

ISST-TANL Tagsets

The morpho-syntactic and dependency tagsets were jointly developed by the Istituto di Linguistica Computazionale “Antonio Zampolli” (ILC-CNR) and the University of Pisa in the framework of the [TANL \(Text Analytics and Natural Language processing\)](#) project and was used for the annotation of the ISST-TANL dependency annotated corpus, which originates as a revision of the ISST-[CoNLL corpus](#) which was used in the CoNLL 2007 Shared Task on Dependency Parsing, and which was built in its turn starting from the [Italian Syntactic-Semantic Treebank \(ISST\)](#), a multi-layered corpus annotated at the orthographic, morpho-syntactic, syntactic and lexico-semantic levels.

ISST-TANL morpho-syntactic tagset

The ISST-TANL part-of-speech tags are based on the ILC/PAROLE tagset and are conformant to the EAGLES international standard. The table below documents the 14 coarse-grained pos tags (column 1) and the 37 fine-grained tags (column 2) used for ISST-TANL annotation.

Coarse-grained tag	Fine-grained tag	Description	Examples	Contexts of use
A	A	adjective	<i>bello, buono, pauroso, ottimo</i>	una bella passeggiata un ottimo attaccante una persona paurosa
	AP	possessive adjective	<i>mio, tuo, nostro, loro</i>	a mio parere il tuo libro
B	B	adverb	<i>bene, fortemente, malissimo, domani</i>	arrivo domani sto bene
	BN	negation adverb	<i>non</i>	non sto bene
C	CC	coordinative conjunction	<i>e, o, ma, ovvero</i>	i libri e i quaderni vengo ma non rimango
	CS	subordinative conjunction	<i>mentre, quando</i>	quando ho finito vengo mentre scrivevo ho finito l’inchiestro
D	DE	exclamative determiner	<i>che, quale, quanto</i>	che disastro! quale catastrofe!
	DI	indefinite determiner	<i>alcuno, certo, tale, parecchio, qualsiasi</i>	alcune telefonate parecchi giornali qualsiasi persona
	DQ	interrogative determiner	<i>che, quale, quanto</i>	che cosa quanta strada quale formazione
	DR	relative determiner	<i>cui, quale</i>	i cui libri
	DD	demonstrative determiner	<i>questo, codesto, quello</i>	questo denaro quella famiglia
E	E	preposition	<i>di, a, da, in, su, attraverso, verso, prima_di</i>	a casa del poeta prima_di giorno

Coarse-grained tag	Fine-grained tag	Description	Examples	Contexts of use
				verso sera
	EA	articulated preposition	<i>del, alla, dei, nelle</i>	nella casa il prezzo del pane
F	FB	“balanced” punctuation	() “ ” ‘ ’ - -	il gatto – che conoscete –
	FC	clause boundary punctuation	, ;	ha detto : Vieni!
	FF	comma, hyphen	,	mele, pere e banane due-trecento persone
	FS	sentence boundary punctuation	. ? !	mele, pere e banane. cosa vuoi?
I	I	interjection	<i>ahimè, beh, ecco, grazie</i>	Beh , che vuoi?
N	N	cardinal number	<i>uno, due, cento, mille, 28, 2000</i>	due partite 28 anni
	NO	ordinal number	<i>primo, secondo, centesimo</i>	secondo posto
P	PD	demonstrative pronoun	<i>questo, quello, costui</i>	quello di Roma costui uccide
	PE	personal pronoun	<i>egli, lui, esso noialtri, voialtri, essi io, me, tu, te</i>	io parto lo mangio
	PI	indefinite pronoun	<i>chiunque, ognuno, molto</i>	chiunque venga i diritti di ognuno
	PP	possessive pronoun	<i>mio, tuo, suo, loro, proprio</i>	il mio è qui più bella della loro
	PQ	interrogative pronoun	<i>che, chi, quanto</i>	non so chi parta quanto costa? che ha fatto ieri?
	PR	relative pronoun	<i>che, cui, quale</i>	ciò che dice il quale afferma a cui parlo
	PC	clitic pronoun	<i>ci, vi, mi, ti, la, le</i>	lo vidi li ho sentiti aver la le dissero, le videro mi dicono ci sposiamo vi credo si sente, si sentono ci vado spesso
R	RD	determinative article	<i>il, lo, la, i, gli, le</i>	il libro i gatti
	RI	indeterminative article	<i>uno, un, una</i>	un amico una bambina
S	S	common noun	<i>amico, insegnante, verità</i>	l' amico la verità
	SA	abbreviation	<i>ndr, a.C., d.o.c., km</i>	30 km

Coarse-grained tag	Fine-grained tag	Description	Examples	Contexts of use
				sesto secolo a.C.
	SP	proper noun	<i>Monica, Pisa, Fiat, Sardegna</i>	Monica scrive
T	T	predeterminer	<i>tutto, entrambi, ambedue</i>	tutte le notizie ambedue le idee
V	VA	auxiliary verb	<i>avere, essere, venire</i>	il peggio è passato ho scritto una lettera viene fatto domani
	VM	modal verb	<i>volere, potere, dovere, solere</i>	non posso venire vuole il libro
	V	main verb	<i>mangio, avere, passato, camminando</i>	il peggio è passato ho scritto una lettera vengo domani
X	X	residual class	it includes formulae, unclassified words, alphabetic symbols and the like	distanziare di 43'' mi piacce

Here follows the set of morph-syntactic features complementing the part of speech information: the first table documents the association between morpho-syntactic features and part of speech information (where features marked between square brackets are optionally specified); the second table reports the typology of features and their possible values.

POS	Associated features
A	gen, num, [sup]
AP	gen, num
B	[sup]
EA	[gen], [num]
DD, DE, DI, DQ, DR	gen, num
N, NO	[gen], [num]
P, PD, PI, PP, PQ, PR	gen, num
PC, PE	gen, num, per
RD, RI	gen, num
S, SA, SP	gen, num
V, VA, VM	num, per, [mod], tmp

Tag	Feature values	Meaning	Example wordform pos feats
gen	m f n	masculine feminine underspecified tag subsuming M and F values	caso S gen=M num=S giustizia S gen=F num=S ospite S gen=N num=S
num	s p n	singular plural underspecified tag subsuming S and P values	casa S gen=F num=S donne S gen=F num=P le PQ gen=F num=N per=3
mod	g f	gerundive infinitive	dedicando V mod=G essere V mod=F

Tag	Feature values	Meaning	Example wordform pos feats
	i c d m p	indicative subjunctive conditional imperative participle	trovava V num=S per=3 mod=I tmp=I possegga V num=S per=3 mod=C tmp=P sarebbero V num=P per=3 mod=D tmp=P cercate V num=P per=2 mod=M tmp=P rapita V gen=F num=S mod=P tmp=R
per	1 2 3	first person second person third person	possiamo V num=P per=1 mod=I tmp=P sapete V num=P per=2 mod=I tmp=P le PQ gen=F num=N per=3 vede V num=S per=3 mod=I tmp=P
ten	p f i r	present tense future tense imperfect tense past tense	ha V num=S per=3 mod=I tmp=P sarà V num=S per=3 mod=I tmp=F trovava V num=S per=3 mod=I tmp=I rapita V gen=F num=S mod=P tmp=R
sup	s	superlative	gravissimi A gen=M num=P sup=S benissimo B sup=S