XIP Spanish Grammar

Author: E. López García
XRCE
9/26/2006

Introduction

The aim of this document is to explain category, chunk and dependency names that are most widely used in the XIP Spanish Grammar.

Chunks are basic syntactic domains that are computed by the Spanish grammar and that serve as a basis for the dependency extraction.

A XIP dependency is a n-ary relation (unary, binary) that binds one or two nodes together from the chunk tree. Usually, dependencies apply on lexical nodes. Each dependency is named with a specific label. A dependency label denotes the grammatical relation that has been computed between the nodes that have been bound together.

1. Categories & Features

A category in a XIP grammar must be declared as a set of features (attribute-value pairs).

The PoS (part-of-speech) features are distinguished from other features.

Lexical categories and feature names do not display on screen in the default output mode. However, in the XIP API all the linguistic information (e.g., categories and features) is available within the XipResult object.

1.1 Part-of-Speech Features

ADJ	adjective
ADV	adverb
ART	article
CONJ	conjunction
DET	determiner
INTERJ	interjection
NOUN	common or proper noun
NUM	numeral
PREP	preposition
PRON	pronoun
PUNCT	punctuation
QUANT	quantifier
SYMBOL	special symbol

VERB	verb
MISC	word parts, foreign words, etc.
PREDET	pre-determiner
LOCUT	locution

1.2 Other Features

Features can appear together with Part-of-Speech tags when appropriate.

Here follows the list of the most important features used in the Spanish grammar.

//Morphological features

PL	plural
SG	singular
SP	singular or plural
1p	first person
2p	second person
3p	third person
FEM	feminine
MASC	masculine
MF	feminine or masculine
NEUT	neuter
NOM	nominative
ACC	accusative
DAT	dative
GEN	genitive
REF	reflexive
ENCLIT	enclitic pronoun

//Verbal

SUBJ	subjunctive
FUT	future
IMP	imperative
IMPERF	imperfect
IND	indicative
INF	infinitive
PARTICIPIO	participle
PASSIF	passive form
PASTPART	past participle
PRES	present
PRESPART	present participle
PERF	past simple
PLUPERF	past subjunctive

COND	conditional
ASPECT	aspectual verbal periphrasis
ASPECTGER	aspectual verbal periphrasis (build with gerundive)
ASPECTINF	aspectual verbal periphrasis (build with infinitive)
ASPECTPREP	aspectual verbal periphrasis (build with preposition)
ASPECTPART	aspectual verbal periphrasis (build with past participle)
ASPECTPOT	verb that can build an aspectual periphrasis
MODAL	modal verbal periphrasis
MODALPOT	verb that can build a modal periphrasis
AUX	auxiliary
COP	copulative
CAUS	causative
NONPER	impersonal verbal forms

//Syntactic features for verbs

SUBJ_INF	verbs with infinitive subject
OD_INF	verbs with infinitive direct object
TRANS	transitive
INTRANS	intransitive
DOUBLE	double complement
VERBAL	verbal
PRED	verb that can have a predicative complement

//Semantic features for verbs

PAROLE	verbs of speech (say, think…)
MOUV	verb of movement

//Semantic features for noun

ANIM	animated

//Syntactic features for noun

DE_INF	nouns that can have a complement de + infinitive

//Syntactic features for various categories

PP_A	categories that require a prepositional complement starting by ‘a’
PP_DE	categories that require a prepositional complement starting by ‘de’
PP_EN	categories that require a prepositional complement starting by ‘en’

//Determiners

DEF	definite
INDEF	indefinite
DEM	demonstrative
POSS	possessive
POSTDET	post-determiner
PREDET	pre-determiner

//Numbers

CARD	cardinal
DIG	digit
ORD	ordinal
ROMAN	roman
PERCENT	percentage
DEG	degree
TEL	telephone number
TIME	time

//Punctuations:

ACCENT	accent
BRACKET	bracket
COMMA	comma
DASH	dash
DOTS	dots
LEFT	left (paren or bracket)
PAREN	parenthesis
QUOTE	quotation mark
RIGHT	right (paren or bracket)
SENT	sentence final punctuation
SLASH	slash or backslash

//Pronouns

COMPL	complement of preposition

//Verbal complements

EXP_TEMP	temporal expression
TEMP	temporal
MAN	manner
MEAS	measure
MENTE	adverbs in –mente
REAS	reason

//Conjunctions

COORD	coordinated
ENUM	enumeration

//Derivational morphology

AUM	augmentative
DIM	diminutive
SUP	superlative

//Letters

INIT	initials
LET	letter

LIT	literal letters
//Modifiers	
POS	postponed
PRE	ante posed
//Nouns	
PROP	proper (for proper names)
ORG	organization
SOC	company name
TITLE	title
ZODIAC	zodiac names
CORT	courtesy phrase
CONTINENT	continent
COUNTRY	country
LOC	location
LUGAR	place
USASTATE	USA state
//Sentences	
REL	relative
STC	sentence
SUBORD	subordinated
INTERROG	interrogative
//Various features.	
ABREV	abbreviation
ACRON	acronym
APOC	apocopate
APOS	apposition
COMPAR	comparative
NOT_NP	deep prepositional phrases
ONOM	onomatopoeia
PREPCOUR	current preposition
WORDPART	not a word by itself

2. Chunk Nomenclature

Here follows the list of chunks used in the XIP Spanish Grammar.

AP	adjectival phrase
NP	noun phrase
PP	prepositional phrase
AVP	adverbial phrase
REFLEX	reflexive phrase
FV	finite verb phrase
VPART	past participle phrase
VINF	infinitive phrase
VGER	gerundive phrase
SC	chunk sentence

3. List of dependencies

Below is the list of dependencies extracted by the Spanish XIP parser.

HEAD

This is a binary relation relating the nucleus of some chunk with the chunk itself.

Example
Todos mis vecinos

HEAD(vecinos, todos mis vecinos)

Deben de haber muerto

HEAD(muerto, deben de haber muerto)

DETD

This binary dependency links a nominal head and a determiner.

Example
Mi novio

DETD(novio, mi)

De la huerta

DETD(huerta, la)

Dame el mío

DETD(mío, el)

NUMD

This binary dependency links a nominal head and a numeral.

Example
Dos días

NUMD(días, dos)

QUANTD

This binary dependency links a nominal head and a quantifier.

Example
Todos los días

QUANTD(días, todos)

TETE_MORPH

This binary dependency links the inflected verb and the verbal chunk itself. In simple forms the HEAD and the TETE_MORPH will be the same word.

Example
Has venido

TETE_MORPH(has, has venido)

Debe de haber sido

TETE_MORPH(debe, debe de haber sido)

DOMIN

This binary dependency links the inflected verb to the main verb of a compound form.

Example
Has venido

DOMIN(venido, has)

Acaba de ser nombrado

DOMIN(ser nombrado, acaba)

MAIN

This unary dependency extracts the main verb of a sentence

Example
Ven cuando puedas

MAIN(Ven)

AGENT

This binary dependency links a passive verb (the main one) and the agent.

Example
Fue secuestrado por los soldados

AGENT(secuestrado, soldados)

ATTRIB

This binary dependency links a copulative verb and its attribute.

Example
Es muy importante

ATTRIB(es, importante)

PRED

This binary dependency links a noun and a predicative complement.

Example
María se quedó estupefacta

PRED(María, estupefacta)

SUBJ

This dependency links a verb and its subject.

Example
María duerme

SUBJ(duerme, María)

María ha dormido durante horas

SUBJ(dormido, María)

OBJ

This binary dependency links a verb and its direct object.

Example
Cría cuervos

OBJ(cría, cuervos)

No ha querido venir

OBJ(querido, venir)

IOBJ

This binary dependency links a verb and its indirect object.

Example
Cuenta todo a su madre

IOBJ(cuenta, madre)

MOD

This binary dependency links a noun and any kind of complements or adjuncts attached to this noun. Note that we do not make any distinction between complements and adjuncts.

Example
El gran desastre

MOD(desastre, gran)

El día de la bestia

MOD(día, bestia)

MODADJ

This binary dependency links an adjective and any kind of complements or modifiers of this adjective.

Example
Muy importante

MODADJ(importante, muy)

MODV

This binary dependency links a verb and any kind of complements or adjuncts attached to this verb. Note that we do not make any distinction between complements and adjuncts

Example
Has venido muy deprisa

MODV(venido, deprisa)

Se quedó en París

MODV(quedó, París)

ANT

This binary dependency links the relative pronoun and its antecedent.

Example
La chica que ves es mi hermana

ANT(chica, que)

INTROD

This binary dependency links the article and the relative pronoun in relative sentences introduced by an article.

Example
Quiero el que ves

INTROD(el, que)

COORD

This binary relation links a coordination conjunction to one of the coordinated elements. **All** the coordinated elements of a coordinated expression have to be linked by a COORD dependency.

Example
Fue a cenar, bailar y beber

COORD(y, cenar)
COORD(y, bailar)
COORD(y, beber)

The Dependency Features

A list of suffixes is sometimes added to the dependency labels (described above) in order to give more specific information about the syntactic relation. Here follows the list of the possible suffixes (features).

_PRE

This suffix expresses that the modifier is on the left of the governor in the sentence.

Examples
La gran fiesta tendría lugar el viernes.

MOD_PRE(fiesta, gran)

Siempre duerme.

MOD_PRE(duerme, siempre)

_POS

This suffix expresses that the modifier is on the right of the governor in the sentence.

Examples

La version nueva va a aparecer pronto.

MOD_POS(versión, nueva)

_TEMP

This suffix is added to MODV dependency when the modifier is a temporal expression

Example

Se marchó el viernes.

MODV_TEMP(marchó, viernes)

_STC

This suffix expresses that the modifier is sentential (excluding the case of relatives). The sentential modifier is represented by its main verb.

Example

Se marchó porque quería.

MODV_STC(marchó, quería)

Dijo que quería venir

OBJ_STC(dijo, quería)
OBJ_STC(quería, venir)