Part-of-Speech Tagging Guidelines for the Penn Treebank Project (3rd Revision, 2nd printing)

Beatrice Santorini

June 1990 [1]

[1] Second printing (February 1995) updated and slightly reformatted by Robert MacIntyre. The text of this version appears to be the same as the first printing, but subtle differences may exist. The tags for proper noun and personal pronoun were altered in late 1992 in order to avoid conflicts with bracketing tags; this version reflects the new tag names.

Contents

1 Introduction

This section addresses the linguistic issues that arise in connection with annotating texts by part of speech ("tagging"). Section 2 is an alphabetical list of the parts of speech encoded in the annotation system of the Penn Treebank Project, along with their corresponding abbreviations ("tags") and some information concerning their definition. This section allows you to find an unfamiliar tag by looking up a familiar part of speech. Section 3 recapitulates the information in Section 2, but this time the information is alphabetically ordered by tags. This is the section to consult in order to find out what an unfamiliar tag means. Since the parts of speech are probably familiar to you from high school English, you should have little difficulty in assimilating the tags themselves. However, it is often quite difficult to decide which tag is appropriate in a particular context. The two sections 4.1 and 4.2 therefore include examples and guidelines on how to tag problematic cases. If you are uncertain about whether a given tag is correct or not, refer to these sections in order to ensure a consistently annotated text. Section 4.1 discusses parts of speech that are easily confused and gives guidelines on how to tag such cases, while Section 4.2 contains an alphabetical list of specific problematic words and collocations. Finally, Section 5 discusses some general tagging conventions. One general rule, however, is so important that we state it here. Many texts are not models of good prose, and some contain outright errors and slips of the pen. Do not be tempted to correct a tag to what it would be if the text were correct; rather, it is the incorrect word that should be tagged correctly.

- If you have questions that you do not find covered, be sure to let us know so that we can incorporate a discussion of them into updates of this guide.

2 List of parts of speech with corresponding tag

Adjective—JJ

Hyphenated compounds that are used as modifiers are tagged as adjectives (JJ).

EXAMPLES: happy-go-lucky/JJ
one-of-a-kind/JJ
run-of-the-mill/JJ

Ordinal numbers are tagged as adjectives (JJ), as are compounds of the form *n-th X-est*, like *fourth-largest*.

Adjective, comparative—JJR

Adjectives with the comparative ending *-er* and a comparative meaning are tagged JJR. *More* and *less* when used as adjectives, as in *more or less mail*, are also tagged as JJR. *More* and *less* can also be tagged as JJR when they occur by themselves; see the entries for these words in Section 4.2. Adjectives with a comparative meaning but without the comparative ending *-er*, like *superior*, should simply be tagged as JJ. Adjectives with the ending *-er* but without a strictly comparative meaning ("more X"), like *further* in *further details*, should also simply be tagged as JJ.

Adjective, superlative—JJS

Adjectives with the superlative ending *-est* (as well as *worst*) are tagged as JJS. *Most* and *least* when used as adjectives, as in *the most or the least mail*, are also tagged as JJS. *Most* and *least* can also be tagged as JJS when they occur by themselves; see the entries for these words in Section 4.2. Adjectives with a superlative meaning but without the superlative ending *-est*, like *first*, *last* or *unsurpassed*, should simply be tagged as JJ.

Adverb—RB

This category includes most words that end in *-ly* as well as degree words like *quite*, *too* and *very*, posthead modifiers like *enough* and *indeed* (as in *good enough*, *very well indeed*), and negative markers like *not*, *n't* and *never*.

Adverb, comparative—RBR

Adverbs with the comparative ending *-er* but without a strictly comparative meaning, like *later* in *We can always come by later*, should simply be tagged as RB.

Adverb, superlative—RBS

Article—DT (see "Determiner")

Cardinal number—CD

Common noun, plural—NNS (see "Noun, plural")

Common noun, singular or mass—NN (see "Noun, singular or mass")

Comparative adjective—JJR (see "Adjective, comparative")

Comparative adverb—RBR (see "Adverb, comparative")

Conjunction, coordinating—CC (see "Coordinating conjunction")

Conjunction, subordinating—IN (see "Preposition or subordinating conjunction")

Coordinating conjunction—CC

This category includes *and*, *but*, *nor*, *or*, *yet* (as in *Yet it's cheap*, *cheap yet good*), as well as the mathematical operators *plus*, *minus*, *less*, *times* (in the sense of "multiplied by") and *over* (in the sense of "divided by"), when they are spelled out.

For in the sense of "because" is a coordinating conjunction (CC) rather than a subordinating conjunction (IN).

EXAMPLE: He asked to be transferred, for/CC he was unhappy.

So in the sense of "so that," on the other hand, is a subordinating conjunction (IN).

Determiner—DT

This category includes the articles *a(n)*, *every*, *no* and *the*, the indefinite determiners *another*, *any* and *some*, *each*, *either* (as in *either way*), *neither* (as in *neither decision*), *that*, *these*, *this* and *those*, and instances of *all* and *both* when they do not precede a determiner or possessive pronoun (as in *all roads* or *both times*). (Instances of *all* or *both* that do precede a determiner or possessive pronoun are tagged as predeterminers (PDT).) Since any noun phrase can contain at most one determiner, the fact that *such* can occur together with a determiner (as in *the only such case*) means that it should be tagged as an adjective (JJ), unless it precedes a determiner, as in *such a good time*, in which case it is a predeterminer (PDT).

Exclamation—UH (see "Interjection")

Existential *there*—EX

Existential *there* is the unstressed *there* that triggers inversion of the inflected verb and the logical subject of a sentence.

EXAMPLES: There/EX was a party in progress.
There/EX ensued a melee.

Foreign word—FW

Use your judgment as to what is a foreign word. For me, *yoga* is an NN, while *bête noire* and *persona non grata* should be tagged *bête/FW noire/FW* and *persona/FW non/FW grata/FW*, respectively.

Gerund—VBG (see "Verb, gerund or present participle")

Interjection—UH

This category includes *my* (as in *My, what a gorgeous day*), *oh*, *please*, *see* (as in *See, it's like this*), *uh*, *well* and *yes*, among others.

List item marker—LS

This category includes letters and numerals when they are used to identify items in a list.

Modal verb—MD

This category includes all verbs that don't take an *-s* ending in the third person singular present: *can*, *could*, (*dare*), *may*, *might*, *must*, *ought*, *shall*, *should*, *will*, *would*.

Negation—RB (see "Adverb")

Noun, plural—NNS

Noun, singular or mass—NN

Numeral, cardinal—CD (see "Cardinal number")

Numeral, ordinal—JJ (see "Adjective")

Ordinal number—JJ (see "Adjective")

Participle, past—VBN (see "Verb, past participle")

Participle, present—VBG (see "Verb, gerund or present participle")

Particle—RP

This category includes a number of mostly monosyllabic words that also double as directional adverbs and prepositions. Consult the headings "IN or RB," "IN or RP" and "RB or RP" in Section 4.1 for further details.

Past participle—VBN (see "Verb, past participle")

Past tense verb—VBD (see "Verb, past tense")

Personal pronoun—PRP (see also "Possessive pronoun")

This category includes the personal pronouns proper, without regard for case distinctions (*I*, *me*, *you*, *he*, *him*, etc.), the reflexive pronouns ending in *-self* or *-selves*, and the nominal possessive pronouns *mine*, *yours*, *his*, *hers*, *ours* and *theirs*. The adjectival possessive forms *my*, *your*, *his*, *her*, *its*, *our* and *their*, on the other hand, are tagged PRP$.

Possessive ending—POS

The possessive ending on nouns ending in *'s* or *'* is split off by the tagging algorithm and tagged as if it were a separate word.

EXAMPLES:	John/NNP 's/POS idea
	the parents/NNS '/POS distress

Possessive pronoun—PRP$ (see also "Personal pronoun")

This category includes the adjectival possessive forms *my*, *your*, *his*, *her*, *its*, *one's*, *our* and *their*. The nominal possessive pronouns *mine*, *yours*, *his*, *hers*, *ours* and *theirs* are tagged as personal pronouns (PRP).

Possessive wh-pronoun—WP$

This category includes the wh-word *whose*.

Predeterminer—PDT

This category includes the following determinerlike elements when they precede an article or possessive pronoun.

EXAMPLES:	all/PDT his marbles	nary/PDT a soul
	both/PDT the girls	quite/PDT a mess
	half/PDT his time	rather/PDT a nuisance
	many/PDT a moon	such/PDT a good time

Preposition or subordinating conjunction—IN

We make no explicit distinction between prepositions and subordinating conjunctions. (The distinction is not lost, however—a preposition is an IN that precedes a noun phrase or a prepositional phrase, and a subordinate conjunction is an IN that precedes a clause.)

The preposition *to* has its own special tag TO.

Present participle—VBG (see "Verb, gerund or present participle")

Present tense verb—VBP or VBZ (see "Verb, present tense, other than 3rd person singular" and "Verb, present tense, 3rd person singular)"

Pronoun, personal—PRP (see "Personal pronoun")

Pronoun, possessive—PRP$ (see "Possessive pronoun")

Proper noun, plural—NNPS

Proper noun, singular—NNP

Subordinating conjunction—IN (see "Preposition or subordinating conjunction")

Superlative adjective—JJS (see "Adjective, superlative")

Superlative adverb—RBS (see "Adverb, superlative")

Symbol—SYM

This tag should be used for mathematical, scientific and technical symbols or expressions that aren't words of English. It should not used for any and all technical expressions. For instance, the names of chemicals, units of measurements (including abbreviations thereof) and the like should be tagged as nouns.

There, **existential**—EX (see "Existential *there*")

to—TO

To is tagged TO, regardless of whether it is a preposition or an infinitival marker.

Verb, base form—VB

This tag subsumes imperatives, infinitives and subjunctives.

EXAMPLES:	Imperative:	Do/VB it.
	Infinitive:	You should do/VB it.
		We want them to do/VB it.
		We made them do/VB it.
	Subjunctive:	We suggested that he do/VB it.

Verb, past tense—VBD

This category includes the conditional form of the verb *to be*.

EXAMPLES:	If I were/VBD rich, ...
	If I were/VBD to win the lottery, ...

Verb, gerund or present participle—VBG

Verb, past participle—VBN

Verb, present tense, other than 3rd person singular—VBP

Take care to correct VB to VBP where appropriate.

Verb, present tense, 3rd person singular—VBZ

Wh-determiner—WDT

This category includes *which*, as well as *that* when it is used as a relative pronoun.

Wh-pronoun—WP

This category includes *what*, *who* and *whom*.

Wh-pronoun, possessive—WP$ (see "Possessive wh-pronoun")

Wh-adverb—WRB

This category includes *how*, *where*, *why*, etc.

When in a temporal sense is tagged WRB. In the sense of "if," on the other hand, it is a subordinating conjunction (IN).

EXAMPLES: When/WRB he finally arrived, I was on my way out.
I like it when/IN you make dinner for me.

3 List of tags with corresponding part of speech

This section contains a list of tags in alphabetical order and the parts of speech corresponding to them.

1.	CC	Coordinating conjunction
2.	CD	Cardinal number
3.	DT	Determiner
4.	EX	Existential *there*
5.	FW	Foreign word
6.	IN	Preposition or subordinating conjunction
7.	JJ	Adjective
8.	JJR	Adjective, comparative
9.	JJS	Adjective, superlative
10.	LS	List item marker
11.	MD	Modal
12.	NN	Noun, singular or mass
13.	NNS	Noun, plural
14.	NNP	Proper noun, singular
15.	NNPS	Proper noun, plural
16.	PDT	Predeterminer
17.	POS	Possessive ending
18.	PRP	Personal pronoun
19.	PRP$	Possessive pronoun
20.	RB	Adverb
21.	RBR	Adverb, comparative
22.	RBS	Adverb, superlative
23.	RP	Particle
24.	SYM	Symbol
25.	TO	*to*
26.	UH	Interjection
27.	VB	Verb, base form
28.	VBD	Verb, past tense
29.	VBG	Verb, gerund or present participle

30.	VBN	Verb, past participle
31.	VBP	Verb, non-3rd person singular present
32.	VBZ	Verb, 3rd person singular present
33.	WDT	Wh-determiner
34.	WP	Wh-pronoun
35.	WP$	Possessive wh-pronoun
36.	WRB	Wh-adverb

4 Problematic cases

This section discusses difficult tagging decisions. Section 4.1 discusses parts of speech that are easily confused and guidelines on how to tag such cases. Section 4.2 contains an alphabetical list of specific problematic words and collocations.

4.1 Confusing parts of speech

This section discusses parts of speech that are easily confused and gives guidelines on how to tag such cases.

CC or DT

When they are the first members of the double conjunctions *both* ... *and*, *either* ... *or* and *neither* ... *nor*, *both*, *either* and *neither* are tagged as coordinating conjunctions (CC), not as determiners (DT).

EXAMPLES: Either/DT child could sing.

But:

Either/CC a boy could sing or/CC a girl could dance.
Either/CC a boy or/CC a girl could sing.
Either/CC a boy or/CC girl could sing.

Be aware that *either* or *neither* can sometimes function as determiners (DT) even in the presence of *or* or *nor*.

EXAMPLE: Either/DT boy or/CC girl could sing.

CD or JJ

Number-number combinations should be tagged as adjectives (JJ) if they have the same distribution as adjectives.

EXAMPLES: a 50–3/JJ victory (cf. a handy/JJ victory)

Hyphenated fractions *one-half*, *three-fourths*, *seven-eighths*, *one-and-a-half*, *seven-and-three-eighths* should be tagged as adjectives (JJ) when they are prenominal modifiers, but as adverbs (RB) if they could be replaced by *double* or *twice*.

EXAMPLES:	one-half/JJ cup;	cf. a full/JJ cup
	one-half/RB the amount;	cf. twice/RB the amount; double/RB the amount

CD or NN

Sometimes, it is unclear whether *one* is cardinal number or a noun. In general, it should be tagged as a cardinal number (CD) even when its sense is not clearly that of a numeral.

EXAMPLE: one/CD of the best reasons

But if it could be pluralized or modified by an adjective in a particular context, it is a common noun (NN).

EXAMPLE: the only (good) one/NN of its kind
(cf. the only (good) ones/NNS of their kind)

In the collocation *another one*, *one* should also be tagged as a common noun (NN).

Hyphenated fractions *one-half*, *three-fourths*, *seven-eighths*, *one-and-a-half*, *seven-and-three-eighths* should be tagged as adjectives (JJ) when they are prenominal modifiers, but as adverbs (RB) if they could be replaced by *double* or *twice*.

EXAMPLES:	one-half/JJ cup;	cf. a full/JJ cup
	one-half/RB the amount;	cf. twice/RB the amount; double/RB the amount

CD or RB

Number-number combinations should be tagged as adverbs (RB) if they have the same distribution as adverbs.

EXAMPLES: They won 50–3/RB. (cf. They won handily/RB.)

Hyphenated fractions *one-half*, *three-fourths*, *seven-eighths*, *one-and-a-half*, *seven-and-three-eighths* should be tagged as adjectives (JJ) when they are prenominal modifiers, but as adverbs (RB) if they could be replaced by *double* or *twice*.

EXAMPLES:	one-half/JJ cup;	cf. a full/JJ cup
	one-half/RB the amount;	cf. twice/RB the amount; double/RB the amount

DT or CC—see CC or DT

DT or NN

When determiners are used pronominally, i.e. without a head noun, they should still be tagged as determiners (DT)—not as common nouns (NN).

EXAMPLES: I can't stand this/DT.
I'll take both/DT.
Either/DT would be fine.

DT or PDT

When potential predeterminers precede an article or possessive pronoun, they are predeterminers (PDT). When they do not, they are determiners (DT).

EXAMPLES: all/DT girls; all/DT young girls
The girls all/DT left.
both/DT boys; both/DT little boys
The boys both/DT left.

all/PDT the girls; all/PDT the young girls
both/PDT the boys; both/PDT the little boys

EX or RB

Existential *there* is unstressed and triggers inversion of the inflected verb and the logical subject of a sentence.

EXAMPLES: There/EX was a party in progress.
There/EX ensued a melee.

By contrast, when *there* is used adverbially, it receives at least some stress and does not trigger inversion.

EXAMPLES: There/RB, a party was in progress.
There/RB, a melee ensued.

Existential and adverbial *there* can both occur together in the same sentence.

EXAMPLE: There/EX was a party in progress there/RB.

IN or RB

It is often difficult to distinguish prepositions and adverbs. In general, prepositions are associated with an immediately following noun phrase. However, they may be "stranded," i.e. their object may occur someplace other than immediately following the preposition. For instance, in the example below, the stranded preposition *without* is associated with *the credit card.*

EXAMPLE: the credit card you won't want to do without/IN

Prepositions may also immediately precede prepositional phrases. This means that one preposition can precede another (*to* counts as a regular preposition in this context), as in the following examples:

EXAMPLES: blaze out/IN into/IN space
come out/IN of/IN the woodwork
look up/IN to/TO someone
because/IN of/IN her late arrival
to plant on/IN into/IN spring

About and *around* when used to mean "approximately" should be tagged as adverbs (RB), not as prepositions (IN).

Close(r) and *near(er)* in collocation with *to* should be tagged as adverbs (RB), not as prepositions (IN).

If a putative preposition is not associated with an explicitly expressed object anywhere in the clause, it should be tagged as an adverb (RB)—or as a particle (RP) (see "RB or RP").

EXAMPLE: We'll just have to do without/RB.

IN or RP

Both prepositions and particles occur in collocation with verbs and are often difficult to distinguish from one another. It is important to realize that the idiomaticity of a collocation is not a foolproof criterion that a word is a particle. After briefly discussing the syntactic properties of prepositions, we give some diagnostic tests for the distinction between prepositions and particles.

As noted above ("IN or RB"), prepositions are generally associated with an immediately following noun phrase. However, they may be "stranded," i.e. their object may occur at the beginning of a clause rather than immediately following the preposition. For instance, in the examples below, the stranded prepositions *at* and *against* are associated with *the picture* and *what*, respectively.

EXAMPLES:	the picture (which/that) we will look at/IN next
	He doesn't know what he is up against/IN.

Prepositions may also immediately precede prepositional phrases. This means that one preposition can precede another (*to* counts as a regular preposition in this context), as in the examples below. To be tagged as IN rather than as RP, a putative preposition must be more closely associated with the following prepositional phrase than with the verb.

EXAMPLES:	blaze out/IN into/IN space
	come out/IN of/IN the woodwork
	look up/IN to/TO someone
	because/IN of/IN her late arrival
	take millions of dollars out/IN of/IN circulation
	(cf. *take out millions of dollars of circulation)

If a putative preposition is not associated with an object anywhere in the clause, it should be tagged either as a particle (RP)—or as an adverb (RB) (see "RB or RP").

A word is a particle (RP) rather than a preposition (IN):

- if it can either precede or follow a noun phrase object.

EXAMPLE:	She told off/RP her friends;
	she told her friends off/RP.

- if when you replace a noun phrase object by a pronoun, the pronoun must precede the word.

EXAMPLES:	She told them off/RP; *she told off/RP them.
	He peeled it off/RP; *he peeled off/RP it.

If the results of this test conflict with the results of the first test, go by the results of the second.

EXAMPLE:	???to run a bill up/RP; to run up/RP a bill;
	to run it up/RP; *to run up/RP it

- if it can be part of a noun that is derived from a particle-verb collocation.

EXAMPLES:	to break down/RP;	breakdown
	to break through/RP;	breakthrough
	to be left over/RP;	leftovers
	to push over/RP;	pushover
	to put down/RP;	putdown

The results of this test are one-directional only; if there is no related noun, the word can still be a particle.

EXAMPLES: to pass out/RP; *passout
to pull off/RP; *pulloff

- if it bears stress in clause-final position (this criterion only applies to monosyllabic words).

EXAMPLE: Why don't you come by/RP?
(vs. Real bargains are hard to come by/IN)

While particles usually occur in construction with verbs, they can occur together with parts of speech that are derived from verbs as well.

EXAMPLES: the cutting/NN off/RP of the top
the setting/NN up/RP of the problem
He looks worn/JJ out/RP.

A word is a preposition (IN) rather than a particle (RP):

- if it must precede a noun phrase object.

EXAMPLE: She stepped off/IN the train;
*she stepped the train off/IN.

- if when you replace a noun phrase object by a pronoun, the pronoun cannot precede the word.

EXAMPLE: She has been into/IN it for a year;
*she has been it into/IN for a year.

- if it cannot bear stress in clause-final position (this criterion only applies to monosyllabic words).

EXAMPLE: Real bargains are hard to come by/IN.
(vs. Why don't you come by/RP?)

IN or VBG, VBN

Putative prepositions ending in *-ed* or *-ing* should be tagged as past participles (VBN) or gerunds (VBG), respectively, not as prepositions (IN).

EXAMPLES: Granted/VBN that he is coming
Provided/VBN that he comes
According/VBG to reliable sources
Concerning/VBG your request of last week

IN or WDT

When *that* introduces complements of nouns, it is a subordinating conjunction (IN).

EXAMPLES: the fact that/IN you're here
the claim that/IN angels have wings

But when *that* introduces relative clauses, it is a wh-pronoun (WDT), on a par with *which*.

EXAMPLE: a man that/WDT I know

JJ or CD—see CD or JJ

JJ or JJR

Adjectives with the comparative ending *-er* and a comparative meaning are tagged JJR. *More* and *less* when used as adjectives, as in *more or less mail*, are also tagged as JJR. *More* and *less* can also be tagged as JJR when they occur by themselves; see the entries for these words in Section 4.2. Adjectives with a comparative meaning but without the comparative ending *-er*, like *superior*, should simply be tagged as JJ. Conversely, adjectives with the ending *-er* but without a strictly comparative meaning ("more X"), like *further* in *further details*, should also simply be tagged as JJ.

JJ or JJS

Adjectives with the superlative ending *-est* (as well as *worst*) are tagged as JJS. *Most* and *least* when used as adjectives, as in *the most or the least mail*, are also tagged as JJS. *Most* and *least* can also be tagged as JJS when they occur by themselves; see the entries for these words in Section 4.2. Adjectives with a superlative meaning but without the superlative ending *-est*, like *first*, *last* or *unsurpassed*, should simply be tagged as JJ.

JJR or JJ—see JJ or JJR

JJS or JJ—see JJ or JJS

JJ or NN

Nouns that are used as modifiers, whether in isolation or in sequences, should be tagged as nouns (NN, NNS) rather than as adjectives (JJ).

EXAMPLES: wool/NN sweater (vs. woollen/JJ sweater)
terminal/NN type (vs. terminal/JJ cancer)
life/NN insurance/NN company

Hyphenated modifiers, on the other hand, should always be tagged as adjectives (JJ). Thus, we have different part-of-speech assignments in examples like the following—depending on the orthographic conventions used:

EXAMPLES: income-tax/JJ return; income/NN tax/NN return
value-added/JJ tax; value/NN added/VBN tax

Prenominal modifiers that are gradable (i.e. they can be modified by a degree adverb or form a comparative or superlative) should be tagged as adjectives (JJ), not as nouns (NN).

EXAMPLES: a fun/JJ party
(cf. a really fun party, the most fun party I ever went to)

Color words should be tagged as nouns (NN, NNS) when they are used as names since they have the distribution of nouns—i.e. they can be modified by adjectives and they have an overt plural.

EXAMPLES: That's a nice red/NN.
Not too many reds/NNS go with that purple.

Also note the following contrast:

EXAMPLES: These plants are dark green/JJ.
These plants are a dark green/NN.

Generic adjectives should be tagged as adjectives (JJ) and not as plural common nouns (NNS), even when they trigger subject-verb agreement, if they can be modified by adverbs.

EXAMPLES: The (very) rich/JJ in this country pay far too few taxes.
The (multiply) handicapped/JJ

But if a putative adjective can't be modified by an adverb, it should be tagged as a common noun (NN).

EXAMPLES: Little good/NN will come of it.
(cf. *Very good will come of it.)

When used as prenominal modifiers, *top*, *side*, *bottom*, *front* and *back* should be tagged as adjectives (JJ).

JJ or NNP

Words that refer to languages or nations, like *English* or *French*, can be either adjectives (JJ) or proper nouns (NNP, NNPS).

EXAMPLES: English/JJ cuisine tends to be uninspired.
The English/NNPS tend to be uninspired cooks.

In prenominal position, such words are almost always adjectives (JJ). Do not be led to tag such words as proper nouns just because they occur in idiomatic collocations.

EXAMPLES: Chinese/JJ cabbage; Chinese/JJ cooking
Welsh/JJ rarebit; Welsh/JJ poetry

However, note:

EXAMPLE: an English/NNP sentence
(cf. a sentence of English/NNP)

The two parts of compounds such as *West German* or *North Korean* should be tagged identically—either as JJ or NNP.

EXAMPLE: the West/JJ German/JJ mark
He's a West/NNP German/NNP.

Hyphenated compound proper nouns acting as modifiers, such as *Gramm-Rudman* in *the Gramm-Rudman Act*, as well as compounds containing proper nouns as their second constituent, such as *mid-March* or *non-NATO*, should be tagged as proper nouns (NNP) rather than as adjectives (JJ).

JJ or RB

While most adverbs formed from adjectives end in *-ly*, not all do. The crucial criterion is whether a word modifies a noun, in which case it is an adjective (JJ), or a non-noun, in which case it is an adverb (RB).

EXAMPLES:	rapid/JJ growth/NN
	rapid/RB growing/VBG plants

Vexing cases arise in connection with compound adjectives that are spelled as two words, such as *mild flavored.* Tag both parts of such sequences as JJ—thus, *mild/JJ flavored/JJ.*

Take care not to tag predicate adjectives as adverbs. Thus, in *make life simple*, *simple* is an adjective; cf. the unacceptability of *make life simply.*

JJ or VBG

The distinction between adjectives (JJ) and gerunds/present participles (VBG) is often very difficult to make. There are a number of tests that you can use to decide. Be sure to apply these tests to the entire sentence containing the word that you are unsure of, not just the word in isolation, since the context is important in determining the part of speech of a word.

A word is an adjective (JJ):

- if it is gradable—that is, if it can be preceded by a degree adverb like *very*, or if it allows the formation of a comparative.

EXAMPLE:	Her talk was very interesting/JJ.
	Her talk was more interesting/JJ than theirs.

- if there is a corresponding *un-* form with the opposite meaning.

EXAMPLE:	an interesting/JJ conversation;
	an uninteresting/JJ conversation

- if it occurs in construction with *be*, and *be* could be replaced by *become*, *feel*, *look*, *remain*, *seem* or *sound.*

EXAMPLES:	The conversation became depressing/JJ.
	That place feels depressing/JJ.
	That place looks depressing/JJ.
	That place remains depressing/JJ.
	That place seems depressing/JJ.
	That place sounds depressing/JJ.

- if it precedes a noun, and the corresponding verb is intransitive, or does not have the same meaning.

EXAMPLES:	an appealing/JJ face;	*a face that appeals
	an appetizing/JJ dish;	*a dish that appetizes
	a revolving/JJ fund;	*a fund that revolves
	a winning/JJ smile;	*a smile that wins

But:

the existing/VBG safeguards;	safeguards that exist
a holding/VBG company;	a company that holds another one
a managing/VBG director;	a director who manages
the ruling/VBG class;	the class that rules

Note the following contrast:

a striking/JJ hat	*the hat will strike
the striking/VBG teachers	the teachers will strike

- if there is no corresponding verb.

EXAMPLE: a thoroughgoing/JJ investigation;
*thoroughgo

In connection with this point, note that striking meaning differences need not be reflected in terms of different parts of speech:

EXAMPLES: the outgoing/JJ president; an outgoing/JJ type of guy;
*outgo

an outstanding/JJ record; outstanding/JJ debts
*outstand

JJ or VBN

The distinction between adjectives (JJ) and past participles (VBN) is often very difficult to make. There are a number of tests that you can use to decide. Be sure to apply these tests to the entire sentence containing the word that you are unsure of, not just the word in isolation, since the context is important in determining the part of speech of a word.

A word is an adjective (JJ):

- if it is gradable—that is, if it can be preceded by a degree adverb like *very*, or if it allows the formation of a comparative.

EXAMPLE: He was very surprised/JJ.
He was more surprised/JJ than she was.

- if there is a corresponding *un-* form with the opposite meaning.

EXAMPLE: a hurried/JJ meeting;
an unhurried/JJ meeting

Be sure to check whether there is a corresponding verb beginning with *un-*. If there is, you cannot rely on this test to determine whether the word in question is an adjective or a participle, and you will have to use the other tests.

EXAMPLE: Your shoelace has been untied/JJ ever since we started.
I know—it got untied/VBN by accident.

When applying the *un-* test, be sure to take the entire context into account. For instance, *armed* can be either a JJ or a VBN, depending on its context.

EXAMPLE:	We need an armed/JJ guard. (cf. We need an unarmed guard.)
	Armed/VBN with only a knife, ... (cf. *Unarmed with only a knife, ...)

- if the word occurs in construction with *be*, and *be* could be replaced by *become*, *feel*, *look*, *remain*, *seem* or *sound*.

EXAMPLES:	He became interested/JJ. He felt interested/JJ. He looked surprised/JJ. He remained surprised/JJ. He seemed surprised/JJ. He sounded surprised/JJ.

However, if the complement of any of the verbs listed above is modified by a *by*-phrase, it should be tagged as a participle (VBN) rather than as an adjective (JJ).

EXAMPLE:	He remains guided/VBN by these principles.

- if the word occurs in construction with *keep*.

EXAMPLES:	They should be kept well watered/JJ.

- if it refers to a (resultant) state rather than to a (specific) event.

EXAMPLES:	At the time, I was married/JJ. I was mistaken/JJ (= wrong) the other day. a mistaken/JJ decision

- if a collocation of the form "X-ed N" cannot be paraphrased as "N that has been X-ed."

EXAMPLES:	a decided/JJ advantage;	*an advantage that has been decided
	a grown/JJ woman;	*a woman that has been grown
	married/JJ life;	*life that has been married
	worried/JJ faces;	*faces that have been worried

A word is a past participle (VBN):

- if it can be followed by a *by* phrase. If this criterion clashes with the possibility of inserting a degree adverb, tag the word as an adjective (JJ), not as a participle (VBN).

EXAMPLES:	He was invited/VBN by some friends of hers. He was very surprised/JJ by her remarks.

- if it refers to an (specific) event rather than to a (resultant) state.

EXAMPLES:	I was married/VBN on a Sunday.	
	I was mistaken/VBN for you the other day.	
	a case of mistaken/VBN identity	

- if the word occurs in construction with *be*, and *be* could be replaced by *get*, but not by *become*.

EXAMPLE: I was married/VBN on a Sunday.
(cf. I got married, *I became married)

JJR or NN

More and *less* should be tagged as comparative adjectives (JJR), even when they occur without a head noun, as in *more of the same*.

JJR or RBR

In collocations such as *Shares closed higher/lower*, *higher* and *lower* should be tagged as comparative adverbs (RBR). Cf. *Shares closed more reasonably/*reasonable today*; also *Shares closed up/RB by two points*.

JJS or NN

Most should be tagged as a superlative adjective (JJS) even when it occurs without a head noun, as in *most of the time*. The reason is that its distribution is parallel to that of other superlative adjectives; cf. the following:

EXAMPLES:	most apples;	most of the apples
		most of the bunch
	the ripest apples;	the ripest of the apples
		the ripest of the bunch

MD or VB

Forms of the auxiliary verbs *be*, *do* and *have* are tagged on a par with those of other verbs.

NN or CD—see CD or NN

NN or DT—see DT or NN

NN or JJ—see JJ or NN

NN or JJR—see JJR or NN

NN or NNS

Whether a noun is tagged singular or plural depends not on its semantic properties, but on whether it triggers singular or plural agreement on a verb. We illustrate this below for common nouns, but the same criterion also applies to proper nouns.

Any noun that triggers singular agreement on a verb should be tagged as singular, even if it ends in final *-s*.

EXAMPLE: Linguistics/NN is/*are a difficult field.

If a noun is semantically plural or collective, but triggers singular agreement, it should be tagged as singular.

EXAMPLES: The group/NN has/*have disbanded.
The jury/NN is/*are deliberating.

On the other hand, if a noun triggers plural agreement on a verb, it should be tagged as plural, even if it does not end in *-s*.

EXAMPLES: The faculty/NNS are on strike.
The police/NNS have arrived on the scene.

Some nouns, like *data*, trigger variable number agreement. Such nouns should be tagged according to their usage in a particular text. If the agreement pattern cannot be determined, tag such nouns as NN|NNS (see Section 5.2 for the vertical slash convention). Finally, note the following contrast:

Mechanics/NN is an established discipline.
The mechanics/NNS of the system are complex.

The only exception to the agreement rule concerns nouns denoting amounts, which trigger singular agreement even though formally plural. Such nouns should be tagged as plural noun (NNS).

EXAMPLES: Three years/NNS is a long time.
Twelve inches/NNS is a foot.

NN or NNP

Chapter, *Exhibit*, *Figure*, *Table* and the like should be tagged as common nouns (NN) even when capitalized as part of a reference. Abbreviations and initials should be tagged as if they were spelled out. Thus, *S&L* (which stands for *savings and loan*) should be tagged as a common noun (NN), not as a proper noun (NNP). By contrast, *the U.S.* should be tagged as *the U.S./NNP*.

Compounds containing proper nouns as their second constituent, such as *mid-March* or *non-NATO*, should be tagged as proper nouns (NNP).

NN or PRP

The indefinite pronouns *naught*, *none* and compounds of *any-*, *every-*, *no-* and *some-* with *-one* and *-thing* should be tagged as nouns (NN), not as pronouns (PRP). The sequence *no one* should be tagged *no/DT one/NN*; in its hyphenated form *no-one*, it should be tagged NN.

NN or RB

Nouns that are used adverbially should be tagged as nouns (NN, NNS, NNP or NNPS), not as adverbs (RB).

EXAMPLE: He comes by Sundays/NNPS and holidays/NNS.

Words denoting points of the compass are tagged as either nouns (NN) or adverbs (RB), depending on their syntactic properties. For instance, in *The nearest shopping center is two miles to the north of here*, *north* is preceded by an article and should be tagged as a common noun (NN). On the other hand, in the variant *The nearest shopping center is two miles north of here*, *north* is not preceded by an article and should be tagged as an adverb (RB).

The temporal expressions *yesterday*, *today* and *tomorrow* should be tagged as nouns (NN) rather than as adverbs (RB). Note that you can (marginally) pluralize them and that they allow a possessive form, both of which true adverbs do not.

The locative expression *home* when used by itself in an adverbial sense should be tagged as an adverb (RB).

EXAMPLES: Call me when you get home/RB.
Call me when you are at home/NN.

NN or VBG

It is often difficult to tell whether a form in *-ing* is a noun (NN) or a gerund (VBG), especially since both nouns and gerunds can be preceded by an article or a possessive pronoun.

- If a word in *-ing* allows a plural, it is a noun (NN).

  EXAMPLES: The reading/NN for this class is difficult.
  (vs. The readings/NNS for this class are difficult.)

- While both nouns and gerunds can be preceded by an article or a possessive pronoun, only a noun (NN) can be modified by an adjective, and only a gerund (VBG) can be modified by an adverb.

  EXAMPLES: Good/JJ cooking/NN is something to enjoy.
  Cooking/VBG well/RB is a useful skill.

- Similarly, if the direct object of the verb underlying the *-ing* form is expressed in an *of*-phrase, the *-ing* form is a noun (NN), but if it is expressed directly as a noun phrase, the *-ing* form is a gerund (VBG).

  EXAMPLES: GM's closing/NN of the plant
  (cf. GM's repeated/*repeatedly closing of the plant;
  GM's closings of the plant)

  GM's closing/VBG the plant
  (cf. GM's *repeated/repeatedly closing the plant;
  *GM's closings the plant)

- When an *-ing* form is preceded by a noun or a sequence of nouns, it is itself a noun. The resulting combination can in turn precede a noun.

  EXAMPLES: the plant/NN closing/NN
  unsavory plant/NN closing/NN tactics/NNS

- If a collocation *X-ing N* is equivalent (or similar) in meaning to *N X-es*, then the word is a gerund (VBG).

  EXAMPLES: the declining/VBG productivity of U.S. industry
  (cf. The productivity of U.S. industry is declining)

  the acting/VBG vice president
  (cf. a person who is acting as vice president)

- If a collocation *X-ing N* is not equivalent (or similar) in meaning to *N X-es*, then the word is a noun (NN). In such cases, the collocation can often be paraphrased in terms of an infinitive or a more clearly nominal construction.

  EXAMPLES: spending/NN reductions
  (reductions in spending, not: reductions that are spending)

  the mating/NN season
  (the season for mating, not: the season that is mating)

  a holding/NN pattern
  (a pattern of holding, not: a pattern that is holding)

- Finally, in many cases there is nothing to do but to use the vertical slash convention (see Section 5.2). For instance, since *anchoring devices* could be analyzed as either *devices that anchor* (gerund) or as *devices for anchoring* (noun), it should be tagged *anchoring/NN|VBG*.

NNS or NN—see NN or NNS

NNP or JJ—see JJ or NNP

NNP or NNPS—see NN or NNS

NNPS or NNP—see NN or NNS

PDT or DT—see DT or PDT

PRP or NN—see NN or PRP

PRP or PRP$

The nominal possessive pronouns *mine*, *yours*, *his*, *hers*, *ours* and *theirs* are tagged as personal pronouns (PRP), not as possessive pronouns (PRP$).

PRP$ or PRP—see PRP or PRP$

RB or CD—see CD or RB

RB or EX—see EX or RB

RB or IN—see IN or RB

RB or JJ—see JJ or RB

RB or NN—see NN or RB

RB or RBR

Adverbs with the comparative ending *-er* but without a comparative meaning should also simply be tagged as RB. For specific words, see Section 4.2.

RB or RBS

Most, though usually a superlative form, is a simple adverb (RB) in the collocation *most every-*.

RB or RP

Adverbs and particles can be even more difficult to distinguish than particles and prepositions. Again, it is important to realize that the idiomaticity of a particular collocation is not a diagnostic for the distinction. A word is an adverb (RB) rather than a particle (RP) if you can insert a manner adverb between the verb and the word.

EXAMPLE: to sit calmly/RB by/RB

Note that striking meaning differences are not always reflected in different part-of-speech assignment.

EXAMPLES: Bring the girls right up/RP. (= conduct)
Bring the girls right up/RP. (= educate)

Off in *badly off*, *better off*, *well off* and *worse off* is a particle (RP), not an adverb (RB).

In contexts concerning the movement of currency or commodity prices, *up* and *down* should be tagged as adverbs (RB), not as particles (RP).

RBR or RB—see RB or RBR

RBS or RB—see RB or RBS

RP or IN—see IN or RP

RP or RB—see RB or RP

VB or MD—see MD or VB

VB or VBP

If you are unsure whether a form is a subjunctive (VB) or a present tense verb (VBP), replace the subject by a third person pronoun. If the verb takes an *-s* ending, then the original form is a present tense verb (VBP); if not, it is a subjunctive (VB).

EXAMPLE: I recommended that you do/VB it.
(cf. I recommended that he do/*does it.)

VBG or IN—see IN or VBG, VBN

VBG or JJ—see JJ or VBG

VBG or NN—see NN or VBG

VBN or IN—see IN or VBG, VBN

VBN or JJ—see JJ or VBN

VBP or VB—see VB or VBP

WDT or IN—see IN or WDT

WDT or WP

If a wh-word precedes a head noun, it is a wh-determiner.

EXAMPLES: What/WDT kind do you want?
I don't know what/WDT kind you want.

Be sure to wash whatever/WDT fruit you buy.

Which/WDT book do you like better?
I don't know which/WDT book you like better.

Which/WDT one do you like better?
I don't know which/WDT one you like better.

What(ever) is only tagged as a wh-determiner (WDT) when it precedes a head noun; otherwise, it is tagged as a simple wh-word (WP).

EXAMPLES: Tell me what/WP you would like to eat.
I'll get you whatever/WP you want.

Which(ever), on the other hand, is tagged as a wh-determiner (WDT) even when it does not precede a head noun. This is parallel to the tagging of determiners as DT when they are used by themselves.

EXAMPLES: Which/WDT do you like better?
I don't know which/WDT you like better.
I'll get you whichever/WDT you want.

WP or WDT—see WDT or WP

4.2 Specific words and collocations

This section contains an alphabetical list of specific problematic words and collocations.

about when used to mean "approximately" should be tagged as an adverb (RB), rather than a preposition (IN).

all is usually a determiner (DT) or a predeterminer (PDT), but it can also be an adverb (RB).

EXAMPLES: all/RB around the Mediterranean
all/RB through the night

The collocation *at all* is tagged *at/IN all/DT*.

all but is tagged *all/RB but/RB*.

all right is tagged *all/RB right/JJ*.

another one is tagged *another/DT one/NN*.

any can be a determiner (DT).

EXAMPLES: We don't have any/DT.
Don't you want any/DT more/JJR?

However, when it precedes a comparative adverb, it is an adverb (RB).

EXAMPLES: I can't run any/RB further/RBR.
I can't go on like this any/RB more/RBR.

around when used to mean "approximately" should be tagged as an adverb (RB), rather than a preposition (IN).

EXAMPLES: The pound stabilized at around/RB 1.6973 dollars.
(cf. The pound stabilized at approximately 1.6973 dollars.)

But:

The pound stabilized around/IN 1.6973 dollars.
(cf. *The pound stabilized approximately 1.6973 dollars.)

as can be an IN.

EXAMPLES: It's just as/IN I thought.
As/IN an untenured faculty member, ...

But when it has a meaning akin to 'so', it is an adverb (RB).

EXAMPLES: I'm not as/RB hungry.
This one is not as/RB good.

Both types of *as* occur in the collocation *as ... as*.

EXAMPLES: as/RB hungry as/IN me
as/RB many as/IN he has

at all is tagged *at/IN all/DT*.

back should be tagged as an adjective (JJ) when used as a prenominal modifier, as in *back door*.

bottom should be tagged as an adjective (JJ) when used as a prenominal modifier, as in *bottom drawer*.

but, though usually a coordinating conjunction, is a preposition (IN) when it means "except."

EXAMPLE: everybody but/IN me

In very formal usage, it can be an adverb (RB) with the meaning "only."

EXAMPLE: We can but/RB try.

Chapter, as in *Chapter 1*, is an NN, not an NNP.

close(r) In the collocation *close(r) to*, *close(r)* is an adverb (RB) or comparative adverb (RBR).

EXAMPLE:	We were close/RB to home.
	We were closer/RBR to home.

coming should be tagged as an adjective (JJ) by analogy to *upcoming*.

data triggers variable number agreement and should be tagged according to its use in a particular text. If the agreement pattern cannot be determined, tag it as NN|NNS (see Section 5.2 for the vertical slash convention).

dear in *Oh dear* and *Dear me* is an exclamation (UH). In *Yes, dear*, on the other hand, it is used as a true vocative and should be tagged as a noun (NN). Finally, in salutations like *Dear Martin*, it should be tagged as an adjective (JJ).

down should be tagged as an adverb (RB), not as a particle (RP), in contexts concerning the movement of currency or commodity prices.

due in the collocation *due to* is a preposition (IN). Otherwise, it is an adjective (JJ).

EXAMPLES:	due/IN to/TO the storm
	The books are due/JJ on the due/JJ date.

each other should be tagged *each/DT other/JJ*.

far, though usually an adverb (RB), can also be an adjective (JJ).

EXAMPLE:	She lives far/RB away.
	She lives far/RB from/IN here.
	The far/JJ side of the moon

In the collocation *far(ther) from*, *far(ther)* is an adverb (RB) or comparative adverb (RBR).

EXAMPLE:	We were far/RB from home.
	We were farther/RBR from home.

Further is treated on a par with *farther*.

Figure, as in *Figure 1*, is an NN, not an NNP.

fit is an adjective (JJ) in the expression *to see fit*.

for in the sense of "because" is a coordinating conjunction (CC), not a subordinating conjunction (IN).

EXAMPLE: He asked to be transferred, for/CC he was unhappy.

front should be tagged as an adjective (JJ) when used as a prenominal modifier, as in *front door*.

further, like *farther*, is a comparative adverb (RBR) in the context *further from*.

half can be an adjective (JJ), a noun (NN) or a predeterminer (PDT). It is an adjective (JJ) if it immediately precedes a noun, a predeterminer (PDT) if it immediately precedes an article or a possessive pronoun, and a noun (NN) otherwise.

EXAMPLES: a half/JJ point
half/PDT his time; half/PDT the time
half/NN of the time

The hyphenated fraction *one-half* should be tagged as an adjective (JJ) when it is a prenominal modifier, but as an adverb (RB) if it could be replaced by *double* or *twice*.

EXAMPLES:	one-half/JJ cup;	cf. a full/JJ cup
	one-half/RB the amount;	cf. twice/RB the amount; double/RB the amount

hers, as in *That's hers*, is a PRP, not a PRP$.

his, as in *That's his*, is a PRP, not a PRP$.

however is usually a simple adverb (RB).

EXAMPLES: However/RB, that time has not yet come.
There seems to be a problem, however/RB.

It can also (rarely) be a wh-adverb (WRB).

EXAMPLES: However/WRB much he wants to, he can't.
I'll do it however/WRB he wants me to.

later should be tagged as a simple adverb (RB) rather than as a comparative adverb (RBR), unless its meaning is clearly comparative. A useful diagnostic is that the comparative *later* can be preceded by *even* or *still*.

EXAMPLES: I'll get it around to it sooner or later/RB.
If you don't hurry, we'll arrive (even) later/RBR than your mother.

less should be tagged as a comparative adjective (JJR) when it is used without a head noun and it corresponds to the object of a verb or preposition. It should be tagged as a comparative adverb (RBR) when its use is parallel to other adverbs.

EXAMPLES: You should eat less/JJR (in terms of quantity).
(cf. You should eat less/JJR cheese.)

You should eat less/RBR (in terms of frequency).
(cf. You should eat rarely/RB.)

You should work less/RBR.
(cf. You should work harder/RBR.)

Less should be tagged as a comparative adjective (JJR) even when it occurs without a head noun, as in *less of a problem*.

Less in the sense of *minus* should be tagged as a coordinating conjunction (CC).

little is an adjective (JJ) in *a little bit*, *a little bit more* and *a little more*.

lot is a noun (NN) in *a lot* and *a lot more*.

many is a PDT when it immediately precedes an article. In general, however, it is an adjective (JJ), since it can be preceded by an article or a personal pronoun.

EXAMPLES: many/PDT a/DT stormy night
the/DT many/JJ faces of God

maximum, as in *maximum tolerance*, is a noun (NN), not an adjective (JJ).

mine, as in *That's mine*, is a PRP, not a PRP$.

minimum, as in *minimum wage*, is a noun (NN), not an adjective (JJ).

more should be tagged as a comparative adjective (JJR) when it is used without a head noun and it corresponds to the object of a verb or preposition. It should be tagged as a comparative adverb (RBR) when its use is parallel to other adverbs.

EXAMPLES: You should eat more/JJR (in terms of quantity).
(cf. You should eat more/JJR fruit.)

You should eat more/RBR (in terms of frequency).
(cf. You should eat often/RB.)

You should relax more/RBR.
(cf. You should work harder/RBR.)

More should be tagged as a comparative adjective (JJR) even when it occurs without a head noun. Again, however, if it fills the same position as an adverb, it should be tagged as a comparative adverb (RBR).

EXAMPLES: more/JJR of my friends
grows to five feet or more/JJR
It's more/JJR of a vegetable garden.
cf. *It's almost/RB of a vegetable garden.

But:

It's more/RBR a vegetable garden.
It's almost/RB a vegetable garden.

most, should be tagged as a superlative adjective (JJS) even when it occurs without a head noun, as in *most of the time*. The reason is that its distribution is parallel to that of other superlative adjectives; cf. the following:

EXAMPLES:	most apples;	most of the apples
		most of the bunch
	the ripest apples;	the ripest of the apples
		the ripest of the bunch

In the collocation *most every-*, *most* is a simple adverb (RB).

much can be an adjective (JJ) or an adverb (RB).

EXAMPLES: He doesn't have much/JJ energy left.
She said nothing, or at least not very much/JJ.
That's much/RB better already.
I like that quite well; in fact, I like it very much/RB.

near can be an adjective (JJ), an adverb (RB) or a preposition (IN).

EXAMPLES: The near/JJ side of the moon
They had approached quite near/RB.
We were near/IN the station.

But:

We were very/RB near/RB the station.

In the collocation *near(er) to*, *near(er)* is an adverb (RB) or comparative adverb (RBR). The colloquial use of *nearer* without a following *to* should also be tagged as a comparative adverb (RBR).

EXAMPLES: We were near/RB to the station.
We were nearer/RBR to the station.
We were nearer/RBR the station.

next can be an adjective (JJ), an adverb (RB) or (in archaic usage) a preposition (IN).

EXAMPLE: The next/JJ train
They live next/RB to/TO the park.
I grasp the hands of those next/IN me.

no can be a determiner (DT).

EXAMPLE: We have no/DT solution to that difficulty as yet.

It can also be an adverb (RB).

EXAMPLE: She is no/RB longer her old self.

Finally, as the opposite of *yes*, it is an interjection (UH).

no one should be tagged *no/DT one/NN*; in its hyphenated form *no-one*, it should be tagged NN.

not is an adverb (RB), as is its contracted form **n't**.

off is a particle (RP) in the collocations *well off, better off, badly off, worse off.*

one In general, *one* should be tagged as a cardinal number (CD) even when its sense is not clearly that of a numeral.

EXAMPLE: one/CD of the best reasons

However, when it is used as an impersonal third person pronoun, it is a pronoun (PRP).

EXAMPLE: One/PRP doesn't do that kind of thing in public.

If it could be pluralized or modified by an adjective in a particular context, it is a common noun (NN).

EXAMPLE: the only (good) one/NN of its kind
(cf. the only (good) ones/NNS of their kind)

In the collocation *another one*, *one* is a common noun (NN).

only should be tagged as an adverb (RB), unless it can be paraphrased by *sole*.

EXAMPLE: the only/JJ good ones
(cf. the sole/*solely good ones)

only/RB the good ones
(cf. solely/*sole the good ones)

other If *other* could be pluralized in a particular context, it is a common noun (NN).

EXAMPLE: One of them is good, but the other/NN is bad.
(cf. the others/NNS are bad)

ours, as in *That's ours*, is a PRP, not a PRP$.

over in the collocation *It's (all) over* is an adverb (RB).

own in combination with possessive pronouns is an adjective (JJ).

EXAMPLES: a room of one's own/JJ
her own/JJ room

people is a plural noun (NNS), since it triggers plural agreement.

plenty is a common noun (NN), even in collocations like *plenty warm*.

public in the collocation *to go public* is an adjective (JJ).

rather in isolation is an adverb (RB).

EXAMPLE: Tareyton smokers would rather/RB fight than switch.

In the collocation *rather than*, however, it should be tagged as a subordinating conjunction.

EXAMPLE: But often it's wiser to switch rather/IN than/IN to fight.

right The collocation *all right* is tagged *all/RB right/JJ*.

Section, as in *Section 1*, is an NN, not an NNP.

see fit *Fit* is an adjective (JJ) in the expression *to see fit*.

side should be tagged as an adjective (JJ) when used as a prenominal modifier, as in *side door*.

so in the sense of "to such a degree" is an adverb (RB).

EXAMPLES: So/RB many pieces are broken that we need a new one.
He was so/RB irresponsible we fired him.

In the collocations *so as to* and *so that* and when it means *so that* by itself, *so* is a subordinating conjunction (IN).

EXAMPLES: He left the house quietly so/IN as/IN not to wake anyone.
I left early so/IN that/IN I would catch my train.
I gave him money so/IN he could buy it.

So in the sense of "therefore" is an adverb (RB) rather than a subordinating conjunction (IN).

EXAMPLE: He was unhappy, so/RB he asked to be transferred.

sooner If its meaning is not clearly comparative, *sooner* should be tagged as a simple adverb (RB) rather than as a comparative adverb (RBR). A useful diagnostic is that the comparative *sooner* can be preceded by *even*.

EXAMPLES: I'll get it around to it sooner/RB or later.
Let's hurry, so we can arrive (even) sooner/RBR than your mother.

such Since any noun phrase can contain only one determiner, the fact that it can occur together with a determiner (as in *the only such case*) means that it should generally be tagged as an adjective (JJ). However, when it precedes a determiner, it should be tagged as a predeterminer (PDT).

Table, as in *Table 1*, is an NN, not an NNP.

that when used to introduce relative clauses is a WDT, not an IN.

theirs, as in *That's theirs*, is a PRP, not a PRP$.

then can have a strictly temporal sense ("at that point in time") or a more general sense ("in that case"). In either case, it is an adverb (RB). Both uses can occur in the same sentence.

EXAMPLE: Then/RB I'll have to do it before then/RB.

Then can also be an adjective (JJ).

EXAMPLE: The then/JJ governor of Massachusetts

top should be tagged as an adjective (JJ) when used as a prenominal modifier, as in *top drawer* or *top notch*.

up should be tagged as an adverb (RB), not as a particle (RP), in contexts concerning the movement of currency or commodity prices.

very, though usually an adverb (RB), can be an adjective (JJ).

EXAMPLE: the very/JJ idea

vice, as in *vice president*, should be tagged as a common noun (NN) rather than as an adjective (JJ).

well is an adjective (JJ) when it means the opposite of *sick*. It is an adverb (RB) otherwise.

EXAMPLES: I'm quite well/JJ, thank you.
You did very well/RB on the exam.
I try to speak only well/RB of people.

when in a temporal sense is tagged WRB. In the sense of "if," on the other hand, it is a subordinating conjunction (IN).

EXAMPLES: When/WRB he finally arrived, I was on my way out.
I like it when/IN you make dinner for me.

worth is a preposition (IN) when it precedes a measure phrase, as in *worth ten dollars*.

yet can be a coordinating conjunction (CC).

EXAMPLE: It's expensive yet/CC worth it.

It can also be an adverb (RB).

EXAMPLES: I've found yet/RB another error.
We have no solution to that difficulty as yet/RB.

yours, as in *That's yours*, is a PRP, not a PRP$.

5 General tagging conventions

5.1 Part of speech and syntactic function

We adopt the general convention that parts of speech are defined on the basis of their syntactic distribution rather than their semantic function. This convention has several important consequences. One is that nouns in prenominal position that are being used as modifiers are tagged as nouns (NN), not as adjectives (JJ) (see Section 4.1—JJ or NN).

EXAMPLES: a cotton/NN shirt
the nearest book/NN store

Another is that nouns that are used as adverbial modifiers are tagged as nouns, not adverbs (see Section 4.1—NN or RB).

EXAMPLE: This week, I work mornings/NNS only.

5.2 Vertical slash convention

Linguistic or extralinguistic context generally resolves the question of what tag to assign to a token.

EXAMPLES: Sampling/VBG data can be time-consuming.
Sampling/NN data can be full of errors.

Nevertheless, uncertainties can arise. Rather than attempting to forcibly resolve such uncertainties, with the attendant risk of inconsistency, you should simply record them by separating the relevant tags by a vertical slash (this character appears as an interrupted slash on the keyboard). For instance, in the absence of disambiguating context, examples such as the following should be tagged using the vertical slash.

EXAMPLES: Sampling/NN|VBG data can be fun.
The Duchess was entertaining/JJ|VBG last night.
The Duchess was guarded/JJ|VBN last night.

5.3 Capitalized words

If a series of words is capitalized as part of a name, the capitalized words should be tagged as proper nouns (NNP or NNPS).

EXAMPLES: Constitution/NNP Avenue/NNP
the/DT Fulton/NNP County/NNP Grand/NNP Jury/NNP
Kansas/NNP City/NNP
Mount/NNP Everest/NNP
North/NNP Carolina/NNP
Supreme/NNP Court/NNP Justice/NNP
A/NNP Tale/NNP of/IN Two/NNP Cities/NNPS
the/DT United/NNP States/NNPS

Otherwise, titles should be tagged as if they were running text.

If a single word is capitalized because it is used as a title, it should be tagged as a proper noun (NNP). But if a single word or series of words is capitalized as a result of figurative speech, it should be tagged as if it weren't capitalized.

EXAMPLES: Mother/NNP, are you coming?
He was no good at rebelling against Society/NN.

5.4 Abbreviations

Abbreviations and initials should be tagged as if they were spelled out.

EXAMPLES: Mr./NNP John/NNP Warner/NNP
Dr./NNP Elizabeth/NNP Blackwell/NNP
Ave./NNP of/IN the/DT Americas/NNPS
e.g./FW