

# CT Specifications Guide

## Positional and Structural Attributes

POSIT.	<b>word</b>	the token	<i>mangia</i>	... ..
	<b>lemma</b>	the lemma to which the token has been brought back	<i>mangiare</i>	cf. separate list
	<b>pos</b>	the Part of Speech with its Hierarchy-Defining Features (HDF)	<i>v.m.f.ind.pr</i>	cf. FD 1 hereunder
	<b>kat</b>	<ul style="list-style-type: none"> <li>the Morphosyntactic Features (MSF) codes</li> <li>the Hierarchy-Collapsed Features (HCF) codes</li> </ul>	<i>3,0,6,0,0</i> <i>111</i>	cf. FD 2 hereunder cf. FD 1 hereunder
	<b>typ</b>	the structure of a (portion of a) text: {prose; verse; rubrica}. Crossreferenced with type	<i>/P</i>	<i>/P /V /R</i>
	<b>genre</b>	the literary genre of a text: {documentary, didactic, historical, narrative, lyric}	<i>nar</i>	doc; did; stor; nar; lir
	<b>msform</b>	the unaltered token really appearing in manuscript	<i>magia</i>	... ..
.	<b>philform</b>	the philological emendation, with the usual diacritics (round & square brackets, italics)	<i>ma(n)gia; ma[n]gia;</i> <i>ma'n'gia</i>	cf. list hereunder
	STRUCT	<b>author</b>	the author of a text	<i>Anonimo</i>
.	<b>title</b>	title of a text	<i>Novellino</i>	cf. CT texts hereunder
	<b>chapter</b>	chapter number	<i>n</i>	... ..
	<b>par</b>	paragraph number	<i>n</i>	... ..
	<b>s</b>	sentence number	<i>n</i>	... ..
	<b>line</b>	line number (of the page)	<i>n</i>	... ..
	<b>page</b>	page number (of the printed edition)	<i>n</i>	... ..
	<b>type</b>	crossreference with typ	<i>/P</i>	<i>/P /V /R</i>

## Feature Declaration (FD) 1: POS & HDF

Tagset		<i>kat (HCF)</i>	<i>pos (HDF)</i>		
noun		20	<i>n.c</i>	noun.common	
		21	<i>n.p</i>	noun.proper	
	adjective		26	<i>adj</i>	adjective
		pro-det	30	<i>pd.dem.s</i>	pro-det.demonstrative.strong
	31		<i>pd.dem.w</i>	pro-det.demonstrative.weak	
	32		<i>pd.idf</i>	pro-det.indefinit.	
	33		<i>pd.pos.s</i>	pro-det.possessive.strong	
	34		<i>pd.pos.w</i>	pro-det.possessive.weak	
	35		<i>pd.int</i>	pro-det.interrogative	
	36		<i>pd.rel</i>	pro-det.relative	
37	<i>pd.per.s.no</i>		pro-det.personal.strong.nominative		
38	<i>pd.per.s.ob</i>	pro-det.personal.strong.oblique			
39	<i>pd.per.w.ob</i>	pro-det.personal.weak.oblique			
adverb	40	<i>pd.exc</i>	pro-det.exclamative		
	45	<i>adv.gn</i>	adverb.general		
conjunction	46	<i>adv.pc</i>	adverb.particle		
	50	<i>conj.co</i>	conjunction.coordinate		
adposition	51	<i>conj.sb</i>	conjunction.subordinate		
	56	<i>adp.pre</i>	adposition.preposition		
article	57	<i>adp.post</i>	adposition.postposition		
	60	<i>art.d</i>	article.determinative		
	61	<i>art.i</i>	article.indeterminative		

<b>numeral</b>	64	num.car	numeral.cardinal
	65	num.ord	numeral.ordinal
<b>interjection</b>	68	intj	interjection
<b>punctuation</b>	70	punct.fi	punctuation.final
	71	punct.nfi	punctuation.non-final
<b>residuals</b>	75	r.frg	residual.foreign
	76	r.abb	residual.abbreviation
	77	r.for	residual.formulae
	78	r.epe	residual.epenthesis
<b>verb (main)</b>	111	v.m.f.ind.pr	verb.main.finite.indicative.present
	112	v.m.f.ind.ipf	verb.main.finite.indicative.imperfect
	113	v.m.f.ind.pt	verb.main.finite.indicative.past
	114	v.m.f.ind.ft	verb.main.finite.indicative.future
	115	v.m.f.sub.pr	verb.main.finite.subjunctive.present
	116	v.m.f.sub.ipf	verb.main.finite.subjunctive.imperfect
	117	v.m.f.cnd.pr	verb.main.finite.conditional.present
	118	v.m.f.imp.pr	verb.main.finite.imperative.present
	121	v.m.nf.inf.pr	verb.main.non-finite.infinitive.present
	122	v.m.nf.par.pr	verb.main.non-finite.participle.present
	123	v.m.nf.par.pt	verb.main.non-finite.participle.past
	124	v.m.nf.ger.pr	verb.main.non-finite.gerunde.present
<b>verb (auxiliar)</b>	211	v.a.f.ind.pr	verb.auxiliar.finite.indicative.present
	212	v.a.f.ind.ipf	verb.auxiliar.finite.indicative.imperfect
	213	v.a.f.ind.pt	verb.auxiliar.finite.indicative.past
	214	v.a.f.ind.ft	verb.auxiliar.finite.indicative.future
	215	v.a.f.sub.pr	verb.auxiliar.finite.subjunctive.present
	216	v.a.f.sub.ipf	verb.auxiliar.finite.subjunctive.imperfect
	217	v.a.f.cnd.pr	verb.auxiliar.finite.conditional.present
	218	v.a.f.imp.pr	verb.auxiliar.finite.imperative.present
	221	v.a.nf.inf.pr	verb.auxiliar.non-finite.infinitive.present
	222	v.a.nf.par.pr	verb.auxiliar.non-finite.participle.present
	223	v.a.nf.par.pt	verb.auxiliar.non-finite.participle.past
	224	v.a.nf.ger.pr	verb.auxiliar.non-finite.gerunde.present
<b>verb (modal)</b>	311	v.md.f.ind.pr	verb.modal.finite.indicative.present
	312	v.md.f.ind.ipf	verb.modal.finite.indicative.imperfect
	313	v.md.f.ind.pt	verb.modal.finite.indicative.past
	314	v.md.f.ind.ft	verb.modal.finite.indicative.future
	315	v.md.f.sub.pr	verb.modal.finite.subjunctive.present
	316	v.md.f.sub.ipf	verb.modal.finite.subjunctive.imperfect
	317	v.md.f.cnd.pr	verb.modal.finite.conditional.present
	318	v.md.f.imp.pr	verb.modal.finite.imperative.present
	321	v.md.nf.inf.pr	verb.modal.non-finite.infinitive.present
	322	v.md.nf.par.pr	verb.modal.non-finite.participle.present
	323	v.md.nf.par.pt	verb.modal.non-finite.participle.past
	324	v.md.nf.ind.pr	verb.modal.non-finite.gerunde.present

## Feature Declaration (FD) 2: MSF

		<i>kat (MSF)</i>		
MSF	<b>person</b>	1	pers=1	position 1
		2	pers=2	
		3	pers=3	
	<b>gender</b>	4	gend=masc	position 2
		5	gend=fem	
		4;5	gend=c	
	<b>number</b>	6	numb=sg	position 3
		7	numb=pl	
		6;7	numb=n	
	<b>degree</b>	8	degr=pos	position 4
		9	degr=comp	
		10	degr=sup	
	<b>multiword</b>	11-19	loc=11-19	position 5

## CT texts

	<i>author</i>	<i>title</i>	<i>genre</i>	<i>type</i>
CT Texts	<b>MaestroRinuccino</b>	Sonetti	lir	/V
	<b>BonoGiamboni</b>	LibroViziVirtù	did	/P
	<b>BonoGiamboni</b>	TrattatoViziVirtù	did	/P
	<b>BrunettoLatini</b>	Favolello	did	/V
	<b>BrunettoLatini</b>	Tesoretto	did	/V
	<b>BrunettoLatini</b>	Rettorica	did	/P
	<b>Anonimi</b>	CapitoliCompagniaSanGilio(Statuti84)	doc	/P
	<b>DanteAlighieri</b>	VitaNuova	lir	/P /V
	<b>Anonimi</b>	CapitoliCompagniaMadonnaOrsanmichele(Statuti94/97)	doc	/P
	<b>ConsiglioDe'Cerchi</b>	Lettera	doc	/P
	<b>Consiglio&amp;LapoDe'Cerchi</b>	Lettera	doc	/P
	<b>CastraGualfredi&amp;c</b>	LibroDareEdAvere	doc	/P
	<b>LapoRiccomanni</b>	LibroDareEdAvere	doc	/P
	<b>Anonimo</b>	FioreDiFilosafi	nar	/P
	<b>Anonimi</b>	LibroOrdinamentiCompagniaSMariaCarmine(Statuti80)	doc	/P
	<b>Anonimo</b>	CronicaFiorentina	stor	/P
	<b>Anonimo</b>	VolgarizzamentoDisciplinaClericalis	nar	/P
	<b>Anonimo</b>	Novellino	nar	/P /V
	<b>GuidoCavalcanti(?)</b>	DueBallate(I'Vidi/SolPerPietà)	lir	/P
	<b>GuidoCavalcanti</b>	Rime	lir	/V
	<b>JacopoCavalcanti</b>	TreSonetti	lir	/V

## Special conventions

<b>Symbols</b>	¬	logicalnot	compounds	Mei¬di¬donna
	̣	brokenbar	n\d-phonosyntax	nonn¬ è, foglia¬d è
	·	periodcenter	phyllological italics	oṛo!
	÷	divide	proclisis (with assimilation)	de· regno, Be· ll' ho
	~	tilde	graphoclisiss	porta ÷l ÷te ÷ne
	^	caret	compendia	ka~ agosto
			ellipsis	^^^ (corresponding to usual "...")
			lacuna	molto [^^^] francamente
	×	multiply	deperdita	××osta (unreadable characters)
	*	asterisk	vacua	die *** d' aprile (blank in ms.)
	Ø	Oslash	zero morphemes	a ÷Ø demonî 'ai demonii'
	©...®	copyright&register	typographic italics	le credenze del © Credo in Deo ®
	(...)	round brackets	solved abbreviation	mante(n)gna
	[...]	square brackets	integration	rag[g]io
	{...}	brace brackets	graphical symbols	{SN} 'signum notarii'

## Nota Bene

- + **Graphoclitics** are treated as individual tokens. They are marked by the divide (ASCII 246 = ANSI 247) “÷”
- + **Lemmas** are introduced by the formula “lemma=”
- + **Cassationes** and **expunctiones** were not included in the CT texts.
- + The **template of annotation** in CT is “*token\_lemma=lemma,HDF/HCF,MSF1,MSF2,MSF3,MSF4,MSF5*”

Manuel Barbera, 29 August 2000.