This page pertains to UD version 2.

Universal POS tags

These tags mark the core part-of-speech categories. To distinguish additional lexical and grammatical properties of words, use the universal features.

Open class words	Closed class words	Other
ADJ	ADP	PUNCT
ADV	AUX	SYM
INTJ	CCONJ	X
NOUN	DET	
PROPN	NUM	
VERB	PART	
	PRON	
	SCONJ	

ADJ: adjective

Definition

Adjectives are words that typically modify nouns and specify their properties or attributes:

*The **oldest French** bridge*

They may also function as predicates, as in:

*The car is **green**.*

The ADJ tag is intended for ordinary adjectives only. See DET for determiners and NUM for (cardinal) numbers. ADJ is used for "proper adjectives" such as *European*.

Numbers vs. Adjectives: In general, cardinal numbers receive the part of speech NUM, while *ordinal numbers* (more precisely *adjectival* ordinal numerals) receive the tag ADJ.

There are words that may traditionally be called numerals in some languages (e.g., Czech) but which are treated as adjectives in our universal tagging scheme. In particular, the *adjectival* ordinal numerals (note: Czech also has adverbial ones) behave both morphologically and syntactically as adjectives and are tagged ADJ.

Nouns vs. Adjectives: A noun modifying another noun to form a compound noun is given the tag NOUN not ADJ.

Participles: Participles are word forms that may share properties and usage of any of adjectives, nouns, and verbs. Depending on the language and context, they may be classified as any of ADJ, NOUN or VERB.

Adjectival modifiers of adjectives: In general, an ADJ is modified by an ADV (*very strong*). However, sometimes a word modifying an ADJ is still regarded as an ADJ. These cases include: (i) ordinal numeral modifiers of a superlative adjective (*the **third** oldest bridge*) and (ii) when a pair of adjectives form a compound adjectival modifier (*an **African American** mayor*).

Examples

- *big*
- *old*
- *green*

- *African*
- *incomprehensible*
- *first, second, third*

References

- Loos, Eugene E., et al. 2003. Glossary of linguistic terms: What is an adjective?
- Wikipedia

edit ADJ

ADP: adposition

Definition

Adposition is a cover term for prepositions and postpositions. Adpositions belong to a closed set of items that occur before (preposition) or after (postposition) a complement composed of a noun phrase, noun, pronoun, or clause that functions as a noun phrase, and that form a single structure with the complement to express its grammatical and semantic relation to another unit within a clause.

In many languages, adpositions can take the form of fixed multiword expressions, such as *in spite of*, *because of*, *thanks to*. The component words are then still tagged according to their basic use (*in* is `ADP`, *spite* is `NOUN`, etc.) and their status as multiword expressions are accounted for in the syntactic annotation.

Note that in Germanic languages, some prepositions may also function as verbal particles, as in *give **in*** or *hold **on***. They are still tagged `ADP` and not `PART`.

Examples

- *in*
- *to*
- *during*

References

- Loos, Eugene E., et al. 2003. Glossary of linguistic terms: What is an adposition?
- Wikipedia

edit ADP

ADV: adverb

Definition

Adverbs are words that typically modify verbs for such categories as time, place, direction or manner. They may also modify adjectives and other adverbs, as in ***very briefly*** or ***arguably*** *wrong*.

There is a closed subclass of *pronominal adverbs* that refer to circumstances in context, rather than naming them directly; similarly to pronouns, these can be categorized as interrogative, relative, demonstrative etc. Pronominal adverbs also get the `ADV` part-of-speech tag but they are differentiated by additional features.

Note that in Germanic languages, some adverbs may also function as verbal particles, as in *write **down*** or *end **up***. They are still tagged `ADV` and not `PART`.

Note that there are words that may be traditionally called numerals in some languages (e.g. Czech) but they are treated as adverbs in our universal tagging scheme. In particular, *adverbial ordinal numerals* ([cs] *poprvé* "for the first time") and *multiplicative numerals* (e.g. *once, twice*) behave syntactically as adverbs and are tagged `ADV`.

Note that there are verb forms such as *transgressives* or *adverbial participles* that share properties and usage of adverbs and verbs. Depending on language and context, they may be classified as either `VERB` or `ADV`.

Examples

- *very*
- *well*
- *exactly*
- *tomorrow*
- *up, down*
- interrogative adverbs: *where, when, how, why*
- demonstrative adverbs: *here, there, now, then*
- indefinite adverbs: *somewhere, sometime, anywhere, anytime*
- totality adverbs: *everywhere, always*
- negative adverbs: *nowhere, never*

References

- Loos, Eugene E., et al. 2003. Glossary of linguistic terms: What is an adverb?
- Wikipedia

edit ADV

AUX: auxiliary

Definition

An auxiliary is a function word that accompanies the lexical verb of a verb phrase and expresses grammatical distinctions not carried by the lexical verb, such as person, number, tense, mood, aspect, voice or evidentiality. It is often a verb (which may have non-auxiliary uses as well) but many languages have nonverbal TAME markers and these should also be tagged `AUX`. The class `AUX` also include copulas (in the narrow sense of pure linking words for nonverbal predication).

Modal verbs may count as auxiliaries in some languages (English). In other languages their behavior is not too different from the main verbs and they are thus tagged `VERB`.

Note that not all languages have grammaticalized auxiliaries, and even where they exist the dividing line between full verbs and auxiliaries can be expected to vary between languages. Exactly which words are counted as `AUX` should be part of the language-specific documentation.

Examples

- Tense auxiliaries: ***has*** *(done),* ***is*** *(doing),* ***will*** *(do)*
- Passive auxiliaries: ***was*** *(done),* ***got*** *(done)*
- Modal auxiliaries: ***should*** *(do),* ***must*** *(do)*
- Verbal copulas: *He* ***is*** *a teacher.*

References

- Loos, Eugene E., et al. 2003. Glossary of linguistic terms: What is an auxiliary verb?
- Wikipedia

edit AUX

CCONJ: coordinating conjunction

Definition

A coordinating conjunction is a word that links words or larger constituents without syntactically subordinating one to the other and expresses a semantic relationship between them.

For *subordinating conjunctions,* see `SCONJ`.

Examples

- *and*
- *or*
- *but*

References

- Loos, Eugene E., et al. 2003. Glossary of linguistic terms: What is a coordinating conjunction?
- Wikipedia

edit CCONJ

DET: determiner

Definition

Determiners are words that modify nouns or noun phrases and express the reference of the noun phrase in context. That is, a determiner may indicate whether the noun is referring to a definite or indefinite element of a class, to a closer or more distant element, to an element belonging to a specified person or thing, to a particular number or quantity, etc.

Determiners under this definition include both *articles* and *pro-adjectives (pronominal adjectives)*, which is a slightly broader sense than what is usually regarded as determiners in English. In particular, there is no general requirement that a nominal can be modified by at most one determiner, although some languages may show a strong tendency towards such a constraint. (For example, an English nominal usually allows only one `DET` modifier, but there are occasional cases of *addeterminers*, which appear outside the usual determiner, such as [en] *all* in ***all*** *the children survived*. In such cases, both *all* and *the* are given the POS `DET`.)

Note that the `DET` tag includes (pronominal) *quantifiers* (words like *many, few, several),* which are included among determiners in some languages but may belong to numerals in others. However, *cardinal numerals* in the narrow sense *(one, five, hundred)* are not tagged `DET` even though some authors would include them in quantifiers. Cardinal numbers have their own tag `NUM`.

Also note that the notion of determiners is unknown in traditional grammar of some languages (e.g. Czech); words equivalent to English determiners may be traditionally classified as pronouns and/or numerals in these languages. In order to annotate the same thing the same way across languages, the words satisfying our definition of determiners should be tagged `DET` in these languages as well.

It is not always crystal clear where pronouns end and determiners start. Unlike in UD v1 it is no longer required that they are told apart solely on the base of the context. The words can be pre-classified in the dictionary as either `PRON` or `DET`, based on their *typical* syntactic distribution (and morphology, when applicable). Language-specific documentation should list all determiners (it is a closed class) and point out ambiguities, if any.

See also general principles on pronominal words for more tips on how to define determiners. In particular:

- Articles *(the, a, an)* are always tagged `DET`; their PronType is `Art`.
- Pronominal numerals (quantifiers) are tagged `DET`; besides `PronType`, they also use the NumType feature.
- Words that behave similar to adjectives are `DET`. Similar behavior means:
  - They are more likely to be used attributively (modifying a noun phrase) than substantively (replacing a noun phrase). They may occur alone, though. If they do, it is either because of ellipsis, or because the hypothetical modified noun is something unspecified and general, as in *All [visitors] must pay.*
  - Their inflection (if applicable) is similar to that of adjectives, and distinct from nouns. They agree with the nouns they modify. Especially the ability to inflect for gender is typical for adjectives and determiners. (Gender of nouns is determined lexically and determiners may be required by the grammar to agree with their nouns in gender; therefore they need to inflect for gender.)
- Possessives vary across languages. In some languages the above tests put them in the `DET` category. In others, they are more like a normal personal pronoun in a specific case (often the genitive), or a personal pronoun with an adposition; they are tagged `PRON`.

Examples

- articles (a closed class indicating definiteness, specificity or givenness): *a, an, the*
- possessive determiners (which modify a nominal): [cs] *můj, tvůj, jeho, její, náš, váš, jejich*; [en] *my, your*
- demonstrative determiners: *this* as in *I saw* ***this*** *car yesterday.*
- interrogative determiners: *which* as in *"****Which*** *car do you like?"*
- relative determiners: *which* as in *"I wonder* ***which*** *car you like."*

- quantity determiners (quantifiers): indefinite *any*, universal: *all*, and negative *no* as in *"We have no cars available."*

References

- Loos, Eugene E., et al. 2003. Glossary of linguistic terms: What is a determiner?
- Wikipedia

edit DET

`INTJ`: interjection

Definition

An interjection is a word that is used most often as an exclamation or part of an exclamation. It typically expresses an emotional reaction, is not syntactically related to other accompanying expressions, and may include a combination of sounds not otherwise found in the language.

Note that words primarily belonging to another part of speech retains their original category when used in exclamations. For example, *God* is a `NOUN` even in exclamatory uses.

As a special case of interjections, we recognize feedback particles such as *yes*, *no*, *uhuh*, etc.

Examples

- *psst*
- *ouch*
- *bravo*
- *hello*

References

- Loos, Eugene E., et al. 2003. Glossary of linguistic terms: What is an interjection?
- Wikipedia

edit INTJ

`NOUN`: noun

Definition

Nouns are a part of speech typically denoting a person, place, thing, animal or idea.

The `NOUN` tag is intended for common nouns only. See `PROPN` for proper nouns and `PRON` for pronouns.

Note that some verb forms such as *gerunds* and *infinitives* may share properties and usage of nouns and verbs. Depending on language and context, they may be classified as either `VERB` or `NOUN`.

Examples

- *girl*
- *cat*
- *tree*
- *air*
- *beauty*

References

- Loos, Eugene E., et al. 2003. Glossary of linguistic terms: What is a noun?
- Wikipedia

edit NOUN

NUM: numeral

Definition

A numeral is a word, functioning most typically as a determiner, adjective or pronoun, that expresses a number and a relation to the number, such as quantity, sequence, frequency or fraction.

Note that cardinal numerals are covered by NUM whether they are used as determiners or not (as in *Windows Seven*) and whether they are expressed as words *(four)*, digits *(4)* or Roman numerals *(IV)*. Other words functioning as determiners (including quantifiers such as *many* and *few*) are tagged DET.

Note that there are words that may be traditionally called numerals in some languages (e.g. Czech) but which are not tagged NUM. Such non-cardinal numerals belong to other parts of speech in our universal tagging scheme, based mainly on syntactic criteria: ordinal numerals are adjectives *(first, second, third)* or adverbs ([cs] *poprvé* "for the first time"), multiplicative numerals are adverbs *(once, twice)* etc.

Examples

- *0, 1, 2, 3, 4, 5, 2014, 1000000, 3.14159265359*
- *one, two, three, seventy-seven*
- *I, II, III, IV, V, MMXIV*

References

- Loos, Eugene E., et al. 2003. Glossary of linguistic terms: What is a numeral?
- Wikipedia: Numeral

edit NUM

PART: particle

Definition

Particles are function words that must be associated with another word or phrase to impart meaning and that do not satisfy definitions of other universal parts of speech (e.g. adpositions, coordinating conjunctions, subordinating conjunctions or auxiliary verbs). Particles may encode grammatical categories such as negation, mood, tense etc. Particles are normally not inflected, although exceptions may occur.

Note that the PART tag does not cover so-called *verbal particles* in Germanic languages, as in *give in* or *end up*. These are adpositions or adverbs by origin and are tagged accordingly ADP or ADV. Separable verb prefixes in German are treated analogically.

Note that not all function words that are traditionally called particles in Japanese automatically qualify for the PART tag. Some of them do, e.g. the question particle か / *ka.* Others (e.g. に / *ni*, の / *no*) are parallel to adpositions in other languages and should thus be tagged ADP.

In general, the PART tag should be used restrictively and only when no other tag is possible. The the language-specific documentation should list the words classified as PART in the given language.

Examples

- Possessive marker: [en] *'s*
- Negation particle: [en] *not;* [de] *nicht*
- Question particle: [ja] か / *ka* (adding this particle to the end of a clause turns the clause into a question); [tr] *mu*
- Sentence modality: [cs] *ať, kéž, nechť* (*Let's* do it! *If only* I could do it over. *May you* have an enjoyable stay!)

References

- Loos, Eugene E., et al. 2003. Glossary of linguistic terms: What is a particle?
- Wikipedia

edit PART

PRON: pronoun

Definition

Pronouns are words that substitute for nouns or noun phrases, whose meaning is recoverable from the linguistic or extralinguistic context.

Pronouns under this definition function like nouns. Note that some languages traditionally extend the term *pronoun* to words that substitute for adjectives. Such words are not tagged `PRON` under our universal scheme. They are tagged as determiners in order to annotate the same thing the same way across languages.

It is not always crystal clear where pronouns end and determiners start. Unlike in UD v1 it is no longer required that they are told apart solely on the base of the context. The words can be pre-classified in the dictionary as either `PRON` or `DET`, based on their *typical* syntactic distribution (and morphology, when applicable). Language-specific documentation should list all pronouns (it is a closed class) and point out ambiguities, if any.

See also general principles on pronominal words for more tips on how to define pronouns. In particular:

- Non-possessive personal, reflexive or reciprocal pronouns are always tagged `PRON`.
- Possessives vary across languages. In some languages the above tests put them in the `DET` category. In others, they are more like a normal personal pronoun in a specific case (often the genitive), or a personal pronoun with an adposition; they are tagged `PRON`.

Examples

- personal pronouns: *I, you, he, she, it, we, they*
- reflexive pronouns: *myself, yourself, himself, herself, itself, ourselves, yourselves, theirselves*
- interrogative pronouns: *who, what* as in ***What*** *do you think?*
- relative pronouns: *who, what* as in *I wonder* ***what*** *you think.* (Note, however, that some relative clause introducing words, such as [en] *that* are better analyzed as subordinating conjunctions (otherwise known as "complementizers" in the literature), and so are tagged as `SCONJ`.)
- indefinite pronouns: *somebody, something, anybody, anything*
- total pronouns: *everybody, everything*
- negative pronouns: *nobody, nothing*
- possessive pronouns (which usually stand alone as a nominal): *mine, yours, (his), hers, (its), ours, theirs*

References

- Loos, Eugene E., et al. 2003. Glossary of linguistic terms: What is a pronoun?
- Wikipedia

edit PRON

PROPN: proper noun

Definition

A proper noun is a noun (or nominal content word) that is the name (or part of the name) of a specific individual, place, or object.

Note that `PROPN` is only used for the subclass of nouns that are used as names and that often exhibit special syntactic properties (such as occurring without an article in the singular in English). When other phrases or sentences are used as names, the component words retain their original tags. For example, in *Cat on a Hot Tin Roof*, *Cat* is `NOUN`, *on* is `ADP`, *a* is `DET`, etc.

A fine point is that it is not uncommon to regard words that are etymologically adjectives or participles as proper nouns when they appear as part of a multiword name that overall functions like a proper noun, for example in *the Yellow Pages*, *United*

Airlines or *Thrall Manufacturing Company*. This is certainly the practice for the English Penn Treebank tag set.

Acronyms of proper nouns, such as *UN* and *NATO*, should be tagged `PROPN`. Even if they contain numbers (as in various product names), they are tagged `PROPN` and not `SYM`: *130XE*, *DC10*, *DC-10*. However, if the token consists entirely of digits (like *7* in *Windows 7*), it is tagged `NUM`.

Examples

- *Mary*, *John*
- *London*
- *NATO*, *HBO*

References

- Loos, Eugene E., et al. 2003. Glossary of linguistic terms: What is a proper noun?
- Wikipedia

edit PROPN

PUNCT: punctuation

Definition

Punctuation marks are non-alphabetical characters and character groups used in many languages to delimit linguistic units in printed text.

Punctuation is not taken to include logograms such as *$*, *%*, and *§*, which are instead tagged as `SYM`.

Examples

- Period: *.*
- Comma: *,*
- Parentheses: *()*

References

- Wikipedia

edit PUNCT

SCONJ: subordinating conjunction

Definition

A subordinating conjunction is a conjunction that links constructions by making one of them a constituent of the other. The subordinating conjunction typically marks the incorporated constituent which has the status of a (subordinate) clause.

We follow Loos et al. 2003 in recognizing these three subclasses as subordinating conjunctions:

- Complementizers, like [en] *that* or *if*
- Adverbial clause introducers, like [en] *when*, *since*, or *before* (when introducing a clause not a nominal)
- Relativizers, like [he] *še*. (Note that these words, which simply introduce a relative caluse, and normally don't inflect, need to be distinguished from relative or resumptive pronouns, which have a nominal function within the relative clause and which we analyze as `PRON`.)

For *coordinating conjunctions,* see `CCONJ`.

Examples

- *that* as in *I believe **that** he will come.*

- *if*
- *while*

References

- Loos, Eugene E., et al. 2003. Glossary of linguistic terms: What is a subordinating conjunction?
- Wikipedia

edit SCONJ

SYM: symbol

Definition

A symbol is a word-like entity that differs from ordinary words by form, function, or both.

Many symbols are or contain special non-alphanumeric characters, similarly to punctuation. What makes them different from punctuation is that they can be substituted by normal words. This involves all currency symbols, e.g. *$ 75* is identical to *seventy-five dollars*.

Mathematical operators form another group of symbols.

Another group of symbols is emoticons and emoji.

Strings that consists entirely of alphanumeric characters are not symbols but they may be proper nouns: *130XE*, *DC10*; others may be tagged PROPN (rather than SYM) even if they contain special characters: *DC-10*. Similarly, abbreviations for single words are not symbols but are assigned the part of speech of the full form. For example, *Mr.* (mister), *kg* (kilogram), *km* (kilometer), *Dr* (Doctor) should be tagged nouns. Acronyms for proper names such as *UN* and *NATO* should be tagged as proper nouns.

Characters used as bullets in itemized lists *(•, ‣)* are not symbols, they are punctuation.

Examples

- *$, %, §, ©*
- *+, −, ×, ÷, =, <, >*
- :), ♥‿♥, 😁
- *john.doe@universal.org, http://universaldependencies.org/, 1-800-COMPANY*

edit SYM

VERB: verb

Definition

A verb is a member of the syntactic class of words that typically signal events and actions, can constitute a minimal predicate in a clause, and govern the number and types of other constituents which may occur in the clause. Verbs are often associated with grammatical categories like tense, mood, aspect and voice, which can either be expressed inflectionally or using auxilliary verbs or particles.

Note that the VERB tag covers main verbs *(content verbs)* but it does not cover *auxiliary verbs* and verbal *copulas* (in the narrow sense), for which there is the AUX tag. *Modal verbs* may be considered VERB or AUX, depending on their behavior in the given language. Language-specific documentation should specify which verbs are tagged AUX in which contexts.

Note that *participles* are word forms that may share properties and usage of adjectives and verbs. Depending on language and context, they may be classified as either VERB or ADJ.

Note that some verb forms such as *gerunds* and *infinitives* may share properties and usage of nouns and verbs. Depending on language and context, they may be classified as either VERB or NOUN.

Note that there are verb forms such as *converbs (transgressives)* or *adverbial participles* that share properties and usage of adverbs and verbs. Depending on language and context, they may be classified as either VERB or ADV.

Examples

- *run, eat*
- *runs, ate*
- *running, eating*

References

- Loos, Eugene E., et al. 2003. Glossary of linguistic terms: What is a verb?
- Wikipedia

edit VERB

X: other

Definition

The tag X is used for words that for some reason cannot be assigned a real part-of-speech category. It should be used very restrictively.

A special usage of X is for cases of code-switching where it is not possible (or meaningful) to analyze the intervening language grammatically (and where the dependency relation flat:foreign is typically used in the syntactic analysis). This usage does not extend to ordinary loan words which should be assigned a normal part-of-speech. For example, in *he put on a large sombrero*, *sombrero* is an ordinary NOUN.

Examples

- *And then he just **xfgh pdl jklw***

edit X