

The ICE Tagging Manual

Revised Version, 2005

Gerald Nelson

1. Introduction

This manual is an updated version of the *ICE Tagset Manual*, by Sidney Greenbaum (1995). The ICE tagging scheme was devised by the Survey of English Usage, in collaboration with the TOSCA Research Group at the University of Nijmegen. It is based in part on Quirk *et al*, *A Comprehensive Grammar of the English Language* (1985), though there are some significant differences between the two.

1.1 The ICE Tagset

The repertoire of wordclass tags is known as the ICE Tagset. The tagset consists of 20 main wordclasses. These are listed below, followed by their corresponding tags:

Wordclass	ICE Tag
Adjective	ADJ
Adverb	ADV
Article	ART
Auxiliary verb	AUX
Cleft <i>it</i>	CLEFTIT
Conjunction	CONJUNC
Connective	CONNEC
Existential <i>there</i>	EXTHERE
Formulaic expression	FRM
Genitive marker	GENM
Interjection	INTERJEC
Nominal Adjective	NADJ
Noun	N
Numeral	NUM
Preposition	PREP
Proform	PROFM
Pronoun	PRON
Particle	PARTCL
Reaction signal	REACT
Verb (lexical)	V

Punctuation marks and pauses are also tagged.

Wordclass tags consist of one of these main wordclass symbols, in upper case, followed (usually) by tag features in lower case and in parentheses. Tags, then, have the general form

WORDCLASS(feature)

For example, adjectives carry the main wordclass symbol **ADJ**, followed by a feature indicating their form. So comparative adjectives are labelled

ADJ(comp)

If the tag carries more than one feature, these are separated by a comma:

WORDCLASS(feature1,feature2)

For example, verbs carry the main wordclass tag **V**. This is followed by one feature indicating their complementation pattern and another feature indicating their form. So an intransitive (**intr**) verb in the present tense (**pres**) is tagged

V(intr,pres)

1.2 Multi-word Expressions

In general, each lexical item has been assigned its own grammatical tag. However, certain compound expressions have been assigned compound tags if they are considered to function grammatically as single units. Each word in the expression is assigned the tag of the expression as a whole.

<i>in</i>	PRTCL(to):1/2
<i>order</i>	PRTCL(to):2/3
<i>to</i>	PRTCL(to):3/3

Personal names, titles of books, and headings, are tagged in this way, as singular, proper nouns, without any internal analysis:

<i>John</i>	N(prop,sing):1/2
<i>Brown</i>	N(prop,sing):2/2
<i>The</i>	N(prop,sing):1/4
<i>Duke</i>	N(prop,sing):2/4
<i>of</i>	N(prop,sing):3/4
<i>Norfolk</i>	N(prop,sing):4/4
<i>Of</i>	N(prop,sing):1/4
<i>Mice</i>	N(prop,sing):2/4
<i>and</i>	N(prop,sing):3/4
<i>Men</i>	N(prop,sing):4/4

We refer to these compound tags as *ditto tags*. In ICECUP's main display window, ditto tagged items are indicated by yellow underlining. In the tree window, they are indicated by a yellow brace. The numbers (1/2, 2/2, etc) appears only when you save the results of a search, selecting the Tagged Text option.

2. ADJECTIVE (ADJ)

Adjectives carry the main wordclass label **ADJ**, followed by a label for one of the following formal features:

2.1 Compararive Adjective (comp)

<i>brighter</i>	ADJ(comp)
<i>happier</i>	ADJ(comp)

2.2 Superlative Adjective (sup)

<i>brightest</i>	ADJ(sup)
<i>happiest</i>	ADJ(sup)

2.3 Participial Adjective

Participial adjectives are tagged either **edp** or **ingp**, depending on their form

an <i>endangered</i> species	ADJ(edp)
a <i>broken</i> arm	ADJ(edp)
a <i>surprising</i> result	ADJ(ingp)
a <i>rewarding</i> experience	ADJ(ingp)

2.4 General Adjective (ge)

This subclass consists of all adjectives that do not belong to any of the other subclasses.

<i>bright</i>	ADJ(ge)
<i>happy</i>	ADJ(ge)

NOTE: Adjectives in periphrastic, comparative constructions, such as *more expensive* and *most expensive*, are tagged **ADJ(ge)**, since they are not formally marked.

3. ADVERB (ADV)

Adverbs carry the main wordclass label **ADV**. The class is divided into eight subclasses, which appear as features in the tag. These subclasses are

additive	add
exclusive	excl
intensifier	inten
particularizer	partic
phrasal	phras
relative	rel
<i>wh-</i>	wh
general	ge

We discuss each of these subclasses below.

3.1 Additive Adverb (add)

Additive adverbs are tagged **ADV(add)**.

Warfare was <i>both</i> a demonstration of right and a gesture	ADV(add)
Lansbury was <i>neither</i> afraid nor resentful.	ADV(add)
I think he worked in a bank <i>too</i> at one stage.	ADV(add)

3.2 Exclusive (excl)

Exclusive adverbs are tagged **ADV(excl)**.

It's <i>only</i> a novel	ADV(excl)
I was <i>merely</i> pointing out what he already knew	ADV(excl)
You <i>just</i> don't understand	ADV(excl)

3.3 Intensifier (inten)

Intensifiers denote a place on a scale of comparison, and include amplifiers and downtoners. They are tagged **ADV(inten)**.

She's <i>very</i> clever	ADV(inten)
It's <i>too</i> late to go out	ADV(inten)
The hall was <i>fairly</i> full	ADV(inten)

3.4 Particularizer (partic)

Particularizers emphasize that the utterance is restricted to the focused part. They are tagged **ADV(partic)**. *At least*, *at most*, and *in particular* are compound particularizers.

I am <i>mainly</i> concerned about his attitude	ADV(partic)
The youngsters, <i>in particular</i> , enjoyed the show	ADV(partic)
I knew <i>at least</i> ten people there	ADV(partic)

3.5 Phrasal adverb (phras)

Adverbs are tagged **ADV(phras)** when they enter into a combination traditionally known as a phrasal verb.

He gives <i>up</i> too easily	ADV(phras)
She looked it <i>up</i> in a dictionary	ADV(phras)
She looked <i>up</i> the word	ADV(phras)
He caught <i>on</i>	ADV(phras)

The adverb in the traditional phrasal-prepositional verb is tagged in the same way. For example, *up* in *put up with* is tagged **ADV(phras)**, and the phrasal preposition *with* is tagged **PREP(phras)**.

3.6 Relative adverb (rel)

Relative adverbs are tagged **ADV(rel)**. The relative adverbs *when*, *where*, *whereby*, and *why* introduce postmodifying relative clauses.

It happened during the week <i>when</i> I was in bed with the flu	ADV(rel)
I can't find the phone number of the hotel <i>where</i> she is staying	ADV(rel)
It provides a method <i>whereby</i> differences may be identified	ADV(rel)
They explained the reason <i>why</i> they were absent	ADV(rel)

Relative adverbs may be used to introduce the second part of a cleft sentence:

Was it a year ago *when* you were having back pain? **ADV(rel)**

3.7 Wh-adverbs (wh)

Wh-adverbs are tagged **ADV(wh)**. This subclass comprises all adverbs beginning with *wh*- plus the adverbs *how* and *however*. The adverbs in this subclass introduce clauses that are (a) independent exclamatory, (b) dependent exclamatory, (c) independent interrogative, (d) dependent interrogative, (e) nominal relative.

(a) <i>How</i> tall she is!	ADV(rel)
(b) You can't imagine <i>how</i> tall she is	ADV(rel)
(c) <i>When</i> can we meet for lunch?	ADV(rel)
(d) He asked her <i>how</i> to complete the form	ADV(rel)
(e) Paris is <i>where</i> you should go in the spring	ADV(rel)

If *when*, *whenever*, *where*, *whenever* introduce an adverbial clause, they are tagged as subordinating conjunctions.

3.8 General Adverb (ge)

This subclass consists of all adverbs that do not belong to any of the other subclasses. They are tagged **ADV(ge)**. The subclass includes *arguably*, *often*, *recently*, *slowly*, *there*, *yesterday*, as well as *AD*, *BC*, *am.*, *pm.*, *ibid.*, *etc.* *et al*, *per cent*.

Inflected adverbs - mostly general adverbs - have an additional feature indicating comparative **comp** or superlative **sup** form:

<i>fast</i>	ADV(ge)
<i>faster</i>	ADV(ge,comp)
<i>fastest</i>	ADV(ge,sup)

4. ARTICLE (ART)

Articles are assigned the main wordclass label **ART**, and they carry one of the feature labels **def** (definite) or **indef** (indefinite).

the **ART(def)**

a/an **ART(indef)**

5. AUXILIARY VERB (AUX)

Auxiliary verbs are tagged **AUX** for wordclass. This is followed by at least two features. The first feature indicates the subclass of the auxiliary. The subclasses are shown below, with their corresponding feature label:

<i>do</i> auxiliary	do
<i>let</i> auxiliary	let
modal	modal
passive	pass
perfective	perf
progressive	prog
semi-auxiliary	semi
semi-auxiliary followed by an <i>-ing</i> participle	semip

The second feature indicates the form, selected from the following:

<i>-ed</i> participle form	edp
<i>-ing</i> participle form	ingp
past	past
present	pres
infinitive	infin
imperative	imp

Where appropriate, additional features indicate the following:

negative	neg
enclitic	encl
proclitic	procl
elliptical	ellipt
discontinuous	disc

Here we will look at each of the auxiliary subclasses.

5.1 *Do* auxiliary (do)

This subclass consists of the dummy operator *do* and the introductory imperative marker *do*. All instances are marked **AUX(do,...)**.

How *did* it start? **AUX(do,past)**

It just *doesn't* work **AUX(do,pres,neg)**

<i>Do</i> come in	AUX(do,imp)
<i>Don't</i> sit there	AUX(do,imp,neg)
<i>D'</i> you want some cake?	AUX(do,pres,procl)

5.2 *Let* auxiliary (let)

The introductory imperative marker *let* is tagged **AUX(let,imp)**.

<i>Let</i> me just see if I can give you a little test	AUX(let,imp)
<i>Let's</i> see if there are any buyers	AUX(let,imp)

Auxiliary *let* is distinguished from the lexical verb *let* ('allow'), as in '*Let* me go'.

5.3 Modal Auxiliary (modal)

Modal auxiliaries are tagged **AUX(modal,...)**. The modal auxiliaries are *can, may, shall, will, must, could, might, should, would*

We <i>can</i> leave early tomorrow morning	AUX(modal,pres)
The surveyor <i>will</i> be here soon	AUX(modal,pres)
You really <i>shouldn't</i> speak to me like that	AUX(modal,past,neg)
I <i>'ll</i> be there soon	AUX(modal,pres,encl)

5.4 Passive Auxiliary (pass)

The passive auxiliaries: *be* and *get* are tagged **AUX(pass,...)**.

They <i>were</i> arrested	AUX(pass,past)
I don't want to <i>be</i> poisoned.	AUX(pass,infin)
Everything else has <i>been</i> stopped.	AUX(pass,edp)
These temples <i>got</i> abandoned in medieval times	AUX(pass,past)

5.5 Perfect Auxiliary (perf)

The perfect auxiliary *have* is tagged **AUX(perf,...)**.

He has put his certificate on the wall	AUX(perf,pres)
It <i>hasn't</i> got a good reputation	AUX(perf,pres,neg)
Nothing <i>had</i> been done	AUX(perf,past)
She <i>'d</i> already eaten when I arrived	AUX(perf,past,encl)

5.6 Progressive Auxiliary (prog)

The progressive auxiliary *be*. is tagged **AUX(prog,...)**.

You <i>were</i> working closely with her	AUX(prog,past)
They <i>'re</i> just pretending	AUX(prog,pres,encl)
You must <i>be</i> running out of time	AUX(prog,infin)
I think that's <i>been</i> playing on my mind	AUX(prog,edp)

5.7 Semi-auxiliary (semi)

Semi-auxiliaries are tagged **AUX(semi,...)**. The class includes modal idioms and catenatives. The subclass includes *appear to*, *be about to*, *be likely to*, *have to*, *tend to*. All the semi-auxiliaries are ditto-tagged:

<i>is</i>	AUX(semi,pres):1/3
<i>about</i>	AUX(semi,pres):2/3
<i>to</i>	AUX(semi,pres):3/3

If the parts of a semi-auxiliary do not occur adjacent to each other, they carry an additional feature **disc** (discontinuous). For example:

<i>was</i>	AUX(semi,past,disc):1/3
<i>just</i>	ADV(excl)
<i>going</i>	AUX(semi,past,disc):2/3
<i>to</i>	AUX(semi,past,disc):3/3

However, if modifiers of the adjectives in semi-auxiliaries are present, they are included in the ditto tags:

<i>was</i>	AUX(semi,past):1/4
<i>almost</i>	AUX(semi,past):2/4
<i>certain</i>	AUX(semi,past):3/4
<i>to</i>	AUX(semi,past):4/4

A repeated semi-auxiliary may be elliptical, but it is tagged in the same way as the full form, except that it is given the added feature **ellipt**:

Are you supposed to take your exam this June?
Yes, I *am*.

AUX(semi,pres,ellipt)

He's apt to exaggerate, *isn't* he?

AUX(semi,pres,neg,ellipt)

NOTE: When *be* performs more than one function at the same time, by convention we tag only the first function. In the following example, *was* is both a progressive auxiliary (*was playing*) and a passive auxiliary (*was judged*). The tag is determined by the first function:

Nigel *was* playing beautifully but judged too inexperienced. **AUX(prog,past)**

5.8 Semi-auxiliary followed by *-ing* participle (semip)

These resemble the semi-auxiliaries (4.1.7) except that they are followed by an *-ing* participle. They are tagged **AUX(semip,...)**. The *-ing* participle is not part of the semi-auxiliary.

I *keep* thinking I must do something about it. **AUX(semip,pres)**

They *began* loading bombs. **AUX(semip,past)**

She can't *stop* talking **AUX(semip,infinitive)**

I've *started* wearing T-shirts. **AUX(semip,edp)**

6. CLEFT *IT* (CLEFTIT)

The *it* in cleft constructions is tagged CLEFTIT, without any further features.

It was my sister who saw him **CLEFTIT**

Was *it* last year we met? **CLEFTIT**

NOTE: The main verb in cleft constructions is tagged *intransitive*

7. CONJUNCTION (CONJUNC)

We distinguish two type of conjunctions, coordinating conjunctions (**coord**) and subordinating conjunctions (**subord**). Both carry the main wordclass label **CONJUNC**.

7.1 Coordinating Conjunction (coord)

Coordinating conjunctions are labelled **CONJUNC(coord)**. The following items are tagged as coordinating conjunctions: *and, as well as, but, for, let alone, nor, or, plus, rather than, yet*.

NOTE: When the conjunctions *and, but, for, nor, or, plus, and yet* occur at the beginning of a text unit, they are tagged as general **connectives**, not as coordinators. *Nor* and *yet* are also tagged connectives when they follow *and* or *but*.

7.2 Subordinating Conjunction (subord)

The subordinators are tagged **CONJUNC(subord)**. They include *after, if, since, so, that, unless, until, when(ever)*.

Multi-word subordinators may be discontinuous, requiring the additional feature **disc**:

<i>provided</i>	CONJUNC(subord,disc)
<i>always</i>	ADV(ge)
<i>that</i>	CONJUNC(subord,disc)

They may also be elliptical:

our letters Your card has been cancelled, seeing that you won't pay and *that* you refuse to answer
CONJUNC(subord,ellipt)

8. CONNECTIVE (CONNEC)

The ICE grammar distinguishes two type of connectives, general connectives **CONNEC(ge)** and appositive connectives **CONNEC(appos)**.

8.1 General connective (ge)

General connectives are tagged **CONNEC(ge)**. They are used to establish a relation between the current clause or sentence and (one or more) previous clauses or sentences.

Finally, the candidates were criticised for 'woolly answers' **CONNEC(ge)**

Also I note that the floor plan is grossly inaccurate **CONNEC(ge)**

There are, *however*, a couple of obstacles to be overcome first **CONNEC(ge)**

For the tagging of coordinating conjunctions as **CONNEC(ge)**, see 5.1.

8.2 Appositive connective (appos)

Appositive connectives are tagged **CONNEC(appos)**. They typically occur between items which are in apposition.

Things *like* basketball and football are great exercise **CONNEC(appos)**

Other factors, *such as* lack of amenities, are also important **CONNEC(appos)**

The government resists all pressure groups *particularly* trade unions **CONNEC(appos)**

The feature **disc** indicates discontinuous appositive connectors:

<i>that</i>	CONNEC(appos,disc):1/4
<i>is</i>	CONNEC(appos,disc):2/4
perhaps	ADV(ge)
<i>to</i>	CONNEC(appos,disc):3/4
<i>say</i>	CONNEC(appos,disc):4/4

9. EXISTENTIAL *THERE* (EXTHERE)

Existential *there* is tagged **EXTHERE**. This tag does not carry any features.

There's nobody else he can trust. **EXTHERE**

That's the kind of area where *there* is uncertainty. **EXTHERE**

NOTE: The main verb in existential constructions is tagged *intransitive*.

10. FORMULAIC EXPRESSION (FRM)

Formulaic expressions are tagged **FRM**, without any further features.

The class includes greetings and farewells (*adieu, bye, goodbye, hello, Merry Christmas*), thanks (*cheers, thanks, thank you*), and apologies (*excuse me, I beg your pardon, sorry*). It also include expletives (*Christ, damn, fuck, shit*) and miscellaneous expressions (*bravo, congratulations, okay, please*).

11. GENITIVE MARKER (GENM)

The genitive marker 's or ' is separated from the word to which it is attached, and is assigned its own tag, **GENM**.

Susan

's **GENM**
book

the
students
' **GENM**
essays

12. INTERJECTION (INTERJEC)

Interjections are emotive words that do not enter into syntactic relations. Examples include *aha*, *boo*, *ha*, *oops*, *wow*. The class also includes the voiced pauses *uh* and *uhm*. All interjections are tagged **INTERJEC**, without any features.

See also Reaction Signals.

13 NOUN (N)

Nouns carry the wordclass label **N**. This is followed by two features. The first distinguishes between common (**com**) and proper (**prop**) nouns, and the second indicates number - singular (**sing**) or plural (**plu**).

13.1 Singular (**sing**) and Plural (**plu**) Nouns

The assignment of singular and plural relies predominantly on form, and no distinction is made between singular count nouns and noncount (or mass) nouns. The following italicised nouns are therefore tagged **sing**:

your <i>furniture</i>	industrial <i>research</i>
this <i>information</i>	all kinds of <i>wine</i>
the <i>wine</i>	the <i>truth</i>
<i>coffee</i>	some more <i>paper</i>
some <i>difficulty</i>	your <i>courage</i>

Singular collective nouns are tagged **sing**; for example: *board*, *gang*, *team*, *cornmittee*. So too are *news*; names of disciplines, etc., ending in *-ics* (e.g. *mathematics*, *physics*, *politics*, *athletics*); names of diseases ending in *-s* (e.g. *measles*, *mumps*); and names of certain games ending in *-s* (e.g. *dominoes*, *darts*). However, some of these nouns may be used with number contrast, and in such cases the final *-s* marks the plural; for example: *a statistic*, *some statistics*; *a dart*, *two darts*.

Some nouns that are not morphologically marked as plural are tagged **plu** because they require a plural verb:

the <i>cattle</i> are stampeding	N(com,plu)
the <i>gentry</i> were overthrown in the rebellion	N(com,plu)

13.2 Common (com) and Proper (prop) Nouns

The distinction between common noun and proper noun is made simply on the basis of the absence or presence of an initial capital letter. If a noun begins without a capital it is a common noun, if it begins with a capital it is a proper noun (unless the capital is only required to mark the beginning of a sentence):

in		in	
<i>time</i>	N(com,sing)	<i>June</i>	N(prop,sing)

13.3 Compound Nouns

To facilitate the parsing process, the concept of compound noun has been broadened to encompass every sequence of two or more nouns with a noun as Head that constitutes a unit. The nouns in the sequence are assigned *ditto tags*, determined by Head of the sequence.

<i>railway</i>	N(com,sing):1/2
<i>station</i>	N(com,sing):2/2
<i>potato</i>	N(com,plu):1/2
<i>crisps</i>	N(com,plu):2/2
<i>English</i>	N(com,sing):1/3
<i>Department</i>	N(com,sing):2/3
<i>office</i>	N(com,sing):3/3
<i>European</i>	N(com,plu):1/4
<i>Community</i>	N(com,plu):2/4
<i>finance</i>	N(com,plu):3/4
<i>ministers</i>	N(com,plu):4/4

13.4 Mentions

Expressions that are mentioned as linguistic objects are treated as common singular nouns. They are explicitly marked as *mentions* in the structural markup.

<i>Cats</i> is a noun	N(com,sing)
I don't like the word <i>prioritize</i> .	N(com,sing)
<i>Sorry</i> is hardly the right word	N(com,sing)

13.5 Genitive Nouns

Genitive nouns with determiner function (or in a noun phrase with determiner function) are not part of the sequence and are therefore tagged independently; for example, *husband's* in *her husband's death*:

<i>her</i>	PRON(poss,sing)
<i>husband</i>	N(com,sing)
<i>'s</i>	GENM
<i>death</i>	N(com,sing)

13.6 Appositive Nouns

Nouns in apposition are tagged independently; for example, *George Brown, Professor of Physics*:

<i>George</i>	N(prop,sing):1/2
<i>Brown</i>	N(prop,sing):2/2
<i>,</i>	PUNC(com)
<i>Professor</i>	N(prop,sing)
<i>of</i>	PREP(ge)
<i>Physics</i>	N(prop,sing)

Contrast the tagging of *Professor George Brown*, where *Professor* is a title:

<i>Professor</i>	N(prop,sing):1/3
<i>George</i>	N(prop,sing):2/3
<i>Brown</i>	N(prop,sing):3/3

Similarly, *the word prioritize* in 12.4 above:

<i>the</i>	ART(def)
<i>word</i>	N(com,sing)
<i>prioritize</i>	N(com,sing)

13.7 Adjective-Noun Sequences

Some noun compounds consist of an adjective plus a noun. They are treated as compounds on the basis of their stress pattern (main stress on the first word) or their idiomaticity:

<i>hot</i>	N(com,plu):1/2
<i>dogs</i>	N(com,plu):2/2

If an adjective + noun compound premodifies a noun, the compound is tagged with the Head noun under the sequence rule stated at the beginning of this section:

<i>a</i>	ART(indef)
<i>High</i>	N(com,sing):1/3

<i>Court</i>	N(com,sing):2/3
<i>judge</i>	N(com,sing):3/3

13.8 Titles and Names

The title of a book, play, song, newspaper, etc. is tagged as a compound singular proper noun, without regard to the word classes of its constituents:

<i>The</i>	N(prop,sing):1/4
<i>Interpretation</i>	N(prop,sing):2/4
<i>of</i>	N(prop,sing):3/4
<i>Dreams</i>	N(prop,sing):4/4

<i>Making</i>	N(prop,sing):1/6
<i>a</i>	N(prop,sing):2/6
<i>Reality</i>	N(prop,sing):3/6
<i>of</i>	N(prop,sing):4/6
<i>Home</i>	N(prop,sing):5/6
<i>Care</i>	N(prop,sing):6/6

Punctuation, including the genitive marker, is included in the ditto tags:

<i>English</i>	N(prop,sing):1/5
<i>Grammar</i>	N(prop,sing):2/5
:	N(prop,sing):3/5
<i>An</i>	N(prop,sing):4/5
<i>Introduction</i>	N(prop,sing):5/5

14 NOMINAL ADJECTIVE (NADJ)

Nominal adjectives carry the main wordclass label **NADJ**, and some additional features. A major subclass denotes members of a nationality and has plural reference. These carry the feature **prop** (proper) because they begin with a capital:

the <i>English</i>	NADJ(prop)
the <i>Welsh</i>	NADJ(prop)
the <i>French</i>	NADJ(prop)

Nominal adjectives with the **prop** feature are not open to any other features.

We distinguish three further subclasses:

1. Words with plural reference to classes of people: These are tagged **NADJ(plu)**.

the *young* **NADJ(plu)**

the *jobless* **NADJ(plu)**

the *careless* **NADJ(plu)**

2. Words with abstract and singular reference.

The *worst* is yet to come. **NADJ(sing)**

A glimpse of the *obvious* **NADJ(sing)**

3. Words with a participial ending. They carry a form feature **edp** or **ingp**, and a number feature **sing** or **plu**.

Police are interviewing the *accused*. **NADJ(edp,sing)**

The *wounded* were carried away. **NADJ(edp,plu)**

Judgement is left to the *discerning*. **NADJ(ingp,plu)**

Like other adjectives, nominal adjectives may be marked for comparison **comp** or superlative **sup** form.

the *easier* of the two choices **NADJ(comp,sing)**

at my *best* **NADJ(sup,sing)**

15 NUMERAL (NUM)

Numerals carry the main wordclass label **NUM**. This is followed by a feature label indicating the subclass. We distinguish the following subclasses of numerals:

cardinal numeral	card
ordinal numeral	ord
fraction	frac
hyphenated	hyph
multiplier	mult

All numerals are marked for one of these subclasses. Where relevant, they are also marked for number according to their form; hence, *thousand* is singular (**sing**) and *thousands* is plural (**plu**).

NOTE: In written texts, numerals may appear as words (*a hundred*) or as digits (*100*). In spoken texts, they should always appear as words (*nineteen ninety-eight*, or *nineteen hundred and ninety-eight*, not *1998*).

15.1 Cardinal Numeral (card)

Cardinal numerals carry the feature label **card**, and a number feature **sing** or **plu**. Examples include *one* (with singular nouns), *two*, *threes*, *forty-two*, *one hundred*, *a hundred*, *two thousand*, *thousands*, *millions*, *a dozen*, *scores*. The subclass also includes *zero* and its synonyms.

<i>a</i>	NUM(card,sing):1/2
<i>hundred</i>	NUM(card,sing):2/2
<i>one</i>	NUM(card,sing):1/2
<i>hundred</i>	NUM(card,sing):2/2
<i>ten</i>	NUM(card,sing):1/2
<i>thousand</i>	NUM(card,sing):2/2
<i>the</i>	ART(def)
<i>sixties</i>	NUM(card,plu)

Combinations of cardinals with fractions (see 14.2) have been separated into two units. For example, $2\frac{1}{2}$ appears in the texts as *2 1/2*. These combination are ditto-tagged as cardinals:

<i>2</i>	NUM(card,sing):1/2
<i>1/2</i>	NUM(card,sing):2/2

If combinations of cardinal and fraction are set out as words, the words are tagged independently:

<i>one</i>	NUM(card,sing)
<i>and</i>	CONJUNC(coord)
<i>two-fifths</i>	NUM(frac,plu)

15.2 Ordinal numeral (ord)

The subclass of ordinals includes the primary ordinals, such as *first*, *second*, *10th*, *twenty-first*, and the following:

<i>additional</i>	<i>further</i>	<i>others</i>
<i>another</i>	<i>last</i>	<i>preceding</i>
<i>extra</i>	<i>latter</i>	<i>previous</i>
<i>following</i>	<i>next</i>	<i>same</i>
<i>former</i>	<i>other</i>	<i>subsequent</i>

15.3 Fraction (frac)

Fractions include *a half*, *one fifth*, *three-quarters*, *four-fifths*, *1/8*, *3/5*. They carry the feature label **frac** and a number feature **sing** or **plu**.

<i>one</i>	NUM(frac,sing):1/2
<i>fifth</i>	NUM(frac,sing):2/2
<i>a</i>	NUM(frac,sing):1/2
<i>half</i>	NUM(frac,sing):2/2
<i>three-quarters</i>	NUM(frac,plu)
<i>1/8</i>	NUM(frac,sing)

15.4 Hyphenated numeral (hyph)

Hyphenated numerals denote an inclusive range . They are labelled **NUM(hyph)**, with no other feature. The ‘hyphen’ is more properly in print a short dash or en-dash.

<i>40-65_</i>	NUM(hyph)
<i>1750-1830</i>	NUM(hyph)
<i>9-12 a.m</i>	NUM(hyph)
<i>1995-98</i>	NUM(hyph)
<i>March 1st-25th</i>	NUM(hyph)
<i>aged 14-16</i>	NUM(hyph)

The range may be indicated by a slash:

<i>1992/93</i>	NUM(hyph)
----------------	------------------

15.5 Multiplier (mult)

Multipliers include *once*, *twice*, *double*, *triple*. They carry the feature label **mult**, and no number feature.

<i>twice</i>	NUM(mult)
<i>two</i>	NUM(card)
<i>is</i>	V(cop,pres)
<i>four</i>	NUM(card)

16. PREPOSITION (PREP)

Prepositions carry the main wordclass label **PREP**, followed by the feature label *ge* (general), **phras** (phrasal), or **inter** (interrogative)

16.2 General preposition (*ge*)

General prepositions are tagged **PREP(*ge*)**. They may be simple prepositions, consisting of just one word, such as *about*, *by*, *for*, *of*, *to*, and *with*. We also recognise a large number of complex prepositions. This group includes *according to*, *by means of*, *except for*, *prior to*, *with reference to*, *thanks to*. Complex prepositions are ditto-tagged

<i>according</i>	PREP(<i>ge</i>):1/2
<i>to</i>	PREP(<i>ge</i>):2/2
<i>with</i>	PREP(<i>ge</i>):1/3
<i>reference</i>	PREP(<i>ge</i>):2/3
<i>to</i>	PREP(<i>ge</i>):3/3

Complex prepositions may be elliptical (**ellipt**):

We remain in contact with you and *with* your family **PREP(*ge*,*ellipt*)**

They may also be discontinuous (**disc**)

He is entitled to payment, *subject only to* the usual conditions **PREP(*ge*,*disc*)**

16.1 Phrasal Preposition (**phras**)

Prepositions that combine with verbs to form intransitive prepositional verbs are labelled **PREP(**phras**)**.

She	PRON(<i>pers</i> , <i>sing</i>)
looked	V(<i>intr</i> , <i>past</i>)
<i>at</i>	PREP(phras)
it	PRON(<i>pers</i> , <i>sing</i>)

Similarly, the preposition is tagged **PREP(**phras**)** in transitive prepositional constructions:

It	PRON(<i>pers</i> , <i>sing</i>)
protected	V(<i>montr</i> , <i>past</i>)
us	PRON(<i>pers</i> , <i>plu</i>)
<i>from</i>	PREP(phras)
the	ART(<i>def</i>)
rain	N(<i>com</i> , <i>sing</i>)

The preposition may be stranded after the verb, without its complement:

Who are you talking <i>about</i> ?	PREP(phras)
How long is she staying <i>for</i> ?	PREP(phras)
It's so hard to concentrate <i>on</i>	PREP(phras)

16.3 Interrogative Preposition (inter)

What about, how about, and what of are ditto-tagged **PREP(inter)**:

<i>What about</i> the new guy?	PREP(inter)
<i>How about</i> some ice-cream?	PREP(inter)
<i>What of</i> twentieth-century writers?	PREP(inter)

17. PROFORM (PROFM)

Proforms carry the main wordclass label **PROFM**. Three subclasses are distinguished:

conjoin	conj
<i>so</i>	so
<i>one</i>	one

17.1 Proform conjoin (conj)

Proform conjoins are tagged **PROFM(conj)**. The subclass includes the following items, all introduced by a coordinating conjunction. The conjunction is not part of the proform.

(or) <i>so</i>	(and/or) <i>both</i>
(and) <i>so forth</i>	(or) <i>neither</i>
(and) <i>so on</i>	(or) <i>not</i>
(and) <i>the like</i>	(or) <i>otherwise</i>
(and/or) <i>such like</i>	(and/or) <i>whatever</i>
(and/but/or) <i>the reverse</i>	(or) <i>what have you</i>
(and/or) <i>vice versa</i>	(or) <i>something</i>

It may go malignant in twenty years or *so* **PROFM(conj)**

We might have people to supper or *whatever* **PROFM(conj)**

I spoke about his qualities and work and *so on* **PROFM(conj)**

17.2 Proform *so* (so)

Proform *so* replaces phrases and clauses, and is tagged **PROFM(so)**.

I hope <i>so</i>	PROFM(so)
It says <i>so</i> right here	PROFM(so)
I'm hungry and <i>so</i> is everyone else	PROFM(so)

The clausal proform has a negative counterpart in the word *not*:

I hope <i>not</i>	PROFM(so)
-------------------	------------------

17.3 Proform *one* (one)

Proform *one* includes substitute *one* and generic *one*. Both uses are tagged **PROFM(one)**. Substitute *one* may be singular (**sing**) or plural (**plu**):

Can I borrow <i>one</i> ?	PROFM(one, sing)
I prefer the sweet <i>ones</i>	PROFM(one, plu)

Generic *one* is always singular:

What can <i>one</i> expect?	PROFM(one, sing)
<i>One</i> really shouldn't laugh	PROFM(one,sing)

18. PRONOUN (PRON)

Pronouns carry the main word class label **PRON** and a feature label for the subclass. We distinguish the following subclasses of pronoun:

anticipatory <i>it</i>	antit
assertive	ass
demonstrative	dem
exclamative	exclam
negative	neg
nonassertive	nonass
personal	pers
possessive	poss
quantifying	quant
reciprocal	recip
reflexive	ref
relative	rel
universal	univ

Where a distinction in number is relevant, the feature **sing** or **plu** is assigned. There is no assignment of case features.

18.1. Anticipatory *it*

Anticipatory *it* is tagged **PRON(antit)**:

It's a good thing (that) you came here **PRON(antit)**

I find *it* really difficult to sit down **PRON(antit)**

He made *it* clear that he intended to resign **PRON(antit)**

The proclitic form carries an additional feature **procl**:

'Tis OK to phone me at work **PRON(antit,procl)**

18.2 Assertive Pronoun (ass)

All assertive pronouns except *some* are tagged **PRON(ass,sing)**; *some* is left unmarked. The assertive pronouns are *some*, *somebody*, *someone*, and *something*. The genitive inflection is tagged in the usual way:

somebody **PRON(ass,sing)**
's GENM

18.3 Demonstrative Pronoun (dem)

The demonstrative pronouns are *that*, *these*, *this*, *those* and *such*. Except for *such*, they are marked for number as **PRON(dem,sing)** or **PRON(dem,plu)**.

18.4 Exclamative Pronoun (exclam)

Exclamative *what* is tagged **PRON(exclam)**

What a great week it's been **PRON(exclam)**

What fun! **PRON(exclam)**

18.5 Negative Pronoun (neg)

Negative pronouns are tagged **PRON(neg)**. The negative pronouns are *neither*, *nobody*, *no one*, *nothing*, *no*, *none*

Neither, *nobody*, *no one*, and *nothing* are tagged **PRON(neg,sing)**. *No* and *none* are tagged **PRON(neg)**.

18.6 Nonassertive pronoun (nonass)

The nonassertive pronouns are *any, anyone, either, anybody, anything*

All except *any* are tagged **PRON(nonass,sing)**; *any* is tagged **PRON(nonass)**

18.7 Personal pronoun (pers)

Personal pronouns are tagged **PRON(pers)**, and except for *you* they are also tagged for number but not for case. Examples:

she **PRON(pers,sing)**

us **PRON(pers,plu)**

it **PRON(pers,sing)**

you **PRON(pers)**

Also tagged as **PRON(pers,sing)** are abbreviations or combinations such as *s/he* and *him/her*.

Proclitic *it*, as in *'tis*, is tagged **PRON(pers,sing,procl)**; proclitic *you*, as in *y'know*, **PRON(pers,procl)**; enclitic *us*, as in *let's*, **PRON(pers,plu,encl)**.

Prop *it*, as in *'It's raining* or *'It's nine o'clock*, is tagged **PRON(pers,sing)**.

18.8 Possessive pronoun (poss)

Possessive pronouns are tagged **PRON(poss)**, and except for *your* and *yours* they are also tagged for number. The possessive pronouns are:

<i>my</i>	<i>your</i>	<i>his</i>	<i>their</i>
<i>mine</i>	<i>yours</i>	<i>her</i>	<i>theirs</i>
<i>our</i>	<i>hers</i>		
<i>ours</i>	<i>its</i>		

Also tagged as **PRON(poss,sing)** are abbreviations or combinations such as *his/her*.

18.9 Quantifying pronoun (quant)

Quantifying pronouns are tagged **PRON(quant)**, and some are tagged for number (**sing** or **plu**). The quantifying pronouns are:

<i>enough</i>	PRON(quant)
<i>few</i>	PRON(quant,plu)
<i>fewer</i>	PRON(quant,plu)

<i>fewest</i>	PRON(quant,plu)
<i>least</i>	PRON(quant)
<i>less</i>	PRON(quant)
<i>little</i>	PRON(quant,sing)
<i>many</i>	PRON(quant,plu)
<i>more</i>	PRON(quant)
<i>most</i>	PRON(quant)
<i>much</i>	PRON(quant,sing)
<i>plenty</i>	PRON(quant)
<i>several</i>	PRON(quant,plu)

18.10 Reciprocal pronoun (recip)

There are two reciprocal pronouns, *each other* and *one another*. They are tagged PRON(recip) with ditto tags:

<i>each</i>	PRON(recip):1/2
<i>other</i>	PRON(recip):2/2
<i>one</i>	PRON(recip):1/2
<i>another</i>	PRON(recip):2/2

18.11 Reflexive pronoun (ref)

Reflexive pronouns are tagged **PRON(ref)**, and they carry an additional feature label for number (**sing** or **plu**). The reflexive pronouns are:

<i>myself</i>	<i>yourself</i>	<i>himself</i>	<i>oneself</i>
<i>ourselves</i>	<i>yourselves</i>	<i>herself</i>	
<i>itself</i>			
<i>themselves</i>			

18.12 Relative pronoun (rel)

Relative pronouns are tagged **PRON(rel)**; number and case are not marked. The relative pronouns are:

<i>which</i>	<i>who</i>	<i>whom</i>	<i>whose</i>	<i>that</i>	<i>whereby</i>
--------------	------------	-------------	--------------	-------------	----------------

18.13 Universal pronoun (univ)

Universal pronouns are tagged **PRON(univ)**, and except for *all* they are also tagged for number (**sing** or **plu**).

<i>all</i>	PRON(univ)
<i>both</i>	PRON(univ,plu)
<i>each</i>	PRON(univ,sing)

19. PARTICLE (PRTCL)

Particles carry the main wordclass label **PRTCL** and one of the subclass features: **to**, **for**, or **with**.

18.1 Particle *to*

Particle *to* introduces an infinitive clause. It is tagged **PRTCL(to)**.

I rushed over to the library *to* get a couple of books **PRTCL(to)**

The subclass includes *in order to* and *so as to*:

They are withdrawing services *in order to* reduce capacity **PRTCL(to)**

Be punctual *so as to* reduce waiting time **PRTCL(to)**

The feature **disc** is used if the particle is discontinuous:

<i>in</i>	PRTCL(to,disc):1/3
<i>order</i>	PRTCL(to,disc):2/3
perhaps	ADV(ge)
<i>to</i>	PRTCL(to,disc):3/3

18.2 Particle *for*

Particle *for* introduces the subject of an infinitive clause. It is tagged **PRTCL(for)**

It's really *for* you to decide **PRTCL(for)**

It's time *for* you to go. **PRTCL(for)**

The subclass includes *in order for*:

In order for you to qualify, you must be under eighteen **PRTCL(for)**

The feature **disc** is used if the particle is discontinuous:

<i>in</i>	PRTCL(to,disc):1/3
<i>order</i>	PRTCL(to,disc):2/3
perhaps	ADV(ge)
<u>for</u>	PRTCL(to,disc):3/3

18.3 Particle *with*

Particle *with* introduces the subject of a nonfinite or verbless clause. It is tagged **PRTCL(with)**

Seagram wins the Grand National *with* twenty-five yards to go **PRTCL(with)**

You'll never get a word in *with* me talking **PRTCL(with)**

The subclass includes *without*:

Victory was achieved *without* a drop of blood being spilled **PRTCL(with)**

19. REACTION SIGNAL (REACT)

Reaction signals express agreement or disagreement with a previous speaker. They are tagged **REACT**, without any feature. The class includes *all right, fine, good, no, ok, right, yes*.

20. VERB (V)

Lexical verbs are tagged **V**. This is followed by at least two features. The first feature indicates the complementation pattern. The ICE grammar recognises the following seven patterns:

intransitive	intr
copular	cop
monotransitive	montr
dimonotransitive	dimontr
ditransitive	ditr
complex-transitive	cextr
transitive	trans

These patterns are discussed in the next section, 20.1

The second feature indicates the form of the verb, selected from the following:

<i>-ed</i> participle	edp
<i>-ing</i> participle	ingp
imperative	imp

infinitive	infin
past tense	past
present tense	pres
subjunctive	subjun

Where appropriate, additional features indicate the following:

discontinuous	disc
enclitic	encl
negative	neg

The features **encl** and **neg** apply only to the lexical verbs *be* and *have*.

20.1 Verb Complementation Patterns

20.1.1 Intransitive (**intr**)

Intransitive verbs (**intr**) are not followed by any object or complement.

She <i>works</i> in a library	V(intr,pres)
She graduated last summer	V(intr,past)
<i>Stop</i>	V(intr,imp)

20.1.2. Copular (**cop**)

Copular verbs (**cop**) require the presence of a subject complement.

It really <i>is</i> great fun	V(cop,pres)
It <i>isn't</i> my fault	V(cop,pres,neg)
She <i>'s</i> home	V(cop,pres,encl)
He <i>looks</i> the part as well	V(cop,pres)
I <i>felt</i> tired this morning	V(cop,past)
She <i>seems</i> upset about something	V(cop,pres)
He <i>acts</i> sometimes <i>as</i> my agent	V(cop,pres,disc)

Instances of *be* as a lexical verb are tagged as copular, except in cleft constructions and in existential constructions. In these cases, the verb is tagged intransitive.

20.1.3 Monotransitive (**montr**)

Monotransitive verbs (**montr**) are complemented by a Direct Object only.

I <i>buy</i> too many books	V(montr,pres)
She <i>left</i> her coat in the cinema	V(montr,past)
Please <i>sign</i> both sides of the form	V(montr,imp)
He just <i>hasn't</i> a clue	V(montr,pres,neg)

20.1.4 Dimonotransitive (**dimontr**)

Dimonotransitive verbs (**dimontr**) are complemented by an Indirect Object only. They include *show, ask, assure, grant, inform, promise, reassure, and tell.*

When I <i>asked</i> her, she burst into tears	V(dimontr,past)
I'll <i>tell</i> you tomorrow	V(dimontr,infin)
<i>Show</i> me	V(dimontr,imp)

20.1.5 Ditransitive (**ditr**)

Ditransitive verbs (**ditr**) are complemented by an Indirect Object and a Direct Object.

We <i>tell</i> each other everything	V(ditr,pres)
They <i>built</i> themselves a new theatre	V(ditr,past)
<i>Give</i> us some more drinks	V(ditr,imp)

20.1.6 Complex transitive (**cxtr**)

Complex transitive verbs (**cxtr**) are complemented by a Direct Object and an Object Complement.

He <i>knocked</i> himself unconscious.	V(cxtr,past)
It would <i>make</i> me sick	V(cxtr,infin)
I <i>take</i> that as a compliment.	V(cxtr,pres)
It <i>gets</i> you out of bed.	V(cxtr,pres)

20.1.7 Transitive (**trans**)

The transitivity of a verb may be unclear in some instances if it is complemented by a nonfinite clause:

I expect *John to do the hoovering*

Here, the noun phrase *John* may be the Subject of the nonfinite clause or the Object of the host clause. In all such cases we tag the main verb as **V(trans,...)**. Further examples include:

You <i>wanted</i> them to recognise your experience	V(trans,past)
I've <i>heard</i> him speak on Celtic spirituality	V(trans,edp)
I <i>saw</i> myself launching off into a philosophical treatise	V(trans,past)
I'll <i>have</i> it removed	V(trans,infin)
We <i>keep</i> each other fully informed	V(trans,pres)

NOTE: The **trans** label is not applied

(a) if the verb is *be*

One of my aims *is* to finish my PhD **V(cop,pres)**

All he did *was* repeat what someone else has told him **V(cop,past)**

(b) if the nonfinite clause does not have an overt Subject:

She *helped* design the course **V(montr,past)**

I *enjoy* doing it **V(montr,pres)**

(c) if the noun phrase is followed by a *wh*-clause whose main verb is a *to*-infinitive:

No one can *tell* me what to do **V(ditr,infin)**

The manual *shows* you how to install it **V(ditr,pres)**

20.2 Special Cases

In the tagging of verbs, the following special cases should be noted:

20.2.1 Passive Constructions

The tagging of the main verb in the passive is the same as it would be if the verb were active:

He was *caught* in the end **V(montr,edp)**
(cf. *They caught him in the end*)

He was *told* about it
(cf. *They told him about it*) **V(dimontr,edp)**

I may be *proved* wrong
(cf. *Someone may prove me wrong*) **V(cxtr,edp)**

Constructions that are tagged **V(trans,...)** are generally tagged the same in the passive:

They are not *allowed* to leave before six
(cf. *He does not allow them to leave before six*) **V(trans,edp)**

It is commonly *found* growing wild in Egypt
(cf. *You commonly find it growing wild in Egypt*) **(V(trans,edp))**

20.2.2 Prepositional and Phrasal-prepositional Verbs

Prepositional verbs, phrasal-prepositional verbs, and other verbs which have a collocational link with a preposition are not recognised as multi-word verbs. Therefore prepositional objects (objects introduced by a preposition) are not recognised.

Unless they have a Direct Object, these verbs are tagged as intransitives:

They *operated* on her leg **V(intr,past)**

Someone is *speaking* to you **V(intr,ingp)**

However, these verbs are tagged as monotransitives when they occur in the passive:

Her leg was *operated* on **V(montr,edp)**

You're being *spoken* to **V(montr,edp)**

The prepositions which collocate with these verbs are tagged **PREP(phras)**.

20.4.3 Nonfinite verbs in nonfinite clauses

The tagging of a nonfinite verb in a nonfinite clause is the same as it would be in the corresponding finite clause:

I use a piece of equipment *called* a spectrometer **V(cxtr,edp)**
(cf. *They call it a spectrometer*)

You end up *feeling* quite compromised **V(cop,ingp)**

opening quote	PUNC(oquo)
period (full stop)	PUNC(per)
question mark	PUNC(qm)
semicolon	PUNC(scol)
Other punctuation	PUNC(other)

- **NOTE:** **PUNC(obrack)** and **PUNC(cbrack)** cover all types of bracketing; for example: round brackets (also called parentheses), square brackets (also simply called brackets), diamond brackets (also called angle brackets), braces, and phonemic slants.
- **NOTE 2.** **PUNC(dash)** also covers the swung dash, i.e. ~.
- **NOTE 3.** ‘Other punctuation’, **PUNC(other)** covers any other symbols; for example, bullets, squares, triangles, stars.

22. PAUSES (PAUSE)

Pauses carry the main label **PAUSE**, followed by a feature label indicating the length of the pause - **short** or **long**. In the transcription of spoken texts, short pauses are marked as <,> and long pauses are marked as <,,>. So the tagging is

<,>	PAUSE(short)
<,,>	PAUSE(long)

23. MISCELLANEOUS TAGS

23.1 UNTAG

UNTAG is used to label incomplete words, marked <>...</>. These occur most commonly in spoken texts, though they may also be found in handwritten texts.

I’m <.> *ste* </.> I’m trying to decide **UNTAG**

UNTAG is also used to label complete words whose wordclass is indeterminate because of a “false start” or other discourse phenomenon:

I *have* I’d like to start by saying.... **UNTAG**

Here, *have* is indeterminate between lexical *have* and auxiliary *have*, so it is tagged **UNTAG**.

This tag does not carry any features.

23.2 Tag ?

Items are labelled with a question mark if they so unclear as to make it impossible to decide their wordclass:

This	PRON(dem,sing)
<unclear>	
word	?
</unclear>	

In other cases, words may be unclear, but it may still be possible to determine their wordclass. This is most commonly the case with personal names:

Dr	N(prop,sing):1/3
Alan	N(prop,sing):1/3
<unclear>	
Name	N(prop,sing):1/3
</unclear>	