The Tycho Brahe Corpus Annotation System

Morphological Tags (POS and Inflectional)

- VERBS And Their Iinflectional Tags - 🆕 PERSON AND NUMBER INFORMATION ADDED TO VERBS AS SOON AS POSSIBLE
- AGREEMENT - Inflectional Tags for Gender and Number
- NOUNS
- PRONOUNS
- DETERMINERS
- ADJECTIVES, ADVERBS, and QUANTIFIERS
- CONJUNCTIONS
- INTERROGATIVE/RELATIVE ELEMENTS
- PREPOSITIONS
- CLUSTERS - 🆕 (previously treated as 'WORDS AS UNIT')
- OUTRO - 🆕
- FOCUS PARTICLE
- CARDINAL NUMBERS
- NEGATION - 🆕 NEG TAG ADDED TO ALL NEGATIVE WORDS
- INTERJECTIONS
- FOREIGN and UNKNOWN WORDS
- PUNCTUATION
- SOME HOMOGRAPHIC LEXICAL ITEMS AND THEIR TAGS: Review
- ITEMS WITH FIXED TAGS. Review 🆕
- COMPLETE LIST OF TAGS 🆕

back to the main page of this manual

VERBS

Two different groups of verbs were contemplated. On one hand, under the tag VB, are all verbs that, besides presenting verbal agreement and finite features, assign thematic roles to their arguments, being considered full verbs. On the other hand, separate tags were attributed to ser *(be-individual level)* (SR), estar *(be-stage level)* (ET), ter *(have)* (TR), and haver *(there to be)* (HV), whose features seem to oscillate diachronically so that sometimes they may be employed as full verbs, and sometimes function as mere auxiliary verbs which are restricted to carrying inflectional information. Associated to these main POS tags, inflectional tags, indicating visible verbal morphology, may be added to the primary tags.

Clitic forms CL and SE can also be added to all verbal forms, except participles (-AN or –PP) - see verb-clitic combinations; for future forms, see also mesoclisis.

🆕 As proposed below for pronouns, also verbs will be present person and number information as soon as possible.

TYPE	TAG	VALUE	EXAMPLES
SER (*to be - individual level)*)	SR	Infinitive (in verbal and nominal contexts)	a ser/SR exemplo de emendas
	SR-F	Inflected Infinitive (with explicit morphological marks)	para sermos/SR-F completos e felices
	SR-I	Imperative (with explicit morphological marks)	Sê/SR-I fiel a Deus, e ele sempre te amará.
	SR-P	Present	não é/SR-P muito que assim me acôlha...
	SR-SP	Present Subjunctive	que êles sejam/SR-SP tão teimosos... seja/SR-SP Deus louvado...
	SR-D	Past	Fatalidade foi/SR-D faltar-vos carta minha ... Boa maré era/SR-D esta ...
	SR-RA	Verbal forms with the inflectional morpheme -ra	Past Perfect Ele fôra/SR-RA embora,... Conditional Hypothetical Se fôra/SR-RA Deus, como vós sois,...
	SR-SD	Past Subjunctive	...que fosse/SR-SD contra a minha regra...
	SR-R	Future and Conditional	esta lembrança será/SR-R, mais que necessidade, imprudência ... seria/SR-R possível que ...

	SR-SR	Future Subjunctive	mais ditoso que Quando fordes/SR-SR rei
	SR-G	Gerund	sendo/SR-G tão má cousa como sou
	SR-PP	Perfect Participle	como haver sido/SR-PP tão má

TYPE	TAG	VALUE	EXAMPLES
HAVER (*there to be*)	HV	Infinitive (in verbal and nominal contexts)	em tudo isto pode haver/HV vaidade
	HV-F	Inflected Infinitive (with explicit morphological marks)	quão lastimosa cousa será haverem/HV-F de deixar aquelas almas
	HV-I	Imperative (with explicit morphological marks)	Havei/HV-I vós....
	HV-P	Present	em toda parte há/HV-P sido servido fazer-me
	HV-SP	Present Subjunctive	...fingir que tem reparado em que eu haja/HV-SP faltado assim.
	HV-D	Past	...o amor que vos tem, havia/HV-D de furtar-lho.
	HV-RA	Verbal forms with the inflectional morpheme -ra	Past Perfect Conditional Hypothetical (...), como se houvera/HV-RA trabalhado.
	HV-SD	Past Subjunctive	Não cuidei que houvesse/HV-SD tão pouco coração.
	HV-R	Future and Conditional	...o que haverá/HV-R em contrário. Haveria problemas, caso.....
	HV-SR	Future Subjuctive	Se me houver/HV-SR perdoado as culpas...
	HV-G	Gerund	Havendo/HV-G provimentos, só os envie se...
	HV-PP	Perfect Participle	...tem havido/HV-PP nesse convento.
	HV-NA	Passive Participle ([see agreement](#) and [participles as adjectives](#))	As casas havidas/HV-AN-F-P

TYPE	TAG	VALUE	EXAMPLES
ESTAR (*to be - stage level*)	ET	Infinitive (in verbal and nominal contexts)	é estar/ET nas mãos de meu Senhor Jesu Cristo,...
	ET-F	Inflected Infinitive (with explicit morphological marks)	Assim, para estarmos/ET-F com Nosso Senhor,...
	ET-I	Imperative (with explicit morphological marks)	Estai/ET-I agora junto ao amor de...
	ET-P	Present	que estão/ET-P no côro.
	ET-SP	Present Subjunctive	não esteja/ET-SP como eu cuidava.
	ET-D	Past	que já estavam/ET-D de espera...
	ET-RA	Verbal forms with the inflectional morpheme -ra	Past Perfect Bem aviado estivera/ET-RA eu... Conditional Hypothetical Se estivera/ET-RA

ET-SD	Past Subjunctive	Espero em Deus que nunca estivesse/ET-SD azedo...
ET-R	Future and Conditional	que cêdo estará/ET-R em Deus... que cêdo estaria/ET-R.....
ET-SR	Future Subjuctive	Quando estiveres/ET-SR melhor....
ET-G	Gerund	Estando/ET-G inocente...
ET-PP	Perfect Participle	Tinha estado/ET-PP em Mafra..

TYPE	TAG	VALUE	EXAMPLES
TER (*to have*)	TR	Infinitive (in verbal and nominal contexts)	a ter/TR conta o que se passara..
	TR-F	Inflected Infinitive (with explicit morphological marks)	de terem/TR-F as noviças...
	TR-I	Imperative (with explicit morphological marks)	Tende/TR-I esperança...
	TR-P	Present	Tenho/TR-P a Madre de Deus por mim...
	TR-SP	Present Subjunctive	Não tenhais/TR-SP cuidado...
	TR-D	Past	Que teve/TR-D seu próprio filho...
	TR-RA	Verbal forms with the inflectional morpheme -ra	Past Perfect Tivera/TR-RA bondade e simplicidade... Conditional Hypothetical Se tivera/TR-RA eu....
	TR-SD	Past Subjunctive	Se tivesse/TR-SD desejo...
	TR-R	Future	Como o terá/TR-R para fazer um sermão
	TR-SR	Future Subjunctive	Se Deus não me tiver/TR-SR muito...
	TR-G	Gerund	Tendo/TR-G sido encontrado...
	TR-PP	Perfect Participle	Tido/TR-PP paciência e afeto...
	TR-NA	Passive Participle ([see agreement](#) and [participles as adjectives](#))	Tidas/TR-AN-F-P as cartas...

TYPE	TAG	VALUE	EXAMPLES		
	VB	Infinitive (in verbal and nominal contexts)	...e bastará ao menos dar/VB três vezes Ordens no ano, ...é a que retrata com mais semelhança o falar/VB e conversação dentre os amigos O amar/VB a Deus (see [infinitives as nouns](#))
	VB-F	Inflected Infinitive (with explicit morphological marks)	Segue-se agora tratarmos/VB-F de como uma pessoa se alongará...
	VB-I	Imperative (with explicit morphological marks)	Sabei/VB-I que agora estais no Cabo da Bôa Esperança...
	VB-P	Present	Bem me parece/VB-P isto, porque... Creio/VB-P em Jesu Cristo,...
	VB-SP	Present Subjunctive	..., até que nos vejamos/VB-SP
	VB-D	Past	Só vos digo que estimei/VB-D as vossas novas....
			Past Perfect

VERBS in general	VB-RA	Verbal forms with the inflectional morpheme -ra	Bem quisera/VB-RA eu prestar para vo-lo agradecer,... Conditional E se mais mundo houvera, lá chegara/VB-RA Hypothetical Se ninguém no mundo se salvara/VB-RA, então (...)
	VB-SD	Past Subjunctive	Se quisesse/VB-SD meu Deus que.....
	VB-R	Future	...de quem não verei/VB-R por muitos anos
	VB-SR	Future Subjunctive	Quando eu puder/VB-SR lhe darei algum....
	VB-G	Gerund	Sabendo/VB-G que...
	VB-PP	Perfect Participle	Terá feito/VB-PP os trabalhos...
	VB-AN	Passive Participle ([see agreement](#) and [adjective as participles](#))	Foi servido/VB-AN de reduzir ao seu rebanho esta ovelha perdida .... Aparecera Cristo crucificado/VB-AN

back to the top of the page
back to the main page of this manual

AGREEMENT - Gender and Number in nominal

The following inflectional tags for [gender](#) and [number](#) are contemplated in our annotation system in order to capture the morphological richness Portuguese exhibits in determiners, pronouns, nouns, adjectives, quantifiers, passive participles, and so on.

TYPE	TAG	VALUE	APPLICATION	EXAMPLES
GENDER	none	Masculine	Default	que o/D dia/N
	-F	Feminine	Determiners	que a/D-F noite
			Adjectives	a noite fermosa/ADJ-F
			Possessive Pronouns	de sua/PRO$-F fermosura
			Quantifiers	de toda/Q-F a alma
			Passive Participles	a ser feita/VB-AN-F
			Relative Pronoun 'cuja'	de cuja/WPRO$-F virtude
			Cardinal Number 'duas'	de duas/NUM-F fragatas
	-G	Double gender	Determiner 'tal'	Tal/D-G saída não nos perturbara.
			Adjectives	da grande/ADJ-G revolta de...
			Quantifiers	de cada/Q-G homem...
NUMBER	None	Singular	Default	que o/D dia/N
	-P	Plural	Nouns	as terras/N-P...
			Determiners	Os/D-P mares e as/D-F-P terras....
			Adjectives	Dois fermosos/ADJ-P frutos... Duas fermosas/ADJ-F-P frutas..
			Possessive Pronouns	Seus/PRO$-P favores e....
			Quantifiers	Muitos/Q-P louvores...
			Passive Participles	Foram encontrados/VB-AN-P mulheres e crianças....
			Some Relative Pronouns	As correntes cujos/WPRO$-P elos....

back to the top of the page
back to the main page of this manual

NOUNS

Common and proper nouns have been attributed the POS tags N and NPR respectively, since they behave differently from a syntactic point of view. As for the inflectional features related to nouns, Portuguese exhibits visible morphology concerning gender and number. Nevertheless, common and proper nouns have only been attributed number inflectional tags. Since the syntactic context is irrelevant for gender attribution to nouns, we decided not to contemplate this part of speech class with a gender inflectional tag, differently from what happens to determiners and adjectives. This idiosyncratic noun property is clearly shown in a list to be established of all the nouns contained in the *corpus*. This list will compose a dictionary and serve as reference for the automated morphological tagger.

TYPE	TAG	VALUE	APPLICATION	EXAMPLES
NOUNS	N	Singular	Common nouns	A gente/N de Lisboa Concordo com este parecer/N Este criado/N de Vossa Senhoria contraria ao bem/N commum
			Others	O contrário/N é verdadeiro Sobre este particular/N, penso que
			Collective nouns	A multidão/N
			Infinitive forms largelly recognised as nouns	O jantar/N fôra agradabilíssimo *versus* **Jantar/VB com amigos é algo como...**
	N-P	Plural	Common nouns	Mares/N-P e espumas/N-P Jornais/N-P, meu caro amigo...
PROPER NOUNS	NPR	Singular	Common nouns with initial capital letter (except after a period punctuation – see above)	... leva na maõ direita o Ceptro/NPR , com que ha de governar o Povo/NPR em Paz/NPR .
			All Address Forms, even those with non initial capital letter	Não diga Vossa/PRO$-F Senhoria/NPR..... Por isso, dê a senhora/NPR graças a Nosso Senhor, pois...
			Titles and Epithets	Ilustríssimo/ADJ-S Senhor/NPR Gonçalves/NPR de/P Azevedo/NPR Reverendíssima/ADJ-S-F Madre/NPR Padre/NPR Provincial/ADJ-G *and* O Provincial/ADJ-G informa que... Espírito/NPR Santo/ADJ Frei/NPR António/NPR Dom/NPR Frederico/NPR Ao ver El-Rei/NPR ao longe.... Por Monsieur/NPR Estrada/NPR o Rey/NPR Catholico/NPR
			Dates, days of week, months, and holidays	Sábado/NPR e domingo/NPR Dezembro/NPR e março/NPR Dia/NPR de São/NPR João/NPR O dia/N de Natal/NPR Sexta/NPR Feira/NPR da Paixão/NPR A Paixão/NPR de Cristo/NPR
			Places	O/D Amazonas/NPR Portugal/NPR Rio/NPR de/P Janeiro/NPR

			a rua/N Nova/NPR
		Other Proper Nouns	São/NPR Bertolameu/NPR João/NPR Alves/NPR
NPR-P	Plural	Address Forms	Suas/PRO$-F-P Majestades/NPR-P e Altezas/NPR-P
		Titles	As Condessas/NPR-P de Braga
		Other Proper Nouns	Os/D-P Mendonças/NPR-P

PRONOUNS

Taking into consideration syntactic differences among strong and deficient pronouns commonly found in null subject languages, like European Portuguese and some other Romance languages, different primary tags were attributed to each of the lexical items of this binary division. Strong pronouns in subject or object positions are tagged PRO and clitic pronouns are tagged CL. It is also worth pointing out that in what concerns the clitics, mesoclisis, which is largely productive in the history of Portuguese, was attributed the diacritic signal *!*. This POS class deserved a third and final tag, PRO$, which was attributed to possessive pronouns.

TYPE	TAG	VALUE/APPLICATION	EXAMPLES
PRONOUNS	PRO	Personal pronouns	Eu/PRO, tu/PRO, ele/PRO, ela/PRO, você/PRO, nós/PRO, vós/PRO, eles/PRO, elas/PRO, vocês/PRO
		Oblique forms	Mim/PRO, ti/PRO, si/PRO
	P+PRO	Preposition +oblique forms	Comigo/P+PRO, Contigo/ P+PRO, Consigo/P+PRO, Conosco/P+PRO, Convosco/P+PRO
POSSESSIVE PRONOUNS	PRO$	Masculine singular	Meu/PRO$, teu/PRO$, seu/PRO$, nosso/PRO$, vosso/PRO$
	PRO$-F	Feminine singular	Minha/PRO$-F, tua/PRO$-F, sua/PRO$-F, nossa/PRO$-F, vossa/PRO$-F
	PRO$-P	Masculine plural	Meus/PRO$-P etc...
	PRO$-F-P	Feminine plural	Minhas/PRO$-F-P etc...
CLITICS	CL	Clitics in general *me, te, o(s), a(s), lhe(s), nos, vos*	Já o/CL havia dado a mi.... Quando lhe/CL visse.... Dar-te-ei/VB-R!CL Entregou-me/VB-D+CL antes de...
	SE NB.: In the parsed files, "se" is tagged as **CL**.	Clitic "se" in all its functions	porque todos para comsigo saõ como os olhos, que vendo tudo, naõ se/SE vem a si mesmos;
	CL+CL	Clitics contraction in general	e não ma/CL+CL deu

VERB-CLITIC COMBINATIONS

Clitics can combine with all verb forms except participles and passives. An exhaustive list of combination possibilities is presented in the NEW complete list of tags ; some examples are shown below:

TYPE	TAG	EXAMPLES
verb+one clitic	**...**+CL	haverem-nos/HV-F+CL, terem-na/TR-F+CL, mostrou-o/VB-D+CL, tergiversandolhe/VB-G+CL,
verb+two clitics	**...**+CL+CL	permittindo-lho/VB-G+CL+CL, quadruplicandolho/VB-G+CL+CL

mesoclisis

Analitic future forms *há de*, *hey de* etc. turn into the synthetic *–rá*, *rão*, etc. future forms (tagged with /VB-R). Clitics can appear inserted into the these forms, and are marked as follows:

TAG	EXAMPLES

SR-R!CL	ser-lhe-há/ SR-R!CL
ET-R!CL	estar-me-iam/ ET-R!CL
HV-R!CL	haver-te-ei/ HV-R!CL
TR-R!CL	ter-lhe-ia/ TR-R!CL
VB-R!CL	amar-te-ei/ VB-R!CL

DETERMINERS

Tag D was applied not only to the elements traditionally classified as *definite articles* but also to the *inflected demonstratives*, all of which also receive tags for number and gender. This is due to the fact that, as can be seen in the history of Portuguese, articles and inflected demonstratives share the same syntactic distribution. On the other hand, since uninflected demonstratives show actual pronominal behavior (not a determiner one), the tag DEM was reserved to them. Additionally, some other elements receive tag D, followed by the inflectional tag G, which is assigned to the words which have no distinction form for masculine and feminine. Concerning the case of ***um**(a)*, we opted to attribute it tag D, associated to the inflectional tag -UM, so that its automatic identification with the other members of the determiners class would be facilitated. The associated inflectional tag -UM has the advantage of differentiating this element from the other articles, which is desirable due to its property of being either [+referential] or [+quantified].

NEW The set ***outro, outros***, etc. is now treated in a new form – please refer to outro

TYPE	TAG	VALUE/ APPLICATION	EXAMPLES
DEFINITE DETERMINERS AND INFLECTED DEMONSTRATIVES	D	Masculine singular	O/D homem Este/D homem Aquele/D homem e o mandou para este/D.
	D-F	Feminine singular	A/D-F videira (see 'a' as preposition) Esta/D-F videira Aquela/D-F videira
	D-P	Masculine plural	Os/D-P homens Estes/D-P homens Aqueles/D-P homens
	D-F-P	Feminine plural	As/D-F-P videiras Estas/D-F-P videiras Aquelas/D-F-P videiras
	D-G	Double gender singular	Tal/D-G homem porque em tal/D-G parte vale muito Tal/D-G videira (sse 'tal' as other demonstratives, as comparative adjective, and comparative adverb)
	D-G-P	Double gender plural	Tais/D-G-P homens Tais/D-G-P videiras
INDEFINITE DETERMINERS/ CARDINAL NUMBER 'UM'	D-UM	Masculine singular	Um/D-UM homem Um/D-UM quilo de forragem
	D-UM-F	Feminine singular	Uma/D-UM-F flor Uma/D-UM-F dúzia de ovos...
	D-UM-P	Masculine plural	Uns/D-UM-P homens Uns/D-UM-P quilos de forragem Acabava por não saber como agradá-los: a uns/D-UM-P oferecia amor, a outros dedicação.....
	D-UM-F-P	Feminine plural	Umas/D-UM-F-P flores Umas/D-UM-F-P dúzias
OTHER DEMONSTRATIVES	DEM	Invariable demonstratives	Por isso/DEM a razão derrubou os ídolos... Sobre aquilo/DEM do convento, cêdo nos veremos e então falaremos.
		'tal'	Apenas por tal/DEM faria a guerra

| | | tal | (see 'tal' as comparative adjetives) |

ADJECTIVES, ADVERBS, AND QUANTIFIERS

The lexical items that, from an interpretive point of view, quantify over entities or events receive tag Q, which can be associated with gender or number inflectional tags. In Modern Portuguese, the quantificational property applied to entities can be neutralized depending on the position of the item that expresses it in the nominal phrase. What follows is that items that are generally classified as quantifiers in pre-nominal position, can be interpreted as adjectives when in post-nominal position. In this present annotated system, developed to be applied to Classic *and* Modern Portuguese, this distinction is not considered. Thus, quantified items are always tagged as Q. With respect to adjectives properly, besides gender and number inflectional tags that might apply to them, one can also identify the exclamative, comparative and superlative adjective forms, using tags R and S, respectively. Finally, as far as adverbs are concerned, the so-called *intensity adverbs* (**muito**, **pouco**, and so on) have been classified as event quantifiers in the presently proposed system, in parallel with the entity quantifiers. The tag ADV, therefore, has only been used to codify *time*, *place* and *manner adverbs* as well as the forms **mais** and **menos**. The variation tags R and S apply to adverbs in the same way as to adjectives.

NEW The set **outro, outros**, etc. is no longer tagged as ADJ – please refer to outro

ADJECTIVES

TYPE	TAG	VALUE	APPLICATION	EXAMPLES
ADJECTIVES	ADJ	Masculine singular	General adjectives	Por ser bonito/ADJ Um e meio/ADJ por cento Ano e meio/ADJ Meio/ADJ por cento
			'mesmo' as adjective	O mesmo/ADJ homem foi encontrado... O mesmo/ADJ se pode dizer... (see 'mesmo' as as focus particle, and as member of complex subordinating conjuction)
			Ordinal numbers (see cardinal numbers	O segundo/ADJ homem que partiu
	ADJ-F	Feminine singular		Uma fermosa/ADJ-F igreja ...
	ADJ-G	Double-gender singular	General adjectives	Um homen grande/ADJ-G e uma mulher amável/ADJ-G
			Others	Por conseguinte/ADJ-G, ... (see adverbs as prepositional phrases) NEW não/NEG obstante/ADJ-G (see clusters)
	ADJ-P	Masculine plural		Homens bonitos/ADJ-P
	ADJ-F-P	Feminine plural		Mulheres bonitas/ADJ-F-P
	ADJ	Double-	General adjetives	Homens e mulheres valentes/ADJ-G-P Os demais/ADJ-G-P

	ADJ-G-P	Double gender plural	'demais'	Os demais/ADJ-G-P homens saíram. As demais/ADJ-G-P foram entregues...
EXCLAMATIVE/COMPARATIVE ADJECTIVES	ADJ-R	Masculine singular ('tanto')		E aplicou tanto/ADJ-R esforço NEW que/C impediu o levantar da barreira, onde....
	ADJ-R-F	Feminine singular ('tanta')		A vontade foi tanta/ADJ-R-F NEW que/C acabou por .....
	ADJ-R-P	Masculine plural ('tantos')		...do coração que há tantos/ADJ-R-P tempos vive oculto nos mistérios do recato. E aplicou tantos/ADJ-R-P esforços NEW que/C impediu o levantar da barreira, onde....
	ADJ-R-F-P	Feminine plural ('tantas')		..., pois êle foi servido de reduzir ao seu rebanho esta ovelha perdida, quási no mesmo tempo que entre as trevas do pecado e entre os horrores de tantas/ADJ-R-F-P culpas se aumentavam os descaminhos e creciam os despenhadeiros.
	ADJ-R-G	'tal'		Muito estimara eu que fôsse tal/ADJ-R-G o meu desembaraço, NEW que/C pudesse escrever a Vossa Mercê mais a miúdo. Fez um tal/ADJ-R-G esforço NEW que/C não pode....
		'maior'/'menor'		Este é o maior/ADJ-R-G bem de todos... o grande amor que vos tem é maior/ADJ-R-G do/P+D que/WPRO cuidais.
		'melhor'/'pior'		desejando de empregar em melhor/ADJ-R-G guerra o que me ficava de vida. E isto acabou por ser melhor/ADJ-R-G do/P+D que/WPRO descansar,.... uma forma melhor/ADJ-R-G NEW que/C esta Este é o pior/ADJ-R-G .... O melhor/ADJ-R-G a fazer é isto.....
	ADJ-R-G-P	'tais'		Esses tais/ADJ-R-G-P fatos acabaram por trazer....
		'maiores'/menores'		...projetos maiores/ADJ-R-G-P NEW que/C aqueles....
		'melhores'/'piores'		...melhores/ADJ-R-G-P casas NEW que/C

				essas
SUPERLATIVE ADJECTIVES	ADJ-S	Masculine singular		Homem amabilíssimo/ADJ-S
	ADJ-S-F	Feminine singular		Mulher belíssima/ADJ-S-F
	ADJ-S-P	Masculine plural		Barcos rapidíssimos/ADJ-S-P
	ADJ-S-F-P	Feminine plural		Fragatas imponentíssimas/ADJ-S-F-P
ADJECTIVES AS NOUNS	N	General adjectives		Um pecador/N miserável *versus* Um homem bom ou mau, justo ou pecador/ADJ....
ADJECTIVES AS ADVERBS	ADV			Fragata rápida/ADJ-F ***versus*** Rápida/ADV e sorateiramente/ADV Homem baixo/ADJ ***versus*** Falando baixo/ADV, (....)
PARTICIPLES AS ADJECTIVES	VB-AN			Esta ovelha perdida/VB-AN-F Bemdito/VB-AN e louvado/VB-AN seja Deus Nosso Senhor

ADVERBS

TYPE	TAG	VALUE	APPLICATION	EXAMPLES
ADVERBS	ADV	Adverbs of manner		Completamente/ADV Rapidamente/ADV Infelizmente/ADV Deveras/ADV, Depressa/ADV
		Adverbs of place		Acima/ADV, abaixo/ADV, por/P baixo/ADV, em/P cima/ADV, debaixo/ADV, diante/ADV, adiante/ADV, cá/ADV, lá/ADV, aqui/ADV, ali/ADV, aquém/ADV, além/ADV, atrás/ADV, dentro/ADV, fora/ADV (see fora as [preposition](#)),
		Adverbs of time		Agora/ADV, ontem/ADV, hoje/ADV, amanhã/ADV, cedo/ADV, tarde/ADV, jamais/ADV, logo/ADV, antes/ADV, primeiro (=antes) Antes/ADV d'hontem/P+ADV fui informado.....(see 'antes' as ['comparative adverb'](#) and [more P+ADV](#)) Primeiro/ADV que/C os vossos achaques me viessem (...)
		Others		...que quase/ADV

		'acaso' as 'casualmente'	interrompemos nós....
EXCAMATIVE/ COMPARATIVE ADVERBS	ADV-R	'mais'/'menos'/ (in all syntactic contexts) 'demais'/'bastante'	por mais/ADV-R força **NEW** que/C aplicassem, não resolviam... por mais/ADV-R **NEW** que/C aplicassem força,.... O mais/ADV-R são detalhes que... que representam no que avultam mais/ADV-R do que/WPRO custam no que oprimem. Nada procuro mais/ADV-R **NEW** que/C o amor de Cristo. Por mais/ADV-R **NEW** que/C eu queira, (...) Pelo menos/ADV-R, isto foi feito. Ao menos/ADV-R isso (...) (...) assim nisto como no demais/ADV-R (...)
		'melhor'/'pior'	....e correu melhor/ADV-R do que eu
		'tão'/'tanto'/'tal'	E fico tão/ADV-R consolado com os favores do Céu, **NEW** que/C os maiores Impérios da terra regeitara eu. Só sinto ser-lhe tão/ADV-R ingrato e tão pobrezinho para pagar-lhe, **NEW** que/C nem um suspiro tenho de meu para oferecer-lhe; Quanto mais quer, tanto/ADV-R mais faz para obter Tanto/ADV-R ricos, quanto/WADV pobres... tal/ADV-R qual/WADV sou
		'antes'	Antes/ADV-R tarde do/P+D que/WPRO nunca
SUPERLATIVE ADVERBS	ADV-S		Lentíssimamente
ADVERBS AS PREPOSITIONAL OR ADVERBIAL PHRASES	.../P .../N .../P .../ADJ .../P.../ADV .../ADV.../P		Com/P certeza/N, , sem/D dúvida/N, por/P acaso/N Por/P conseguinte/ADJ, de/P certo/ADJ , Até/P amanhã/ADV, até/P logo/ADV, **NEW** no/P+D entanto/ADV (see clusters) **NEW** acerca/ADV de/P (see clusters)
ADVERBS AS EVENT QUANTIFIERS	Q	'muito'/'pouco'	Frei António trabalhou muito/Q durante ...
NEW NEGATIVE ADVERBS	ADV-NEG	'nunca'	see negation

back to the top of the page
back to the main page of this manual

QUANTIFIERS

TYPE	TAG	VALUE/APPLICATION	EXAMPLES
QUANTIFIERS	Q	Quantifiers in general	Tudo/Q será feito.
			Alguém/Q disse que...
			Outrém/Q
			Vossa Mercê sofrerá muito/Q.
			muito/Q trabalho a fazer...
			pouco/Q trabalho a fazer...
			Um pouco/Q de trabalho a fazer
			Algum/Q motivo apresenta...
	Q-F		por alguma/Q-F razão...
	Q-P		ambos/Q-P os freis...
			todos/Q-P os homens...
			alguns/Q-P trabalhos....
	Q-F-P		ambas/Q-F-P as religiosas...
			todas/Q-F-P as coisas...
			algumas/Q-F-P guerras....
	Q-G		Cada/Q-G um saberá...
			Qualquer/Q-G solução virá....
	Q-G-P		Quaiquer/Q-G-P resoluções...
NEW NEGATIVE QUANTIFIERS	Q-NEG	'nada', 'ninguém', 'nenhum'	se devem vigiar mais, que nenhumas/Q-NEG-F-P outras
	Q-NEG-P, Q-NEG-F, Q-NEG-F-P	'nenhuns'; 'nenhuma'; 'nenhumas'	see negation

CONJUNCTIONS

TYPE	TAG	VALUE/APPLICATION	EXAMPLES
COORDINATING CONJUNCTIONS	CONJ	Aditive *e*	Frei Luís e/CONJ Frei João
			Teria chegado e/CONJ saído sem ninguém perceber....
		Alternatives	Ou........., ou.......
			Ora........, ora.......
			Nem......., nem.....
	NEW CONJ-NEG	Aditive *nem*	Não fez o pedido, nem/CONJ-NEG sabia dizer quem o fizera....
			(see negation)
			Para dizer a Vossa Mercê que/C esta vida/N entra a ser exemplo de emendas...
			Espero que/C Deus me guie....
			Foi ontem que/C encontrei as...

Complementizer	C	'que'	Assim/ADV que/C ,/, do dinheiro que Vossa Excelência me remeteu não posso já fazer cousa alguma, por estarem as letras nas mãos de quem vendeu a fragata. (see 'que' as part of complex subordinating conjunctions)
SUBORDINATING CONJUNCTIONS NB.: In the parsed files, subordinating conjunctions are tagged as **C**.	CONJS	Causals, conditionals, finals, temporals, etc.	Quando/CONJS fores embora, ... Embora/CONJS não saibas,.... Se/CONJS soubesses.... Exceto/P se/CONJS... Caso/CONJS não saibas..... Conforme/CONJS afirmam,... Como/CONJS afirmam, ... Segundo/CONJS Porque/CONJS
COMPLEX SUBORDINATING CONJUNCTIONS	Various	Adverbs/Prepositions/ Participles/Focus Particle + C	já/ADV que/C para/P que/C visto/VB-AN que/C ainda/ADV que/C depois/P que/C mesmo/FP que/C até/P que/C
QUE	degree clauses		E aplicou tantos/ADJ-R-P esforços que/C impediu o levantar da barreira Fez um tal/ADJ-R esforço que/C não pode.... E fico tão/ADV-R consolado com os favores do Céu, que/C os maiores Impérios da terra regeitara eu. por mais/ADV-R força que/C aplicassem, não resolviam... Por mais/ADV-R que/C eu queira, (...)
	comparative clauses		...projetos maiores **que/C** aqueles....

 back to the top of the page
back to the main page of this manual

INTERROGATIVE/RELATIVE ELEMENTS

To relative elements, we apply tags WPRO or WPRO$. To interrogative elements, the following tags apply: WPRO, WADV and WQ for direct and indirect interrogatives, and WD for interrogative determiners.

TYPE	TAG	VALUE/ APPLICATION	EXAMPLES
	WPRO		O homem que/WPRO veio... O modo pelo qual/WPRO...
	WPRO-P WPRO-F-P		As casas nas quais/WPRO-P.... tantos.....quantos/WPRO-P forem os interessados nas companhias não me soube dizer Jerónimo Nunes quantas/WPRO-F-P eram
	WPRO$		O homem cujo/WPRO$ caráter....
RELATIVE ELEMENTS	WPRO$-F		...a consciência, cuja/WPRO$-F glória é aquietar entre as espinhas.

	WPRO$-P		O homem cujos/WPRO$-F modos...
	WPRO$-F-P		...as consciências, cujas/WPRO$-F-P glórias é aquietar entre as espinhas.
	WADV		O modo como/WADV vejo a Deus é tão sublime que..... O deserto onde/WADV me encontrava, quando.... Quanto/WADV à compra dos navios,
EXCLAMATIVE/ INTERROGATIVE ELEMENTS	WPRO		Por/P que/WPRO não me deste...? O/D que/WPRO queres...?
	WPRO-F	quanta	
	WPRO-P	quantos	
	WPRO-F-P	quantas	....porque não me soube dizer Jerónimo Nunes quantas/WPRO-F-P eram....
	WADV		E para saber isto como/WADV é, tende sentido... Quão/WADV verdadeiro é o sentimento (...)
	WQ		E perguntavas se/WQ não havias de se entregar.
EXCLAMATIVE/ INTERROGATIVE DETERMINERS	WD WD WD	Que processo Quanto sofrimento Qual processo	Qual/WD processo/N procuras? Que/WD processo/N procuras? Quanto/WD desamor/N encontrastes, quando.....?
	WD WD-F WD-P	Que processos Quanta dúvida Quantos/quais processos	Que/WD processos/N-P...? terá quanta/WD-F gente quizer Quais/WD-P processos/N-P...? Quantos/WD-P desamores/N-P encontrastes, quando...? Os/D-P quais/WD-P livros/N-P....
	WD-F-P	Quantas	Quantas/WD-F-P esperanças/N-P...!
OTHERS	WADV		Quanto/WADV a/P Vossa Mercê, peço a Deus que..... (see [prepositions + WADV](#))

PREPOSITIONS

Tag P is used to identify prepositions. In cases of lexical contraction, tags D or DEM can be associated to it.

TYPE	TAG	VALUE/ APPLICATION	EXAMPLES
PREPOSITIONS	P	General prepositions: a, de, com, em, para, até, sobre, desde, exceto, fora (=exceto), salvo, mediante, etc.	

	P+D (P+D-P; P+D-F; P+D-F-P; P+D-UM; P+D-UM-P; P+D-UM-F; P+D-UM-F-P)	à(s), ao(s), no(s), na(s), do(s), da(s), desse(s), dessa(s), pelo(s), pela(s)	
	P+PRO	dele, dela, etc	
	P+OUTRO (P+OUTRO-P; P+OUTRO-F; P+OUTRO-F-P)	d´outro (s), doutro(s), doutra, noutra, etc.	quando uma consoante se converte noutra/P+OUTRO-F
CONTRACTIONS	P+Q (P+Q-P; P+Q-F; P+Q-F-P)	n' algum, n'alguns, n'alguma, n'algumas	
	P+WPRO	porque (in interrogatives)	¿ porque/P+WPRO o não fará ?
	P+DEM	disso, disto, nisso, nisto	
	P+ADV	até'gora, antes d'hontem, etc.	
	P+WADV	donde, aonde, adonde	
	P+CL	pela, polo, etc.	trabalhei pela/P+CL servir
	P+NPR	del rey	

Remember that tag P is applied to many other contexts, as *até/P amanhã/ADV* and *até/P logo/ADV* (see adverbs as PP), *até/P que/C* (see complex subordinating conjunctios), *acerca/ADV de/P* and *apesar/P de/P* (see clusters)

NEW CLUSTERS: (previously treated as "WORDS AS UNIT")

Prepositional clusters *apesar de* and *acerca de*, as as well as conjunctional clusters *no entanto* and *não obstante* have had their taggs changed.

The history of languages in general reveals that these groups of words result, in general, of various grammaticalization processes. From the diachronic point of view, this means that their components migrate systematically from a class of words to another.

The decision of what tags to attribute to those items (separately or as a unit) must result from a compromise between linguistic precision and computational efficiency. A possible consequence of tagging them separately is the impossibility of automatic selection of these data from the *corpus*. In order to avoid that, and following the Penn-Helsinki Parsed Corpus of Middle English, the first option in the Tycho Brahe Corpus was to codify those items as words-as-unit (*apesar/P-21 de/P-22*, *acerca/P-21 de/P-22*, *no/CONJ-21 entanto/CONJ-22* and *não/CONJ-21 obstante/CONJ-22*). However, from the computational point of view, this kind of notation represents an idiosyncrasy, and has proven to cause significant noise to the system. Experience revealed, then, that in order to optimize the computational system, the prepositional and conjuncitonal clusters should no longer be tagged as 'words as unit'.

From now on, these items will be individually tagged, as follows:

ITEM	OLD TAG	NEW NEW TAG
apesar de	P-21 ... P-22 apesar/P-21 de/P-22	**P ... P** apesar/P de/P apesar/P do/P+D; apesar/P dos/P+D-P; apesar/P da/P+D-F; apesar/P das/P+D-F-P; apesar/P disso/P+DEM; apesar/P dele/P+PRO, etc.
acerca de	P-21 ... P-22 acerca/P-21 de/P-22	**ADV ... P** acerca/ADV de/P acerca/ADV do/P+D; acerca/ADV dos/P+D-P; acerca/ADV da/P+D-F, acerca/ADV das/P+D-F-P; acerca/ADV disso/P+DEM; acerca/ADV dele/P+PRO, etc.
no entanto	CONJ-21 CONJ-22 no/CONJ-21 entanto/CONJ-22	**P+D ... ADV** no/P+D entanto/ADV
não obstante	CONJ-21 CONJ-22 não/CONJ-21 obstante/CONJ-22	**NEG ... ADJ-G** não/NEG obstante/ADJ-G

(criteria to define the parts of speech were inspired in *Dicionário Contemporâneo da Língua Portuguesa* , F.J Caldas Aulete, 3ª edição atualizada, Lisboa, 1948 (parceria António M. Pereira)

Note that *'já que'*, *'visto que'*, and so on are (still) treated as complex subordinating conjunctions, while the expression *'quanto a'* is tagged as quanto/WADV a/P

OUTRO NEW

The set *outro*, *outros*, *outra*, *outras* receives now a new treatment, aiming at computational adequacy – but most of all, at facilitating search into the corpus.

Previously, a special tag /OUTRO was used to identify *outro* in anaphoric/reflexive use, as in *Abraçaram-se uns aos outros* (English: They held each other), and distinguish this from the adjectival use of the form, as in *Viu o outro homem* (English: he saw the other man), in which case the item was tagged /ADJ. This distinction was devised in order to reflect an important syntactic aspect and facilitate research; however, experience revealed that the distinction is sometimes too hard to make out. Apart from clear-cut cases like the above, the texts revealed a grey area in which the decision of which tag to apply became erratic. Inconsistencies in the application of either tag led to the decision that one of them had to be supressed. Splitting the tags only made sense, from the point of view of searching, if it could be trusted.

Because it is the more specific option, /OUTRO was chosen to be applied (with inflectional tags when necessary) to all occurrences of *outro(s)*, *outra(s)* – in this way, when searching for the more specific use of these items, researchers will not have to go through the open set of words tagged /ADJ. Of course, many adjectival *outros* will be included in the results, but this is certainly less counter-productive.

This is how the tagging will work:

TYPE	TAG	VALUE/APPLICATION	EXAMPLES
OUTRO	OUTRO	Outro (*in all contexts*)	Foram ao encontro um do outro/OUTRO O outro/OUTRO homem que partiu...
	OUTRO-P OUTRO-F OUTRO-F-P	Outros Outra Outras (*in all contexts*)	Acabava por não saber como agradá-los: a uns oferecia amor, a outros/OUTRO-P dedicação..... Abraçaram-se uns aos outros/OUTRO-P Uma e outra/OUTRO-F coisa se deve evitar

---

FOCUS PARTICLES

TYPE	TAG	VALUE	EXAMPLES
FOCUS PARTICLES	FP	'só', 'mesmo', 'até', and so on	Só/FP os padres podem.... ***versus*** Um só/ADJ livro foi enviado por.... Até/FP Sua Senhoria diria que... ***versus*** se pudesse ir até/P a presença de Sua Senhoria, ...

CARDINAL NUMBERS

Ordinal numbers behave as adjectives and, therefore, are tagged ADJ. Thus tag NUM applies to cardinal numbers only. Remember that NUM is not applied to *um/uma* (*one*), since they are treated as determiners, and tagged as D-UM/D-UM-F.

TYPE	TAG	VALUE/ APPLICATION	EXAMPLES
CARDINAL NUMBERS	NUM	General cardinal number expressed by digits or characters	Em 1538/NUM.... Três/NUM toneladas.... 100/NUM quilos.... um/D-UM milhão/N e/CONJ dois/NUM mil/NUM florins/N-P
	NUM-F	The form 'duas'/2	Duas/NUM-F centenas.... Por 20/NUM homens e 2/NUM-F fragatas....

NEGATION

🆕 This is the complete, updated list of items that receive the NEG tag in the Corpus:

TYPE	TAG	VALUE/ APPLICATION	EXAMPLES
NEGATIVE PARTICLES	NEG	'não'	Não/NEG se fará....... dizem/VB-P um/D-UM não/NEG que/WPRO parece/VB-P sim/ADV não/NEG obstante/ADJ (see clusters)
	SENAO	'senão'	Não respira senão/SENAO morte.....
OTHER NEGATIVE ITEMS	CONJ-NEG	'nem' 🆕 in all contexts	Não fez o pedido, nem/CONJ-NEG sabia dizer quem o fizera.... Nem/CONJ-NEG só de pão vive um ser...
	ADV-NEG	'nunca'	
	Q-NEG	'nada', 'ninguém', 'nenhum'	
	Q-NEG-P, Q-NEG-F, Q-NEG-F-P	'nenhuns'; 'nenhuma'; 'nenhumas'.	

NEW As pointed out by Anthony Kroch and [Ana Maria Martins](#), the non-application of NEG tag to negative adverbs, as *nunca* (e.g. *never*), negative quantifiers, as *ninguem* (e.g. *nobody*), and so on can cause searching problems, since it does not permit an automatic selecting of all negative structures on the corpus. A brief solution would be to add the NEG tag to the parsed files only, as in [1], forcing an NEG interpretation of the sentences, and then making the automatic searching of negative structures possible:

```
[1] (IP-MAT (CONJP (IP-MAT (Declamei/VB-D) (PP (contra/P) (DP (a/D-F) (NP    (vaidade/N))))) (,/,) (e/CONJ) ((IP-MAT-NEG (nunca/ADV) (pude/VB-D)
```

Notice, however, that the solution in [1] could not be enough to localize simple negative sentences, excluding sentences with negative agreement, or negative polarity items, *or* just find sentences with negative agreement, or negative polarity items only, *without* simple negative sentences. This is because a way to localize them in parsed files would be to automatically look for IPs that dominate a single NEG word (in order to localize simple negative sentences) *or* IPs that dominate more than one NEG word (in order to localize sentences with negative agreement).

Then, it seems that the best way to solve this problem was to change the present morphological annotation system, and add NEG to all the negative words (*não/NEG*, *nem/NEG*, *nunca/ADV-NEG*, *ninguém/Q-NEG*, etc.) present in the Tycho Brahe texts.

For computational reasons, the tag applied to *senão* is now typed with no diacritics, *i.e.* SENAO, and not SENÃO anymore.

INTERJECTIONS

Tag INTJ applies to items that belong to the interjections class.

TYPE	TAG	VALUE/APPLICATION	EXAMPLES
INTERJECTIONS	INTJ	General interjections	Oh/INTJ

FOREING AND UNKNOWN WORDS

TYPE	TAG	VALUE/APPLICATION	EXAMPLES
FOREIGN WORDS	FW	General foreign words	repartindo-se/ pro/FW rata/FW Pax/FW Christi/FW
UNKNOWN WORDS	XX	General unknown words	

PUNCTUATION

TYPE	TAG	VALUE/APPLICATION	EXAMPLES
FINAL PUCTUATION	.	Period or reticence	./.
		Interrogative mark	?/.
		Exclamation mark	!/.
		Semi-comma	;/.
		Colon	:/.
NON-FINAL PUNCTUATION	,	Comma	,/,
QUOTATION	QT **NEW**	Quotation mark	"/QT
PARENTHESES	(	Parentheses	(/(.................)/(
		Dash	-/(

SOME HOMOGRAPHIC LEXICAL ITEMS AND THEIR TAGS: REVIEW

The following table, displayed in alphabetic order, presents some of the homographic words commented above and their respective tags.

NEW. Note that the item *que* has lost one of its formerly possoble tags: /CONJS no longer applyes (see que)

LEXICAL ITEM	TAGS	EXAMPLES
ATÉ	**/P** **/FP**	meter a maõ até/P o cotovello pelo alheo até/FP nas arvores há ladroens
COMO	**/C** **/WADV**	
MELHOR/PIOR	/**ADJ-R-G** (with inflectional tag -P, when necessary) **/ADV-R** (always with no inflectional tags)	
MESMO	**/ADJ** (with inflexional tags (-F or -P), when necessary) **/FP** (always with no inflectional tags)	
PORQUE	**/C** **/WADV** **/P+WPRO**	
QUE	**/C,** **/WPRO**	
TAL	**/D-G** (with inflexional tag -P, when necessary) **/DEM** (always with no inflectional tags) /**ADJ-R-G** (with inflexional tag -P, when necessary) **/ADV-R** (always with no inflectional tags)	seus primos, filhos do tal/D-G tio quem tal/DEM fez, que naõ faria ? de tal/ADJ-R-G qualidade , que naõ he possível referillas todas e mostra-se , que tal/ADV-R deve ser a paz
TANTO	/**ADJ-R** (with inflectional tags (-F or -P), when necessary) **/ADV-R** (always with no inflectional tags)	

ITEMS WITH FIXED TAGS: REVIEW NEW

The following table displays those items that receive fixed tags – i.e., in all ocurrences, in any context (in alphabetical order):

LEXICAL ITEM	TAG
MAIS/MENOS	**ADV-R**
NADA	**Q-NEG**
NÃO	**NEG**
NEM	**CONJ-NEG**
NINGUÉM	**Q-NEG**
NENHUM nenhuns nenhuma nenhumas	**Q-NEG** (with inflectional tags –P and/or F, when necessary)
NUNCA	**ADV-NEG**
OUTRO	**OUTRO**

outros outra outras	(with inflectional tags –P and/or F, when necessary)
TÃO	**ADV-R**

COMPLETE LIST OF TAGS NEW

Click here for a list of all tags used in the Tycho Brahe Corpus, in alphabetical order.

If you have any doubts or suggestions, please contact galvesc@ime.usp.br