأَنَّمَا	ka>an~amA	Inside S
إِذْ	<i~*	Inside S
إِذَا	<i~*A	Inside S
هَلْ	hal	Inside SQ
أَ	>a	Inside SQ

7.2.13 >akvar/ أكثر and >aglab/أغلب

>akvar and >aglab have two POS values: They are comparative adjectives and noun quantifiers. As such, they also have two distinct syntactic annotations. Although this section takes >akvar as an illustration, the same tests apply to >aglab.

Note on case marking of >akvar/>aglab: note that relative forms in Arabic with the pattern أفعل />afoEal (صيغة التفضيل) never take tanween or *kasra*, but a *fatha* instead (See examples throughout this section.).

7.2.13.1 The noun quantifiers >akvar ‘most of’ and >aglab ‘majority of’ (NOUN_QUANT)

>akvar and >aglab are noun quantifiers annotated as 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 one 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+>amiyrokisy~i::the+American, ::,::nogloss
          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}iy1u + EalaY + kAmili + Almaoqifi + Al>amoriykyi~i + wa+<akovara + Almawoqifi + Al<uwruwbiy~i.

After 9/11 Israel acquired the full American position and most of the European position.

7.2.13.2 The comparative adjectives >akvar ‘more’/’more than’/’most’ and >aglab ‘most’ (ADJ_COMP)

When >akvar carries a meaning of comparison, it is a comparative adjective that should be tagged as ADJ_COMP. As such, >akvar can occupy the following syntactic positions:

- Modifier of a noun

When it modifies a noun, >akvar is headed by an ADJP and can take either an NP complement in accusative (tamyyiz) by itself or followed by a PP complement:

(NP (NP Alrajulu الرَّجُلُ)
 (ADJP Al>akovaru الْأَكْثَرُ
 (NP-ADV juro>apF جُرْأَةً)))

الرجل الأكثر جرأة

Aljuro>u + Al>akovaru + juro>ap
 The man + the most + audacity
The most audacious man

(NP (NP Alrajulu الرَّجُلُ)
 (ADJP Al>akovaru الْأَكْثَرُ
 (NP-ADV juro>apF جُرْأَةً)
 (PP min مِن
 (NP gayori عَيْرُ
 (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

- Predicate

(PP-PRP ل::li-::for

(SBAR 0
 (S (VP يكون::-ya+kuwn+a::him/it_to+be+[sub.]
 (NP-SBJ النقاش::Al+niqA\$u::the+debate)
 (ADJP-PRD اكثر::>akovara::more
 (NP-ADV فاعلية::fAEily~apF::effectiveness/activity))

لِيَكُونَ النِّقَاشُ أَكْثَرَ فَاعِلِيَّةً

li+yakwna + AlniqA\$u + >kovara + fAEily~apF
 for+be + the debate + more + effectiveness
so that the debate becomes more effective

- Adverbial/Sentential modifier

When it comes as a single word in adverbial positions, >akvar is headed by an NP with an adverbial dashtag and is always in the accusative case.

(S (ADVP-PRD هناك::hunAka::there_is/are)
 (NP-SBJ (NP عوامل::EawAmilu::factors/agents
 أخرى::>uxoraY::other/another/additional)
 (SBAR (WHNP-1 (-NONE- *0*))
 (S (VP عَقَدَت::Eaq~adat::complicate+it/they/she_[verb]
 (NP-SBJ-1 (-NONE- *T*))
 (NP-OBJ الوضع::Al+waDoEa::the+situation)
 (NP-ADV اكثر::>akovara::more))))))

هُنَاكَ عَوَامِلُ أُخْرَى عَقَدَتِ الْوَضْعَ أَكْثَرَ

hunAka + EawAmilu + >uxoraY + Eaq~adat + AlwaDEa + >akovara
 there + factors + others + complicated + the situation + more
There are other factors that complicated the situation more.

When it takes a PP complement, >akvar will be headed by ADJP headed by an NP with an adverbial dashtag:

(S (NP-TPC-2 (NP الزيارة::Al+ziyArapu::the+visit)
 (PP ل::li::for/to
 (NP بيروت::-bayoruwta::Beirut)))

(VP (PRT لن::lan::[will]_not/never)
 تستمر::ta+sotamir~+a::it/they/she+continue/last_[time]+[sub.]
 (NP-SBJ-2 (-NONE- *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

- Nominal positions

In subject, object, complement of a noun or complement of a preposition positions, akvar is headed by an NP, when it's a single word, and by ADJP headed by NP when it is a multi-word phrase.

Note: semantically, it's almost impossible to find akvar single word occupying the above mentioned nominal positions.

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 أكثر::>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

العربي::Al+Earabiy~i::
 the+Arab)
 و::wa-::and
 (NP اوروبا::->uwruwb~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)

تَحْتَهُ هَذَا الْعُنْوَانِ اجْتَمَعَ أَكْثَرُ مِنْ 500 طَبِيبٍ وَإِخْتِصَاصِيٍّ مِنَ الْعَالَمِ الْعَرَبِيِّ وَأُورُوبَا وَآمْرِيكَا فِي الْعَاصِمَةِ الْأُرْدُنِيَّةِ عَمَّانَ

taHota + ha*A + AlEunwAni + AijotamaEa + >akovaru + min + 500 + TabybK + wa+AixotiSASy~K + min + AlEAlami +
 AlEaraby~i + wa+Awrb~A + wa+AmorykA fy + AlEASimapi + Al>urodun~api + Eam~An
 under + this + title + gathered + more + of + 500 + doctors + specialist + from + the world + the Arab + and+Europe + and+Amrica +
 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
 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 (-NONE- *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))))))

هَذِهِ الْمَآسَاةُ الْإِنْسَانِيَّةُ تُطَاوَلُ أَكْثَرَ مِنْ بَيْتِ كُوَيْتِيٍّ وَ عَرَبِيٍّ

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 a Kuwaiti and Arab family

Complement of a preposition

```
(S (VP مارس::mAras+a::practice/pursue/exercise+he/it_[verb]
  (NP-SBJ (-NONE- *))
  (NP-OBJ الطب::Al+Tib~a::the+medicine
  (PP-TMP ل::li-::for/to
    (NP (ADJP اكثر::->akovar+a::more/most+[indef.gen.]
      (PP من::min::from
        (NP 25::25::nogloss)
        (NP عاماً::EAm+AF::year+[acc.indef.])))
```

مارسَ الطيبَ لأكثرَ من 25 عاماً

mArasa + AlTib~a + li+>akovara + min + 25 + EAmAF
practice + the medicine + for+more + of + 25 + year
He was a doctor for more than 25 years

Complement of a noun

```
(PP في::-fiy::in
  (NP رأي::ra>oyi::opinion/view/idea
    (NP (ADJP اكثر::>akovar+a::more/most+[def.gen.]
      (PP من::min::from
        (NP مصدر::maSodarK::source
          فلسطيني::filasoTiydiy~+K::Palestinian+[indef.gen.])))
```

في رأي أكثرَ من مصدرَ فلسطينيِّ

fy + ra>oyi + >akovara + min + maSodarK + falasoTiy~K
in + opinion + more + of + source + Palestinian
In the opinion of more than one Palestinian source

8.0 LINGUISTIC INFORMATION

8.1. Constructions of Numerals

The different constructions for cardinal and ordinal numbers are presented because they assist annotators in determining the correct decision as far as the vocalization of certain words in those constructions is concerned.

8.1.1 Cardinal numbers: NOUN_NUM

Pre-nominal cardinal numbers occur in two constructions: Construct state and tamyyz constructions. Post-nominal cardinal numbers occur in an apposition construction. Information about the constructions where cardinal numbers occur provides support to POS annotation with respect to the marking of features such as case, gender, and number.

8.1.1.1 Prenominals

Cardinal numbers in pre-nominal positions occur in two constructions: the construct state and the tamyyz constructions.

8.1.1.1.1 Construct State Structure (<iDAfap/إضافة)

Cardinal numbers from 3 to 10 are heads of a construct state structure and have a noun complement in genitive.

Arabic	Buckwalter	POS Tag	Gloss	Example	POS Annotation	Transliteration	Gloss	Translation
ثلاثة	valAvap	NOUN_NUM	Three	ثلاثة كتب	NOUN_NUM+ NOUN+NSUFF_PL+CASE_GEN	valAvapu kutubK	Three books	Three books
خَمْسَة	xamsap	NOUN_NUM	Five	خمس ساعات	NOUN_NUM+ NOUN+NSUFF_PL+CASE_GEN	xamsu sAEAtK	Five hours	Five hours
سِتَّة	sit~ap	NOUN_NUM	Six	ستة طلبة	NOUN_NUM+ NOUN+NSUFF_PL+CASE_GEN	sit~apu TalabapK	Six students	Six students
أرْبَعَة	>arbaEap	NOUN_NUM	Four	الأربع بنات	NOUN_NUM+ NOUN+NSUFF_PL+CASE_GEN	AlarbaEapu banAtK	The four girls	The Four girls
مائة	mA}ap	NOUN_NUM	One hundred	مائة كتاب	NOUN_NUM+ NOUN+NSUFF_SG+CASE_GEN	mA}apu kitAbK	Hundred book	One hundred book(s)
ألف	>alf	NOUN_NUM	Two	ألفا ساعة	NOUN_NUM+	>alfA sAEapK	Two	two thousand

			thousands		NOUN+NSUFF_SG+CASE_GEN		thousand hour	hour(s)
مليون	milywn	NOUN_NUM	One million	مليون طالب	NOUN_NUM+ NOUN+NSUFF_SG+CASE_GEN	milywnu TALibK	Million student	One million student(s)

8.1.1.1.2 tamyyz (تَمييز) Structure

The tamyyiz construction is a construction that involves cardinal numbers not ordinal numbers. tamyyz is a complement structure whereby a cardinal number has a noun complement in the accusative. Only numbers from 11 to 99 can head a tamyyz construction.

Arabic	Buckwalter	POS Tag	Gloss	Example	POS Annotation	Transliteration	Gloss	Translation
أحد عشر	>aHada Ea\$ar	NOUN_NUM	Eleven	أحد عشر كتابا	NOUN_NUM + NOUN+NSUFF_SG+CASE_ACC	>aHada Ea\$ara kitAbAF	Eleven book	Eleven book(s)
أربع وعشرون	>arbaEuN wa Ei\$orwn	NOUN_NUM	Twenty four	أربع وعشرون ساعة	NOUN_NUM + NOUN+NSUFF_SG+CASE_ACC	sAEapF	Twenty- four hour	Twenty four hour(s)
ستة وثلاثون	sit~apN wa valAvwn	NOUN_NUM	Thirty six	ستة وثلاثون طالبا	NOUN_NUM + NOUN+NSUFF_SG+CASE_ACC	sit~apN wa valAvwna TalibAF	Sixty- three student	Thirty six student(s)

8.1.1.2 Postnominals

Postnominal cardinal numbers occur in a construction called Apposition (بتدل) when they are preceded by a plural noun. There is reverse gender agreement with that noun.

Arabic	Buckwalter	POS Tag	Gloss	Example	Pos Annotation	Transliteration	Gloss	Translation
عَشْرَةٌ	Ea\$rap	NOUN_NUM	Ten	ليال عشر	NOUN+NSUFF_PL+CASE_GEN +	layAlK Ea\$rK	Nights ten	Ten nights
ثلاثة	valAvap	NOUN_NUM	Three	كتب ثلاثة	NOUN+NSUFF_PL+CASE_GEN	kutubN valAvapN	Books three	Three books
أربعة	>arbaEap	NOUN_NUM	Four	غرف أربع	+ NOUN_NUM+NSUFF_CASE_GEN	gurafN >arbaEN	Rooms four	Four rooms

ثلاثونَ	valAvuwna	NOUN_NUM	Thirty	الصفحات الثلاثون الأخيرة	NOUN+NSUFF_PL+CASE_GEN	AlSafaHatu AlvalAvuwna Al>axiyrapu	The pages the thirty the last	The last thirty pages
---------	-----------	----------	--------	--------------------------------	------------------------	--	--	--------------------------

8.1.2 Ordinal numbers: ADJ_NUM

Prenominal ordinal numbers occur in a false construct state structure whereas post-nominal ordinal numbers are adjoined to the noun they modify.

8.1.2.1 Prenominals

These ordinals have the pattern fAEiIN (فاعل) or >afEal (أفعل) for number one. They occur in a false construct state construction. Prenominal adjectives are restricted to numbers from 1 to 10.

Arabic	Buckwalter	POS Tag	Example	Transliteration	POS Annotation	Gloss	Translation
ثالث	vAliv	ADJ_NUM	ثالث الكتب	vAlivu Al kutubi	ADJ_ NUM + NOUN+NSUFF_DEF_PL +CASE_GEN	Third book (s)	The third book
خامس	xAmis	ADJ_NUM	خامس الغرف	xAmisu Algurafi	ADJ_ NUM + NOUN+NSUFF_DEF_PL +CASE_GEN	Fifth room (s)	The fifth room
أول	>aw~al	ADJ_NUM	أول يوم	>aw~alu yawmK	ADJ_NUM+ NOUN+NSUFF_INDEF_ SG+CASE_GEN	First day	

--	--	--	--	--	--	--	--

8.1.2.2 Postnominals

They occur in an adjunction construction. The adjective agrees in number and case with the noun preceding it.

Arabic	Buckwalter	POS Tag	Example	Transliteration	Gloss	Translation
أول	>aw~al	ADJ_NUM	اليوم الأول	Alyawmu Al>aw~alu	The day first	The first day
العشرون	AlEi\$ruwn	ADJ_NUM	الصفحة العشرون	AlSafHapu AlEi\$rwana	The page the twentieth	The twentieth page
خامسة و الثلاثون	xAmisap wa valAvuwn	ADJ_NUM	الذكرى الخامسة و الثلاثون	Al*ikraY AlxAmisapu wa AlvalAvwna	The commemoration the thirty-fifth	The thirty-fifth commemoration
مائة و خمسون	mA}ap wa xamsuwn	ADJ_NUM	السنة المائة و الخمسون	Alsanapu AlmA}apu wa Alxamsuwna	The year the hundredth and the fiftieth	The hundred and fifthyeth year

9. REFERENCES

Badawi, E., Carter, M., Gully, A. (2004). Modern written Arabic: A comprehensive grammar. USA: Routledge

Blasher, R. and Gaudefroy-Demombynes, M. (1975). Grammaire de l'Arabe classique (A grammar of classical Arabic). Paris: G-P. Maisonneuve et Larose

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

Fisher, W. (2002). A grammar of classical Arabic (3rd Ed.). London: Yale University Press

Qabbish, A. (1984). الكامل في النحو و الصرف و الإعراب
(The complete in syntax, morphology, and case/parsing). Beirut, Lebanon: Daar Aljiyl

Radford, A. (1997). Syntactic theory and the structure of English: A minimalist approach. Cambridge: CUP

Ryding, K. (2005). A reference grammar of Modern Standard Arabic. Cambridge: CUP

Sayyd Ahmed, A., and Eabd Alganiiy, I. (?). الشرح المفصل لإبن يعيش (A detailed grammar analysis of Ibn yaEysh) . Egypt: Almaktabap
Altawfiyy~ap

Van Halteren, H. (1999). Syntactic wordclass tagging. Netherlands: Kluwer Academic Publishers