GCS - Grammatical Coding System Manual

Susan Curtiss, Department of Linguistics, UCLA,
Jeff MacSwan, Department of Language and Literacy, College of Education,
Arizona State University
Jeanette Schaeffer, Faculty of Medicine, Ben Gurion University of the Negev
Murat Kural, Department of Linguistics and Cognitive Sciences, University of California, Irvine
Tetsuya Sano, Meiji Gakuin University

* This research was supported by NIH grant DHHS NS28383. We gratefully acknowledge assistance from Jelena Krivapovic in refining the GCS system.

GCS CODING GUIDE

GCS Format

Contents:

1 Overview 4
2 Arguments 9
2.1 NPs 9
2.2 Pronouns 11
2.3 Expletives 12
2.4 Anaphors 13
2.5 Oblique arguments and PPs 13
3 Determiners 15
3.1 Articles 15
3.2 Plurals 15
3.3 Possessive forms 15
3.4 Demonstratives 16
3.5 Quantifiers 16
3.5.1 Basic Quantifiers *some*, *any*, *no*, *every* (*each*). 16
3.5.2 Cardinality Quantifiers *many*, *much*, *few*, *little*, and numerals. 17
4 Auxiliaries 18
4.1 Auxiliary types 18
4.2 Inflected and base form of modals 20
4.3 Quasi-modal auxiliary verbs 21
4.4 Auxiliary inversion 21
4.5 Cliticized forms 22
4.5.1 Subject-Aux Cliticization 22
4.5.2 AUX-Neg Cliticization 22
4.5.3 Verb-To Cliticization 21
5 Other I-related material 24
5.1 Tense 24
5.2 Agreement 26
5.3 Participial forms 27
5.4 Negative 27
5.5 Imperative 28
5.6 Tag questions…… …… 28
6 Predicates 30
6.1 Verbs 30

6.2 Verb-particles 30
6.3 Predicative NPs 30
6.4 Predicative APs 31
6.5 Predicative PPs 32
7 Adverbials 34
7.1 Adverbs 34
7.2 NP adverbials 34
7.3 PP adverbials 35
8 Conjunctions 36
8.1 Discourse conjunctions 36
8.2 Sentence-internal conjunctions 36
9 The C-System 37
9.1 Complementizers 37
9.1.1 Complementizers of Argument Clauses, *that*, *if*, *whether*, *for*: 37
9.1.2 Complementizers of Adjunct Clauses, *because*, *after*, *if*, *when*, etc. 37
9.2 Fronted Wh-phrases 38
9.3 Relative clauses 40
10 Extraction 42
10.1 P-stranding 42
10.2 Long Distance Extraction 42
10.3 Topicalization 43
10.4 Dislocation 43
11 Complementation 44
11.1 Infinitivals 44
11.2 Gerunds 45
11.3 Small Clauses 46
11.4 Participial modification 47
12 Deletion 49
12.1 Grammatical deletion 49
12.2 Ellliptical deletion 49
13 Errors 51
13.1 Placement/Deletion errors 51
13.2 Form errors 54
14 Comparatives 56
14.1 Prepositional form 56
14.2 Elliptical form 57
15 Codes 59

1 Overview

This manual has material regarding the formatting of:

- the %mor tier
- the %syn tier
- the %lex tier

The three tiers have the following general look:

*XXX: and he played with little children when they were fast asleep.
%mor: ONJ|and IPROS|he IT|play-ed P|with little D|child-p CADJ|when IPROS|they IAUX|be-s-d fast asleep
%syn: [ONJ] <ip SNP Vf PP (<cp [CADJ] <ip SNP AUXB ADVP PAP)
%lex: V|play ADJ|little N|child V|be ADV|fast ADJ|asleep

Interjections (*like*, *well*, etc.), frozen expressions (*I don't know*, *I mean*, *you know*, etc.), and idioms are not coded on any tier.

The function and vocabulary of these tiers are as follows.

%mor: Encodes information regarding the C, I, and D systems, as well as some additional predetermined cases of inflectional and derivational morphology. (All words and morphemes that are not excluded (i.e., interjections, frozen expressions, etc.) are written on the morphological tier, as it is the elements on this tier that will be calculated in the determination of a transcript's MLU.

Material that relates to the C, I, or D system receives the initial codes of C, I, or D, followed by a pipe, '|'.

C-System structures:

- Clause introducers, e.g., complementizers like *that* or *whether* for argument clauses, prepositions like *after* and *because,* and general clause introducers like *when* (when it indicates simultaneity or conditional), *if* (conditional), *while* (simultaneity or contrast), or *even though*, etc.
- Fronted Wh-phrases, e.g., *who*, *why*.
- Relative pronouns, e.g., *which*, *who*, *that*.
- Fronted auxiliaries.

I-system structures:

- Subject pronouns.(nominative case-marked; i.e., *I* but not *you*)
- Auxiliary material and inflected verbs.(including AGR and T)
- The infinitive marker *to*.

D-system structures:

- Articles and demonstratives
- Possessive forms, including pronouns, possessive determiners, names, and common nouns.
- Plural forms.
- Quantifiers.

Others (marked with a pipe, 'l'):

- Conjunctions (sentential)
- Particles.
- Prepositions.
- Object pronouns.

Material that does not relate to any of these systems and is not listed under "Others", receives no code or pipe, 'l'.

Cases of affixation are indicated with a hyphen, '-'.

- 3.sg agreement in the present tense is coded as -s, e.g., IA|walk**-s**, IA|do**-s**.
- Possessive form is coded as -'s, e.g., D|bill-**'s**.
- Ordinal numbers are coded as -th attached to the cardinal numbers, e.g., DCAR|four**-th**, DCAR|five**-th**, and DCAR|one**-th** (for *first*).
- Comparative forms are coded as -er, e.g., smart**-er**. This includes suppletive forms like *better* as good**-er**.
- Superlative forms are coded as -est, e.g., smart**-est**, good**-est**.
- Agentive forms are coded as -ag, e.g., open**-ag** for *opener*.
- Adverb-producing morphology is coded as it is, e.g., quick**-ly**.
- Adjective-producing morphology is coded as it is, e.g., risk**-y**.
- Noun-producing morphology is coded as it is, e.g., arrive**-al**.
- Verb-producing morphology is coded as it is, e.g., black**-en**.

In cases of compounds, the internal morphological structure is indicated with the use of a plus, '+', e.g., rail+road.

Cliticization, as in negative-aux or subject-aux contraction, is indicated with a tilde "~",

e.g., *I'm not*, coded as IPROS|I~IAUX|be-s, MEG|not or *wanna* as want~I|to, *gonna* and *gotta* as AUX|going~to and AUX|got~to. (Note: *wanna* and *hafta*. only get coded as AUX|*want~I*|*to* and AUX|have~I|to when there is overt evidence in the transcript of Agreement morphology.)

Regular and irregular morphology are distinguished as follows:

- Past: -ed for regular (*walked*, *jumped*), -d for irregular (*ran*, *kept*). This includes suppletive forms like *went* coded as IT|go**-d**.
- Participle: -en for regular (*walked*, *jumped*), -n for irregular (*taken*, *shown*).
- Plural: -pl for regular (*chickens*, *cats*), -p for irregular (*geese*, *oxen*). This includes suppletive forms like *people* as D|person**-p**.

Zero morphology is not coded: e.g., ***hit*** as the past of *hit*, ***run*** as the participial form of *run*, or ***deer*** as the plural of *deer*.

What is not coded in %mor (but coded in %lex):

- Underived adverbs
- Underived adjectives
- Bare nouns with no plural, derivational morphology, or genitive case marker.
- Anything that is not covered here, including hard to classify items such as *either*, *too*, *as*, etc.

%syn: Codes all information that relates to the position of a given constituent and partial information regarding constituent boundaries.

Every constituent is notated in one way or another. All constituents are given capitalized codes.

Only the distinctions that have any syntactic consequence are coded: For example, the auxiliary *be* (*I am leaving*) and the copula *be* (*I am fine*) are not distinguished because they behave the same way in terms of barring *do*-support: *Am I leaving?* and *Am I fine?* On the other hand, the auxiliary *have* (*I have left*) and the predicate *have* (*I have a problem*) are distinguished because they behave differently: *Have I left?* vs. *Do I have a problem?*

Independent sentences that happen to form part of the same utterance are coded on separate lines.

- Quotations:
  *XXX: He said "I am Mannekin the ghost".
  %syn: <ip SNP Vf DNP

%syn: <ip SNP AUXB PNP

- Parentheticals:
  *XXX: spirits, I've heard from scientists, and everything are possible
  %syn: <ip SNP AUXB PAP
  %syn: <ip SNP AUXH V-en PP

Constituent structure is coded by inserting <ip before an overt Nominative Case subject (SNP) or a clause with overt I-system elements in it, and a <cp before a fronted Wh-phrase and/or inverted auxiliary, or other overt C material ([CCOM], [CADJ] etc.). <ip is coded in both matrix and subordinate clauses, <cp is coded (in both matrix and subordinate cluases) if there is inversion or a fronted Wh-phrase or other overt C material (see section 9 for examples).

%lex: Lists all open-class items, i.e., those elements that a language can acquire without restricting the use of other related items, or modifying the structure in some way; i.e., those elements that are added easily to language (e.g., new content words like *e-mail* or DVD. *Be* is treated as a verb for purposes of the lexicsl tier. Lexicsl diversity and productivity can be indexed via the lexical tier.

Items/categories coded on the lexical tier include the following:

- Nouns (proper or common)
- Verbs
- Adjectives
- Adverbs

They exclude the following closed-class items:

- Pronouns
- Particles
- Complementizers and other clause introducing items
- Wh-phrases
- Determiners
- Auxiliary verbs
- Others

Compounds are treated as a single word, and a single lexical item.

Complex predicates, including verb-particle pairs, idiom chunks, and any group of items with a non-compositional meaning are listed as single combined entries.

*XXX: she is playing the odds in that race
%lex: **Vl|playtheodds** Nl|race

*XXX: I fell asleep
%lex: **V|fallasleep**

*XXX: he woke up .
%lex: **V|wakeup**

*XXX: he let the cat out of the bag
%lex: **V|letthecatoutofthebag**

If a complex item is composed of two (ore more) lexical items (compounds), the parts are separated with a plus, '+'.

*XXX: he was lying on the railroad
%lex: V|lie **N|rail+road**

*XXX: we handwashed the car
%lex: **V|hand+wash** N|car

Caution: The lexical tier contains only the base form of the word, minus any affixes, is put (e.g., for *boys*, N|boy). If the syntactic category of the word is changed by the affix, the derived category is reflectted on the lexical tier: (e.g., *repeatedly* an adverb derived from the verb *repeat* is coded as ADV|repeatedly).

*XXX: He repeatedly misplaced her toys
%mor: IPROS|he **repeat-ed-ly** IT|**mis-place-ed** DPOSD|her D|**toy-pl**
%syn: <ip SNP ADVP Vf DNP
%lex: **ADV|repeatedly V|misplace N|toy**

*XXX: This was the first time
%mor: DPRO|this IAUX|**be-s** DART|the DCAR|**one-th** time
%syn: <ip SNP AUXB PNP
%lex: **V|be ADJ|first** N|time

*XXX:People are strange
%mor: D|**person-p** IAUX|**be-s** strange
%syn: <ip SNP AUXB PAP
%lex: **N|people V|be** ADJ|strange

General note: When unsure regarding how to code an item on the lexical tier, the following presents some guidelines:

Nouns - Often not coded at %mor, they are usually some form of NP on %syn. Most common functions: argument, predicative NP or Adverbial Phrase (ADVP). For helpful information, try one of the following sections: Arguments (ch. 2), Predicative NPs (6.3), Adverbials (7.2), relative clauses (9.3) and participial modification (11.4)

Verbs – On %syn they can be many things: VP, Vf, Vn, etc; Helpful informtion may be provided in the sections on Auxiliaries (4), Other I-related material (5), Predicates (6) and Complementation (11).

Adverbs – Adverbs bear no codes on the morphological tier, but are coded as ADVPs on the syntactic tier (see section 7 for details). Note that nouns can function as adverbs and prepositions in comparatives (see section 14).

Adjectives – Adjectives are not coded on the morphological tier. If modifying the noun phrase, they are not coded separately on %syn. However, if they are used predicatively, the are PAPs on the syntactic tier. .Sections 2, 3.5 and 6.4 may prove helpful.

2 Arguments

At %mor:	I = infl	At %syn:	S = subject
	S = subject		D = direct object
	O = object		I = indirect object
	PRO = pronoun		O = oblique
	E = expletive		NP = NP
	ANAPH = anaphor		EXP = expletive NP
			P = preposition
			PP = PP

2.1 NPs

On the %mor tier, NPs are coded according to their morphological make-up: pronouns as PRO, anaphors as ANAPH, and full NPs as the collection of its parts, e.g., demonstratives, articles, etc., (see the determiner system in section 3).

Names are not coded on the %mor. They are entered as lexical units at %lex:

*XXX: tim was running away from mannekin
%mor: **tim** IAUX|be-s-d run-ing PRT|away P|from **mannekin**
%lex: **N|tim** V|runaway **N|mannekin**

*XXX: he remembered casper the friendly ghost
%mor: IPROS|he IT|remember-ed **casperthefriendlyghost**
%lex: V|remember **N|casperthefriendlyghost**

Common nouns that are not plurals or genitives are also not coded.

*XXX: the boy saw mannekin
%mor: DART|the **boy** IT|see-d **mannekin**
%lex: **N|boy** V|see **N|mannekin**

At %syn, all NP arguments are specified as NPs. Their grammatical function is indicated by prefixing S (subject), D (direct object), I (indirect object), and O (oblique) to the NP, i.e., SNP, DNP, INP, and ONP.

- An overt subject in a full clause marks the IP boundary, which is coded by placing an <ip before an SNP. There is no <ip before small clause subjects (see section 11.3). .

  *XXX: tim was kicking mannekin
  %mor: **tim** IAUX|be-s-d kick-ing **mannekin**
  %syn: <ip **SNP** AUXB V-ing **DNP**

*XXX: tim showed mannekin to jill
%mor: **tim** IT|show-ed **mannekin P|to jill**
%syn: <ip **SNP** Vf **DNP INP**

Caution - 1: Subject clauses like ***That he wanted to make children happy*** *is a good thing* are not SNPs.

Caution - 2: A *to*-phrase is an indirect object if it is obligatory, as is the case with *give* and *show*, but a plain PP if it is optional, as in *I walked* ***to the park*** or *I talked* ***to my brother***.

Caution –3: Predicative PPs are arguments, as in: *He is* ***at home*** coded as PPP (cf. 6.5)

- Dative-shifted INPs are still marked as INPs, though they appear at a different spot.

*XXX: I gave bill some candy
%mor: IPROS|I IT|give-d **bill** DQUA|some candy
%syn: <ip SNP Vf **INP** DNP

Caution: Heavy NP shift has the same INP - DNP order as the dative-shifted examples, but the P is preserved on the %mor tier. The two forms can be differentiated by stating heavy-NP shift cases in the %com line.

*XXX: I gave bill the candy that I brought to school
%mor: IPROS|I IT|give-d **P|to bill** DART|the candy CREL|that IPROS|I IT|bring-d P|to school
%syn: <ip SNP Vf **INP** DNP (<cp [DREL] <ip SNP Vf PP)

- Oblique NPs are reserved for instances where the NP complement of a preposition is found in isolation without the preposition. These typically involve cases where the preposition is either deleted or stranded, see sections 13 and 10.1, respectively.

*XXX: tim is scared mannekin
%mor: tim IAUX|be-s scare-en P|=0of **mannekin**
%syn: <ip SNP AUXB V-en =0P **ONP**

*XXX: who did bill talk to?
%mor: **CWH|who** CAUX|do-d bill talk **P|to**
%syn: <cp **OWH** AUXD <ip SNP V **P**

- NPs that have APs/Adjectives or PPs in them are only marked as XNPs on the %syn tier. The %lex tier contains all the lexical elements as independent entries.

*XXX: I saw a beautiful picture
%mor : IPROS|I IT|see-d **DART|a beautiful picture**
%syn: <ip SNP Vf **DNP**
%lex: V|see **ADJ|beautiful N|picture**

*XXX: I saw the picture of a beautiful house with brown shutters
%mor: IPROS|I IT|see-d **DART|the picture P|of DART|a beautiful house P|with brown D|shut-ag-pl**
%syn: <ip SNP Vf **DNP**
%lex: V|see **N|picture ADJ|beautiful N|house ADJ|brown N|shutter**

Caution 1: predicative NPs (cf. section 6.3):

*XXX: they are students
%mor: IPROS|they **IAUX|be-s D|student-pl**
%syn: <ip SNP **AUXB PNP**
%lex: **V|be N|student**

Caution 2: modifiers ([MDF]) (cf. section 11.4)

*XXX: he is mannekin the ghost
%mor: IPROS|he **IAUX|be-s mannekin DART|the ghost**
%syn: <ip SNP **AUXB PNP ([PMDF] NP)**
%lex: **V|be N|mannekintheghost**

2.2 Pronouns

On the %mor tier, pronouns are marked with PRO.

*XXX: you ate it
%mor: **PRO|you** IT|eat-d **PRO|it**

The presence of a distinct subject form (Nominative Case) of a pronoun is morphological evidence of a case-assigning functional category in the I-system In cases of pronouns that take distinct forms as subjects and objects (the Nominative and Accusative cases), i.e, *I - me*, *he - him*, *she - her*, *we - us*, *they - th*, an I is added to PROS in the sibject forms, and an O in the object forms: IPROS|I, PROO|me, IPROS|he, PROO|him, IPROS|she, PROO|her, etc.

*XXX: he ate them
%mor: **IPROS|he** IT|eat-d **PROO|them**

*XXX: she talked to me

%mor: **IPROS|she** IT|talk-ed P|to **PROO|me**

- Apart from *you* and *it*, there are other pronouns that also lack distinct forms related to grammatical function: *one*, *this*, and *that*, which are indicated as bare PROs.

*XXX: I ate one
%mor: IPROS|I IT|eat-d **PRO|one**

*XXX: this should stay on the shelf
%mor: **DPRO|this** IAUX|should stay P|on DART|the shelf

On the %syn tier, pronouns are coded according to their grammatical functions:

*XXX: she gave me that thing
%mor: **IPROS|she** IT|give-d **PROO|me** DDEM|that thing
%syn: <ip **SNP** Vf **INP** DNP

*XXX: she gave those to me
%mor: **IPROS|she** IT|give-d **DPRO|that-p** P|to **PROO|me**
%syn: <ip **SNP** Vf **DNP INP**

2.3 Expletives

Expletives are of two types:

- Existential *there*, ***There*** *are too many people in this room*
- The *it* that holds the (subject) place in a clause, ***It*** *was in this room that he was sleeping*

The weather *it* is treated as a full subject NP, (i.e., as an SNP at %syn).

Expletives are coded as PROE|xxx at %mor and as EXP at %syn. The postverbal subject in the *there* construction is coded as a PNP at %syn.

Expletives also define the IP boundary, and so they have an <ip before them.

*XXX: there are many people
%mor: **PROE|there** IAUX|be-s DCAR|many D|person-p
%syn: <ip **EXP** AUXB PNP

2.4 Anaphors

Anaphors are coded as ANAPH|xxx on the %mor tier, and by their grammatical function on the %syn tier, e.g., SNP, DNP, INP, etc.:

*XXX: the men pushed each other
%mor: DART|the D|man-p IT|push-ed **ANAPH|leachother**
%syn: <ip SNP Vf **DNP**

*XXX: you talked about yourself
%mor: PROS|you IT|talk-ed P|about **ANAPH|your+self**
%syn: <ip SNP Vf PP

2.5 Oblique arguments and PPs

All optional and oblique arguments are coded as PPs on the %syn tier. On the %mor tier, PPs are broken into P|xxx and whatever complements they have

*XXX: I brought this book for bill
%mor: IPROS|I IT|bring-d DDEM|this book **P|for bill**
%syn: <ip SNP Vf DNP **PP**

*XXX: I brought this book for him
%mor: IPROS|I IT|bring-d DDEM|this book **P|for PROO|him**
%syn: <ip SNP Vf DNP **PP**

As mentioned above, cases of dative-shift are handled on the %syn tier by coding the shifted argument as an INP. The same holds in the case of shifted beneficiaries.

*XXX: I brought bill this book
%mor: IPROS|I IT|bring-d **bill** DDEM|this book
%syn: <ip SNP Vf **INP** DNP

*XXX: I brought her this book
%mor : IPROS|I IT|bring-d **PROO|her** DDEM|this book
%syn: <ip SNP Vf **INP** DNP

In cases of preposition deletion, the remaining NP is marked as an ONP (for the error and deletion in this example, see sections 12 and 13).

*XXX: I am talking mannekin
%mor: IPROS|I IAUX|be-s talk-ing **P|=0about mannekin**
%syn: <ip SNP AUXB V-ing **=0P ONP**

Another case for the O prefix is P-stranding, discussed in section 10.2.

*XXX: what are you talking about?
%mor: **CWH|what** CAUX|be-s PROS|you talk-ing **P|about**

%syn: <cp **OWH** AUXB <ip SNP V-ing **P**

3 Determiners

At %mor: D = determiner
ART = article
POS = possessive
DEM = demonstrative
QUA = basic quantifiers
CAR = cardinality quantifiers

3.1 Articles

Articles, *a* and *the*, are coded as ART|xxx prefixed with a D for the D-system.

*XXX: The boy saw a plane
%mor: **DART|the** boy IT|see-d **DART|a** plane

3.2 Plurals

The plural forms are marked with a D|xxx and morphologically broken down on the %mor tier. Regular plurals are marked with -pl and irregulars with -p.

*XXX: The children saw planes
%mor: DART|the **D|child-p** IT|see-d **D|plane-pl**

Plural pronouns like *these* and *those* are PRO|xxx prefixed with a D, and the morphology broken down.

*XXX: I like those
%mor: IPROS|I like **DPRO|that-p**

3.3 Possessive forms

The possessive pronouns *my*, *your*, *her*, *his*, *her*, *its*, *our*¸ and *their*, are coded as POS with a D attached to the front. Since they also function as determiners, they also have a D (for 'determiner') at the back: **DPOSD**|xxx.

*XXX: my car is chasing your truck
%mor: **DPOSD|my** car IAUX|be-s chase-ing **DPOSD|your** truck

The possessive forms of names and common nouns are marked with an initial D, and the morphology is broken down as -'s.

*XXX: my mother's car is chasing bill's truck
%mor: DPOSD|my D|mother-'s car IAUX|be-s chase-ing D|bill-'s truck

Combinations of plural and possessive forms are coded by combining the parts.

*XXX: the boys' mothers are talking
%mor: DART|the **D|boy-pl-'s** D|mother-pl IAUX|be-s talk-ing

Possessive forms that do not function as determiners, i.e., *mine*, *yours*, *his*, *hers*, *ours*, *yours*, *theirs*, but rather function as pronouns (DPs), do not have the D after their code of DPRO: **DPRO|mine**, **DPRO|hers**, etc.

3.4 Demonstratives

Demonstratives like *this*, *that,* and *those* are coded as DEM|xxx with a D preceding, i.e., as DDEM|xxx. The demonstratives *these* and *those* are marked as the plural forms of *this* and *that*.

*XXX: these books are on that table
%mor: **DDEM|this-p** D|book-pl IAUX|be-s P|on **DDEM|that** table

3.5 Quantifiers

There are two types of quantifiers: the basic ones: *some*, *any*, *no* and *every*, and the cardinality (count or mass) quantifiers: *many*, *much*, *few*, *little*, and numbers.

3.5.1 Basic Quantifiers *some*, *any*, *no*, *every* (*each*).

These quantifiers are derived from the three basic operators: ∃ (existential), ∀ (universal), and ¬ (negation); *some* (∃), *every* (∀), *any* (¬∃), *no* (∀¬).

These quantifiers cannot co-occur with other determiners: **a some book*, **that every student*, etc.

Basic quantifiers are coded as QUA|xxx on the %mor, prefixed with D, i.e., DQUA|xxx.

*XXX: I saw some students
%mor: IPROS|I IT|see-d **DQUA|some** D|student-pl

*XXX: he plays with every toy
%mor: IPROS|he IA|play-s P|with **DQUA|every** toy

Compound forms like *someone*, *everything*, or *nowhere* are also marked as DQUA| with the morphology broken down. Basic quantifiers are not coded at %lex, unless they are compound forms (see below). In cases where a quantifier

combines with *-one*, *-one* is coded as a PRO|xxx.

*XXX: I saw someone
%mor: IPROS|I IT|see-d **DQUA|some+PRO|one**
%lex: V|see **N|some+one**

*XXX: he went nowhere
%mor: IPROS|he IT|go-d P|0to **DQUA|no+where**
%lex: V|go **ADJ|no+where**

3.5.2 Cardinality Quantifiers *all*, *many*, *much*, *few*, *little*, and numerals.

These quantifiers refer to some specified or unspecified quantity. With count nouns, it is the cardinality -- *many*, *few*, *several*, numbers; with mass nouns, it is amount -- *much*, *little*, *lot*, *plenty*.

These quantifiers can occur with other determiners: *those many students*, *a little pie*, *the five books*.

Cardinality quantifiers are coded as CAR|xxx on the %mor, prefixed with D, i.e., DCAR|xxx.

*XXX: she read a few books
%mor: IPROS|she read DART|a **DCAR|few** D|book-pl

*XXX: a lot of water poured down on mannekin
%mor: DART|a **DCAR|lot** P|of water IT|pour-ed PRT|down P|on mannekin

Numbers are DCAR|xxx on %mor, but ADJ|xxx on %lex.

*XXX: There are five people in those two rooms
%mor: PROE|there IAUX|be-s **DCAR|five** D|person-p P|in DDEM|that-p **DCAR|two** D|room-pl
%lex: **ADJ|five** N|people **ADJ|two** N|room

4 Auxiliaries

At %mor:	I = infl	At %syn:	AUX = auxiliary
	AUX = auxiliary		B = be
			H = have
			D = do
			M = modal
			G = get

4.1 Auxiliary types

All auxiliary categories are indicated as AUX on both the %mor and %syn tiers.

On %mor, an auxiliary typically gets an I prefix, i.e., IAUX|xxx.

- If an auxiliary is the only auxiliary in the structure, it appars as an IAUX|xxx. The only exception is the use of *got* without an accompanying auxiliary *have*:

  *XXX: I gotta go now
  %mor: IPROS|I **AUX|got~to** go now

- As many dialects of English now accept *got* as a verb meaning *have*, *got* so used with 1st or 2nd person is treated as a verb on all three tiers, unless the transcript makes clear that the speaker differentiates *got* from *have* as a main verb and uses *get* only as an auxiliary:

  *XXX: I got a new dress
  %mor: IPROS|I **got** DART|a new dress
  %lex: **V|get** Adj|new N|dress

- If there are two or more auxiliaries, only the first one gets an I prefix:

  *XXX: he has been sleeping for six hours
  %mor: IPROS|he **IAUX|have-s AUX|be-en** sleep-ing P|for DCAR|six D|hour-pl

NB: Although they do not show any overt morphology, modals automatically receive the I prefix because they are always finite: They are excluded from all non-finite contexts, such as infinitives (**He appears to can read* vs. *He appears to be able to read*) and bare VP complements with causatives (**She made me must buy a new rug* vs. *She made me have to buy a new rug*).

On the %syn tier, an auxiliary is coded as an AUX, suffixed with B (*be*), H (*have*), D (*do*), M (*modal*), or G (*get*), i.e., AUXB, AUXH, AUXD, AUXM, AUXG. All others, such as *going to*, *have to*, *used to*, etc. are simply AUX.

The auxiliary verbs themselves are analyzed on the %mor tier. The base form of the auxiliary in question is suffixed with an -s and/or a -d in the following conditions:

- -s if it displays agreement morphology, as in *is*, *am*, *are*, *has*, and *does*.
- -d if it displays (past) tense morphology, as in *had*, *did*, and *got*. See the next section for inflection on modals.
- -s-d if it displays both, i.e., in the cases of *was* and *were*.

*XXX: I am leaving
%mor: IPROS|I **IAUX|be-s** leave-ing
%syn: <ip SNP **AUXB** V-ing

*XXX: he has been talking for a long time
%mor: IPROS|he **IAUX|have-s AUX|be-n** talk-ing P|for DART|a long time
%syn: <ip SNP **AUXH AUXB** V-ing PP

*XXX: I am gonna see him next week
%mor: IPROS|I **IAUX|be-s AUX|go-ing~to** see PROO|him next week
%syn: <ip SNP **AUXB AUX** V DNP ADVP

Note 1: Neither the %mor nor the %syn needs to distinguish between the auxiliary and predicative forms of *be*. They can all be coded as AUX, since there is no COP for copula that is recognized as a distinct category in this context. This is not a crucial point because both types of *be*s behave more or less the same way, e.g., *She is not an employee* and *She is not working here*, and Is *she an employee*? and *Is she working here*? However, the difference is syntactically and morphologically meaningful with the auxiliary vs. main verbs *have* and *do*: *She does not have a purse* vs. *She hasn't bought a purse*, and *Does she have a purse* vs. *Has she bought a purse*? So main verb *have* and *do* are marked as Vs.

Note 2: Copula (main verb) *be* is indicated on the %lex tier, but not the auxiliary *be*.

*XXX: she is the teacher
%mor: IPROS|she **IAUX|be-s** DART|the teach-ag
%syn: <ip SNP **AUXB** PNP
%lex: **V|be** N|teacher

*XXX: this is a new toy
%mor: DPRO|this **IAUX|be-s** DART|a new toy
%syn: <ip SNP **AUXB** PNP
%lex: **V|be** ADJ|new N|toy

*XXX: he is playing with his new toy
%mor: IPROS|he **IAUX|be-s** play-ing P|with DPOSD|his new toy

%syn: <ip SNP **AUXB** V-ing PP
%lex: V|play ADJ|new N|toy

- Auxiliary *have* only carries tense and cooccurs with another verb: *I have read that book*; but main verb *have* indicates possession, association, etc, and has no cooccurring verb: *I have that book* (Note: *have* as a main verb does cooccur with verbs in causatives, *I had him read that book*, which are coded as small clauses (cf. 11.3).

*XXX: I had read that book
%mor: IPROS|I **IAUX|have-d** read-n DDEM|that book
%syn: <ip SNP **AUXH** V-n DNP

*XXX: I had that book
%mor: IPROS|I **IT|have-d** DDEM|that book
%syn: <ip SNP **Vf** DNP

- Auxiliary *do* occurs only in the context of questions, negation, or focus on the main verb.
- The verb *get* is an auxiliary with passives and in secondary predicate contexts, as in *I **got** hungry, I am **getting** angry, He **got** hurt*. Auxiliary *get* means *become;* main verb *get* indicates reception: *I **got** a new book*. *Get* is coded as a main verb on the lexical tier if followed by a Predicative Adj.

*XXX: I got hit
%mor: IPROS|I **IAUXG|get-d** hit
%syn: <ip SNP **AUXG** V

*XXX: I got hungry
%mor: IPROS|I **IT|get-d** hungry
%syn: <ip SNP **Vf** PAP
%lex: V|get Adj|hungry

*XXX: I am getting hungry
%mor: IPROS|I IAUX|be-s- **get-ing** hungry
%syn: <ip SNP AUXB **V-ing** AP

4.2 Inflected and base form of modals

The forms *would* and *could* should be treated as inflected modals in only two contexts: (a) When the modal truly refers to past time: *I couldn't go to school yesterday*, *Back in those days, I would go fishing with my grandpa every summer*.
(b) In sequence of tense contexts, where the modal is in a clause that is subordinated under a clause that contains past tense: *I thought you would leave*, *I knew you could do*

that.

Inflected modals are coded as IAUX|will-d and IAUX|can-d at %mor.

All other cases of *would* and *could*, e.g., conditionals, should be treated as simplex modals, and coded as: IAUX|would and IAUX|could on %mor.

Inflection makes no difference at %syn for modals. They are always coded as AUXM.

4.3 Quasi-modal auxiliary verbs

Quasi-modal verbs like *hafta*, *gonna*, *be able to*, etc are simply coded as AUX or IAUX on the %mor level, depending on whether the auxiliary is inflected or not. The infinitive *to* cliticizes on the auxiliary, and it is coded simply as AUX at %syn, and as a cliticized affix, at %mor, except with inflected *has to*.

*XXX: I hafta leave
%mor: IPROS|I **AUX|have~to** leave
%syn: <ip SNP **AUX** V

*XXX: she has to leave
%mor: IPROS|she **IAUX|have-s~I|to** leave
%syn: <ip SNP **AUX** ([TO] V)

*XXX: she is gonna leave
%mor: IPROS|she IAUX|be-s **AUX|go-ing~to** leave
%syn: <ip SNP AUXB **AUX** V

4.4 Auxiliary inversion

An inverted auxiliary is coded as a CAUX|xxx at %mor because it is in C, and at %syn, as a plain AUXB, AUXH, AUXD, or AUXM, whatever the case may be.

The inverted auxiliary (or the fronted wh-phrase, if there is one) receives a <cp before it, since it marks a CP boundary.

*XXX: did you leave early?
%mor: **CAUX|do-d** PROS|you leave early
%syn: **<cp AUXD** <ip SNP V ADVP

*XXX: where did you go?
%mor: CWH|where **CAUX|do-d** PROS|you go
%syn: **<cp** AWH **AUXD** <ip SNP V

4.5 Cliticized forms

The two most common types of cliticized forms are: subject~aux and aux~neg. There is also V~to cliticizations in cases like *wanna*, *gotta*, *gonna*, and *hafta*. In all these cases, the parts are put together with a "~" at %mor. In cases of *has- to* cliticization (inflected *hafta),* agreement and *to* are both coded in the usual way at %mor (see section 5 below). The cliticization is ignored at %syn.

4.5.1 Subject-Aux Cliticization

*XXX: he's leaving now
%mor: **IPROS|he~IAUX|be-s** leave-ing now
%syn: <ip **SNP AUXB** V-ing ADVP

*XXX: she'll leave soon
%mor: **IPROS|she~IAUX|will** leave soon
%syn: <ip **SNP AUXM** V ADVP

4.5.2 AUX-Neg Cliticization

*XXX: he isn't leaving now
%mor: IPROS|he **IAUX|be-s~NEG|not** leave-ing now
%syn: <ip SNP **AUXB NEG** V-ing ADVP

*XXX: she won't leave soon
%mor: IPROS|she **IAUX|will~NEG|not** leave soon
%syn: <ip SNP **AUXM NEG** V ADVP

4.5.3 Verb-To Cliticization

*XXX: I wanna play with this one
%mor: IPROS|I **want~I|to** play P|with DDEM|this PRO|one
%syn: <ip SNP V ([TO] V PP)

*XXX: I gotta play with this one
%mor: IPROS|I **AUX|got~to** play P|with DDEM|this PRO|one
%syn: <ip SNP AUX V PP

5 Other I-related material

At %mor:	I = infl	At %syn:	V = verb
	T = tense		NEG = negative
	A = agreement		
	NEG = negative		

5.1 Tense

On the %mor tier, tense is indicated with IT|xxx and by suffixing an *-ed* or *-d* to the verb, depending on whether the morphology is regular.

- **-ed** is used for regular past forms, as with *walked*, *skipped*, etc.: IT|walk**-ed** and IT|skip**-ed**.
- **-d** is used for irregular pasts like *ran*, *slid*, as in IT|run**-d** and IT|slide**-d**.

*XXX: I coughed all night
%mor: IPROS|I **IT|cough-ed** all night

*XXX: I slept all night
%mor: IPROS|I **IT|sleep-d** all night

- Verbs that should have been tensed, but do not bear tense morphology are coded as IT|xxx, and their missing tense morphology is coded as -=0ed or -=0d (see section 13.1 for deletion errors).

*XXX: I cough a lot yesterday
%mor: IPROS|I **IT|cough-=0ed** DART|a DCAR|lot yesterday

*XXX: I sleep very little last night
%mor: IPROS|I **IT|sleep-=0d** very DCAR|little last night

- Verbs that do not display overt tense morphology, and would not have, either because the sentence is not past tense or the (irregular) verb does not distinguish past and nontensed forms, e.g., *hit*, *bet*, *set*, etc., are not coded for tense on the %mor tier.

*XXX: I always cough at night
%mor: IPROS|I always **cough** P|at night

*XXX: I hit the wall with my head last night
%mor: IPROS|I **hit** DART|the wall P|with DPOSD|my head last night

On the %syn tier, verbs that are inflected for tense are indicated with a lower case ‘f’ following a ‘V’, as in Vf (‘f’’ for ‘finite’), both with regular and irregular verbs.

*XXX: I coughed all night
%mor: IPROS|I **IT|cough-ed** all night
%syn: <ip SNP **Vf** ADVP

*XXX: I slept all night
%mor: IPROS|I **IT|sleep-d** all night
%syn: <ip SNP **Vf** ADVP

- Verbs that should have appeared as tensed verbs but did not, are marked with an 'n' (for 'nonfinite') after the 'V' on the %syn tier.

*XXX: I cough a lot yesterday
%mor: IPROS|I **IT|cough-=0ed** DART|a DCAR|lot yesterday
%syn: <ip SNP **Vn** ADVP ADVP

*XXX: I sleep very little last night
%mor: IPROS|I **IT|sleep-=0d** very DCAR|little last night
%syn: <ip SNP **Vn** ADVP ADVP

Note: Vn is also used in cases of an error, where tense morphology is missing.

- Verbs that should not have had overt tense morphology are not marked on the %syn tier.

*XXX: I always cough at night
%mor IPROS|I always **cough** P|at D|night
%syn: <ip SNP ADVP **V** ADVP

*XXX: I hit the wall with my head kast night
%mor: IPROS|I **hit** DART|the wall P|with DPOSD|my head last night
%syn: <ip SNP **V** DNP PP ADVP

Caution: The following forms do not bear tense morphology and are not marked at either the %mor or the %syn tier.

- Present tense verbs with plural subjects, as in *They talk a lot*.
- Non-finite verbs that follow modals, *do*, or the infinitive marker *to*, as in *They should talk a lot*, *They don't talk a lot*, and *They tend to talk a lot*.
- Past tense forms with null past forms, as in *The ball hit the goalpost five times during the game*.

*XXX: mannekin will appear in a dream
%mor: mannekin IAUX|will **appear** P|in DART|a dream
%syn: <ip SNP AUXM **V** PP

*XXX: I want to leave
%mor: IPROS|I want I|to **leave**
%syn: <ip SNP V ([TO] **V**)

5.2 Agreement

On the %mor tier, 3rd person singular agreement is indicated with IA|xxx and suffixing an *–s* to the stem, as in *walks*, IA|walk**-s**, and *does*, IA|do**-s**.

*XXX: he talks a lot
%mor: IPROS|he **IA|talk-s** DART|a DCAR|lot

- Verbs that should have had agreement morphology but do not are still coded with IA|xxx, but the verb is followed by **-=0s** (see deletion errors in section 13.1).

*XXX: he go to school every day
%mor: IPROS|he **IA|go-=0s** P|to school DQUA|every day

- Verbs that would not have agreed with the subject (either the verb is in the past tense or the subject is not 3rd singular) are not coded for agreement.

*XXX: they go to school every day
%mor: IPROS|they **go** P|to school DQUA|every day

On the %syn tier, agreement is indicated as a Vf, like tensed verbs.

*XXX: he talks a lot
%mor: IPROS|he **IA|talk-s** DART|a DCAR|lot
%syn: <ip SNP **Vf** ADV

*XXX: she likes me
%mor: IPROS|she **IA|like-s** PROO|me
%syn: <ip SNP **Vf** DNP

- Verbs that should have been marked for agreement, but on which the agreement or tense morphology is missing, are coded on the synactic tier as **Vn**, with a lower case 'n' attached to 'V'.

*XXX: he go to school every day
%mor: IPROS|he **IA|go-=0s** P|to school DQUA|every day
%syn: <ip SNP **Vn** PP ADVP

Caution: Vn is used only in cases of an error, where obligatory agreement or tense

morphology is missing.

- Verbs that would not have been overtly marked for agreement are not coded.

*XXX: they go to school every day
%mor: IPROS|they **go** P|to school DQUA|every day
%syn: <ip SNP **V** PP ADVP

Caution: The following forms do not bear agreement morphology, and thus are not coded for agreement/finiteness at either the %mor or the %syn tier.

- Verbs in sentences with plural subjects, *They talk a lot*.
- Past tense verbs, as in *He talked a lot*.
- Non-finite verbs that follow modals, *does*, or the infinitive marker *to*, as in *He should talk a lot*, *He doesn't talk a lot*, and *He tends to **talk** a lot*.

*XXX: mannekin will appear in a dream
%mor: mannekin IAUX|will **appear** P|in DART|a dream
%syn: <ip SNP AUXM **V** PP

*XXX: we live on that street
%mor: IPROS|we **live** P|on DDEM|that street
%syn: <ip SNP **V** PP

5.3 Participial forms

Verbs in participial form do not have a code-pipe combination on the %mor tier. The affixal morphology is indicated by a hyphen separating the verb from the suffix -ing, -en, or -n. At %syn, the V is suffixed by an -ing, -en, or -n.

*XXX: we are leaving now
%mor: IPROS|we IAUX|be-s **leave-ing** now
%syn: <ip SNP AUXB **V-ing** ADVP

*XXX: you have laughed at me
%mor: IPRO|you IAUX|have **laugh-en** P|at PROO|me
%syn: <ip SNP AUXH **V-en** PP

*XXX: we must have left early
%mor: IPROS|we IAUX|must AUX|have **leave-n** early
%syn: <ip SNP AUXM AUXH **V-n** ADVP

5.4 Negative

Negation is coded with NEG|xxx on the %mor tier and as NEG on the %syn tier.

*XXX: we will not leave early
%mor: IPROS|we IAUX|will **NEG|not** leave early
%syn: <ip SNP AUXM **NEG** V ADVP

Negation cliticized onto an auxiliary is indicated with a tilde '~' at %mor: ~NEG|. The cliticization is not coded at %syn.

*XXX: we won't leave early
%mor: IPROS|we **IAUX|will~NEG|not** leave early
%syn: <ip SNP **AUXM NEG** V ADV

*XXX: he doesn't like kids
%mor: IPROS|he **IAUX|do-s~NEG|not** like D|kid-pl
%syn: <ip SNP AUXD **NEG** V DNP

Negative adverbs like *never* are simply treated as adverbs.

*XXX: we will never leave early
%mor: IPROS|we IAUX|will **never** leave early
%syn: <ip SNP AUXM **ADVP** V ADVP
%lex: **ADV|never** V|leave ADV|early

The *no* in constituent negation is treated as a quantifier.

*XXX: no student will leave early
%mor: **DQUA|no** student IAUX|will leave early
%syn: <ip **SNP** AUXM V ADVP

5.5 Imperativse

Imperatives are not coded on the %mor tier. The verb is suffixed with an -imp at %syn.

*XXX: leave early
%mor: leave early
%syn: V-imp ADVP

With negative imperatives, the auxiliary *do* is fronted to C, so they precede the subject as in *Don't anybody leave this room*. In these cases, *do* is marked as a CAUX at %mor, and as AUXD NEG following <cp at %syn.

Subjects are missing in most negative imperatives, but the auxiliary is still expected to be at the head of CP.

*XXX: don't leave early
%mor: **CAUX|do~NEG|not** leave early
%syn: **<cp AUXD NEG** <ip V ADVP

Caution: In child language, do not code *don't* as a C-level aux unless:

(a) There are other, independent reasons to assume that the child projects a CP in the structure and moves auxiliaries there, e.g., other cases of aux inversion in questions, etc.

(b) There is a subject in the structure, e.g., *Don't you...* or *Don't anyone....*

Until there is such evidence elsewhere, do not code the position of the aux.

*XXX: don't leave early
%mor: **AUX|do~NEG|not** leave early
%syn: **AUXD NEG** V ADVP

5.6 Tag questions

Tag questions are indicated on the %syn tier by the code TAG, followed by the category of the auxiliary used in the tag question: e.g., TAGB, TAGD, etc. At the %mor tier an auxiliary is prefixed with T to indicate that it is a tag question AUX: TAUX. *he, she,* etc. are also coded with a T: TPROS|he etc. So tag questions are coded very similarly to root clause questions, but on %mor instead of I you have T. On %syn, instead of AUX, the code TAG is used (e.g., TAGB instead of AUXB). No <ip or <cp boundaries are shown.

Tags with cliticized negation code the cliticization on the %mor tier in the usual way, and the %syn tier simply contains a NEG after the TAGD or TAGB.

*XXX: This is mine, isn't it
%mor: DPRO|this IAUX|be-s DPOS|mine **TAUX|be-s~NEG|not PROS|it**
%syn: <ip SNP AUXB PNP **TAGB NEG** SNP

6 Predicates

6.1 Verbs

See section 5.1 and 5.2 for tense and agreement marking, 5.3 for participials, 5.4 for negatives, and 5.5 for imperatives.

6.2 Verb-particles

Verb particle pairs are coded as distinct elements at %mor and %syn, as PRT, but as a single unit on the %lex tier.

*XXX: I threw it away
%mor: IPROS|I **IT|throw-d** PRO|it **PRT|away**
%syn: <ip SNP **Vf** DNP **PRT**
%lex: **V|throwaway**

*XXX: the little girl ate up her food
%mor: DART|the little girl **IT|eat-d PRT|up** DPOSD|her food
%syn: <ip SNP **Vf PRT** DNP
%lex: ADJ|little N|girl **V|eatup** N|food

*XXX: mannekin fell down on the floor
%mor: mannekin **IT|fall-d PRT|down** P|on DART|the floor
%syn: <ip SNP **Vf PRT** PP
%lex: N|mannekin **V|falldown** N|floor

Caution: Particles look like prepositions but they are different in three ways:
(a) The object can come between the particle and the verb. This is optional with full NPs, as in *Bill took down the picture* and *Bill took the picture down*, but obligatory with pronouns, as in **Bill took down it* vs. *Bill took it down*.
(b) Particles do not have complements:
They can occur with intransitive verbs, as in *The soup cooled down*, *She spoke up*, etc., as opposed to prepositions, as in **I talked with*, **He moved to*.
(c) They cannot be pied-piped, as in **Up which bottle did you fill?* (cf. *You filled up that bottle*), **In which forms did she turn?* (cf. *She turned in those forms*), as opposed to prepositions, as in *Up which ladder did he climb?*, *In which room did you walk?*.

6.3 Predicative NPs

Predicative NPs establish some identity for the subject. The noun is coded on the %mor tier as other nominal categories are (see sections 1, 6, and 7 for more detail on the %mor tier). On the %syn tier, they are indicated as a PNP (predicative NP), and on the %lex tier, they are Ns. The *be* that occurs with them is a copula, but it is coded as the auxiliary *be* on the %mor and %syn tiers. It is coded as a lexical verb at %lex.

*XXX: they are students
%mor : IPROS|they **IAUX|be-s D|student-pl**
%syn: <ip SNP **AUXB PNP**
%lex: **V|be N|student**

*XXX: he is mannekin the ghost
%mor: IPROS|he **IAUX|be-s mannekin DART|the ghost**
%syn: <ip SNP **AUXB PNP ([PMDF] NP)**
%lex: **V|be N|mannekintheghost**

*XXX: he is a jerk
%mor: IPROS|he **IAUX|be-s DART|a jerk**
%syn: <ip SNP **AUXB PNP**
%lex: **V|be N|jerk**

In the existential *there* construction, the NP that follows the verb *be* is a predicative NP:

*XXX: there is some candy on the table
%mor: PROE|there **IAUX|be-s DQUA|some candy** P|on DART|the table
%syn: <ip EXP **AUXB PNP** PP
%lex: **V|be N|candy** N|table

*XXX: once there was a good boy
%mor: once PROE|there **IAUX|be-s-d DART|a good boy**
%syn: ADVP <ip EXP **AUXB PNP**
%lex: ADV|once **V|be ADJ|good N|boy**

Predicative NPs are also found in the context of verbs like *call*, *consider*, and others (small clauses).

*XXX: I will call him mannekin
%mor: IPROS|I IAUX|will **call** PROO|him **mannekin**
%syn: <ip SNP AUXM **V** ([SC] DNP **PNP**)
%lex: V|call **N|mannekin**

*XXX: we consider him a fool
%mor: IPROS|we **consider** PROO|him **DART|a fool**
%syn: <ip SNP **V** ([SC] DNP **PNP**)
%lex: **V|consider N|fool**

6.4 Predicative APs

Predicative APs indicate the state of the subject. Adjectives are not coded on the %mor tier, but they are coded as PAPs on %syn, and as an ADJ|xxx on the %lex tier. In cases

where they are found with a copula, the copula is coded as an auxiliary on the morphological tier, but as a verb on the lexical tier (both 'be' and 'get').

*XXX: she was tired
%mor: IPROS|she **IAUX|be-s-d tired**
%syn: <ip SNP **AUXB PAP**
%lex: **V|be ADJ|tired**

*XXX: and then she got **angry**
%mor: ONJ|landthen IPROS|she IAUX|get-d **angry**
%syn: [ONJ] <ip SNP AUXG **PAP**
%lex: **V|get ADJ|angry**

*XXX: I consider her undecided
%mor: IPROS|I consider PROO|her **un-decide-en**
%syn: <ip SNP V DNP **PAP**
%lex: V|consider **ADJ|undecided**

6.5 Predicative PPs

Predicative PPs assert the relationship between the subject and a PP. Most often the PP is a locative phrase that corresponds to either some location or state metaphorically used as a location. Crucial distinctions between predicative, adverbial, and argument PPs are the following:

- Predicative PPs define the event/state itself, as in *He is in this class*, *She is on vacation*, etc.
- Adverbial PPs provide information that is not essential to the event/state, as in *He is asleep in the class*, *She is in Hawaii for vacation*, etc.
- Argument PPs are integral components of the event/state, as in *I brought this for Bill*, *I am talking to Mary*, etc.

The P and its complement NP are marked the usual way on the %mor tier. On the %syn tier, the predicative PP is a PPP. On the %lex tier the P is not marked, the N is, if it is an open class lexeme.

The *be* in the predicative construction is a copula, but it is coded as an auxiliary on the morphological and syntactic tiers, as in other cases, and as a verb on the lexical tier.

*XXX: we were in that room
%mor: IPROS|we **IAUX|be-s-d P|in DDEM|that room**
%syn: <ip SNP **AUXB PPP**
%lex: V|be N|room

*XXX: she is from arabia
%mor: IPROS|she **IAUX|be-s P|from afghanistan**
%syn: <ip SNP **AUXB PPP**
%lex: V|be N|afghanistan

*XXX: I want them out of here
%mor: IPROS|I want PROO|them **P|out P|of here**
%syn: <ip SNP V DNP **PPP**
%lex: V|want

Caution 1: PPs that simply modify the event are adverbials, as in the following examples.

*XXX: they are sleeping in that room
%mor: IPROS|they IAUX|be-s sleep-ing **P|in DDEM|that room**
%syn: <ip SNP AUXB V-ing **PP**

*XXX: we were sleepy in that room
%mor: IPROS|we IAUX|be-s-d sleep-y **P|in DDEM|that room**
%syn: <ip SNP AUXB PAP **PP**

Caution 2: PPs that are integral to the main event are arguments, as in the following examples.

*XXX: I took them out of there
%mor: IPROS|I IT|take-d PROO|them **P|out P|of there**
%syn: <ip SNP Vf DNP **PP**

*XXX: we talked about the elections
%mor: IPROS|we IT|talk-ed **P|about DART|the D|election-pl**
%syn: <ip SNP Vf **PP**

7 Adverbials

7.1 Adverbs

Adverbs are coded as ADVP on the %syn and %lex tiers. They are not coded on the %mor tier except for their morphological make-up in cases of complex forms.

*XXX: perhaps she will leave early
%mor: **perhaps** IPROS|she IAUX|will leave **early**
%syn: **ADVP** <ip SNP AUXM V **ADVP**
%lex: **ADV|perhaps** V|leave **ADV|early**

*XXX: he quickly read the book
%mor: IPROS|he **quick-ly** IT|read-d DART|the book
%syn: <ip SNP **ADVP** V DNP
%lex: **ADV|quickly** V|read N|book

*XXX: sometimes we knowingly act badly
%mor: **DQUA|some+times** IPROS|we **know-ing-ly** act **bad-ly**
%syn: **ADVP** <ip SNP **ADVP** V **ADVP**
%lex: **ADV|some+times ADV|knowingly** V|act **ADV|badly**

7.2 NP adverbials

NP adverbials are typically time-denoting phrases that appear without prepositions. They are coded on the %mor and %lex tiers like ordinary NPs, but as ADVPs on the %syn tier.

*XXX: tomorrow I will start school
%mor: **tomorrow** IPROS|I IAUX|will start school
%syn: **ADVP** <ip SNP AUXM V DNP
%lex: **N|tomorrow** V|start N|school

*XXX: that year we talked on the phone almost every day
%mor: **DDEM|that year** IPROS|we IT|talk-ed P|on DART|the phone **almost DQUA|every day**
%syn: **ADVP** <ip SNP Vf PP **ADVP**
%lex: **N|year** V|talk N|phone **ADV|almost N|day**

*XXX: mondays she leaves home a little late
%mor: **D|monday-pl** IPROS|she IA|leave-s home **DART|a DCAR|little ADJ|late**

%syn: **ADVP** <ip SNP Vf DNP **ADVP**
%lex: **N|monday** V|leave N|home **ADJ|late**

7.3 PP adverbials

PP adverbials are coded as PPs on all tiers, including %syn.

*XXX: in that room there are a lot of toys
%mor: **P|in DDEM|that room** PROE|there IAUX|be-s DART|a DCAR|lot P|of D|toy-pl
%syn: **PP** <ip EXP AUXB SNP
%lex: **N|room** V|be N|toy

*XXX: **between tuesday and friday** you and I should meet **twice**
%mor: **P|between tuesday and friday** PRO|you and IPRO|I IAUX|should meet **twice**
%syn: **PP** <ip SNP AUXM V **ADVP**
%lex: **N|tuesday N|friday** V|meet

Note: Although *ago* is a postposition, it is treated as a preposition for practical purposes and coded as a P: **P|ago**.

*XXX: two years ago he took that test about three times in a row
%mor: **DCAR|two D|year-pl P|ago** IPROS|he IT|take-d DDEM|that test **about DCAR|three D|time-pl P|in DART|a row**
%syn: PP <ip SNP Vf DNP ADVP
%lex: Adj|two N|year **V|take N|test** ADV|about Adj|three N|time N|row

8. Conjunctions

All clausal conjunction elements are coded as ONJ. They are ONJ|xxx on the %mor tier, [ONJ] on the %syn tier.

- ***Discourse conjunctions***

In the course of telling a story, each sentence may be linked to the previous one with a sentence-initial *and*, *and then*, *then*, *so*, etc. (For the purpose of determining MLUs, a single utterance may contain two conjoined clauses. Every additional combination of two conjoined clauses will constitute a separate utterance.)

*XXX: it started to rain so he started to run.
%mor: PRO|it IT|start-ed I|to rain ONJ|so IPROS|he IT|start-ed ...
%syn: <ip SNP Vf ([TO] rain) [ONJ] <ip SNP Vf ...
*XXX: and then he fell into a puddle of mud and he got wet.
%mor: ONJ|landthen IPROS|he IT|fall-d P|into DART|a puddle P|of mud ONJ|and IPROS|he IAUX|get-d wet
%syn: [ONJ] <ip SNP Vf PP [ONJ] <ip SNP Vf PAP

8.2 Sentence-internal conjunctions

Only CP-level, IP-level, and VP-level conjunctions are coded. NP-level conjunctions are ignored and coded as a single NP, as are all X-level conjunctions.

Two-part conjunction elements like *either...or* and *both...and* are each coded as ONJ|xxx.

*XXX: Either he stops or I leave.
%mor: **ONJ|either** IPROS|he IA|stop-s **ONJ|or** IPROS|I leave
%syn: **[ONJ]** <ip SNP Vf **[ONJ]** <ip SNP V

*XXX: we both order your washer and install it
%mor: IPROS|we **ONJ|both** order DPOSD|your wash-ag **ONJ|and** install PRO|it
%syn: <ip SNP **[ONJ]** V DNP PP **[ONJ]** V DNP

Caution: Two-part NP conjunctions are not coded with ONJ|xxx

*XXX: either you or your parents should come
%mor: **either** PROS|you **or** DPOSD|your D|parent-pl IAUX|should come
%syn: <ip SNP AUXM V

9 The C-System

9.1 Complementizers

There are various types of complementizers.

9.1.1 Complementizers of Argument Clauses, *that*, *if*, *whether*, *for*:

The complementizer of an argument clause is a CCOM. It is a CCOM|xxx at %mor, a [CCOM] at %syn.

*XXX: I know that he is leaving
%mor: IPROS|I know **CCOM|that** IPROS|he IAUX|be-s leave-ing
%syn: <ip SNP V (<cp **[CCOM]** <ip SNP AUXB V-ing)

*XXX: I wonder if he is leaving
%mor: IPROS|I wonder **CCOM|if** IPROS|he IAUX|be-s leave-ing
%syn: <ip SNP V (<cp **[CCOM]** <ip SNP AUXB V-ing)

*XXX: I wonder whether he is leaving
%mor: IPROS|I wonder **CCOM|whether** IPROS|he IAUX|be-s leave-ing
%syn: <ip SNP V (<cp **[CCOM]** <ip SNP AUXB V-ing)

Caution: The CCOM *if* can be confused with the conditional *if* (a CADJ, see below). An *if* is a CCOM if it can be replaced by *whether*.

For null complementizers (*that*-deletion), see section 12 on deletion.

For the infinitival complementizer *for*, see section 11.1.

9.1.2 Complementizers of Adjunct Clauses, *because*, *after*, *if*, *when*, etc.

The complementizer of an adjunct clause is a CADJ. It is a CADJ|xxx at %mor, a [CADJ] at %syn.

*XXX: I stayed at school because it was raining
%mor: IPROS|I IT|stay-ed P|at school **CADJ|because** PRO|it IAUX|be-s-d rain-ing
%syn: <ip SNP Vf PP (<cp **[CADJ]** <ip SNP AUXB V-ing)

*XXX: I stayed at school after I finished my breakfast
%mor: IPROS|I IT|stay-ed P|at school **CADJ|after** IPROS|I IT|finish-ed

DPOSD|my breakfast
%syn: <ip SNP Vf PP (<cp **[CADJ]** <ip SNP Vf DNP)

*XXX: I will stay at school if it rains
%mor: IPROS|I IAUX|will stay P|at school **CADJ|if** IPRO|it IA|rain-s
%syn: <ip SNP AUXM V PP (<cp **[CADJ]** <ip SNP Vf)

*XXX: I stay at school while I finish my homework
%mor: IPROS|I stay P|at school **CADJ|while** IPROS|I finish DPOSD|my home+work
%syn: <ip SNP V PP (<cp **[CADJ]** <ip SNP V DNP)

How to tell adjunct and argument clauses apart:

- Adjunct clauses are typically moveable, argument clauses are less so.
- The adjunct-clause *if* is conditional, the argument-clause *if* is an embedded question.
- The adjunct-clause *when* fixes temporal reference like *before* and *after*, (also *while* in the progressive tense), but the argument-clause *when* is an embedded question.
- Adjunct clauses can be removed without any major damage to the sentence, but removing an argument clause leads to an incomplete sentence.

9.2 Fronted Wh-phrases

A fronted Wh-phrase is a CWH|xxx on %mor, but an XWH on %syn. X ranges over S (subject), D (direct object), I (indirect object), A (adjunct), O (oblique), and P (predicate), i.e., SWH, DWH, IWH, AWH, OWH, PWH.

Note 1: The *who - whom* distinction is ignored and not expected to show up.
Note 2: IWH will show up sparingly because of the tendency of speakers to strand the preposition.

Complex Wh-phrases, e.g., *which NP*, *what NP*, and *whose NP*, are coded as XWHNP

- Main clause Wh-phrases mark the CP boundary, so they are preceded by <cp. In case of aux-inversion, the fronted aux also appears outside the <ip.

*XXX: who did she see?
%mor: **CWH|who** CAUX|do-d IPROS|she see
%syn: <cp **DWH** CAUXD <ip SNP V

*XXX: when did he leave?
%mor: **CWH|when** CAUX|do-d IPROS|he leave
%syn: <cp **AWH** CAUXD <ip SNP V

*XXX: who are they?
%mor: **CWH|who** CAUX|be-s IPROS|they
%syn: <cp **PWH** AUXB <ip SNP

- Subject Wh-questions do not trigger aux-inversion, so it is hard to tell where the subject position, and therefore, the IP boundary is. These questions do not have the <ip indicated.

*XXX: who told you that?
%mor: **CWH|who** IT|tell-d PRO|you DPRO|that
%syn: <cp **SWH** Vf INP DNP

- Preposition stranding moves an oblique (complement of P) Wh-phrase to the CP, leaving behind the preposition.

*XXX: what are you talking about?
%mor: **CWH|what** CAUX|be-s PRO|you talk-ing P|about
%syn: <cp **OWH** AUXB <ip SNP V-ing P

- Fronted Wh-phrases in subordinate clauses are coded roughly the same way as those in main clauses, but are coded with a preceding <cp to mark the embedded CP clausal boundary.

*XXX: I wonder who you saw
%mor: IPROS|I wonder **CWH|who** PROS|you IT|see-d
%syn: <ip SNP V (<cp **[DWH]** <ip SNP Vf)

*XXX: I don't know when you left
%mor: IPROS|I IAUX|do~NEG|not know **CWH|when** PROS|you IT|leave-d
%syn: <ip SNP AUXD NEG V (<cp **[AWH]** <ip SNP Vf)

*XXX: I told you who that was
%mor: IPROS|I IT|tell-d PRO|you **CWH|who** DPRO|that IAUX|be-s-d
%syn: <ip SNP Vf INP (<cp **[PWH]** <ip SNP IAUXB)

- Note that subordinate yes-no questions are introduced by a CCOM.

*XXX: I don't know whether you left
%mor: IPROS|I IAUX|do~NEG|not know **CCOM|whether** PROS|you IT|leave-d
%syn: <ip SNP AUXD NEG V (<cp **[CCOM]** <ip SNP Vf)

- Also note that temporal adjunct clauses may be introduced by a wh-phrase

*XXX: I left when you left
%mor: IPROS|I IT|leave-d **CADJ|when** PROS|you IT|leave-d
%syn: <ip SNP Vf (<cp **[CADJ]** <ip SNP Vf)

Caution 1: On the %syn tier, argument Wh-phrases are SWH, DWH, or IWH. All adjunct Wh-phrases are AWH (*how*, *when*, *how much*, etc), predicative ones are PWH. OWH is reserved for P-stranding.

Caution 2: On the %mor tier, *whose* is a DPOSD| just like any other possesive pronoun/determiner.

Also see stranding, long distance extraction, topicalization, and dislocation, all in section 11.

9.3 Relative clauses

Relative pronouns are CREL at %mor, but XWHREL and XREL at %syn, where X ranges over S, D, I, A, P, and O, i.e., SWHREL, DWHREL, IWHREL, AWHREL, PWHREL, and OWHREL; or SREL, DREL, IREL, AREL, PREL, and OREL. On the syntactic tier, these will all be prefaced with <cp, marking their site as within CP.

Note: IWHREL and IREL will probably be rare because of the preference for speakers to strand the preposition.

*XXX: I saw the man who was leaving
%mor: IPROS|I IT|see-d DART|the man **CREL|who** IAUX|be-s-d leave-ing
%syn: <ip SNP Vf DNP (<cp **[SWHREL]** AUXB V-ing)

*XXX: I saw the man who you pushed
%mor: IPROS|I IT|see-d DART|the man **CREL|who** PROS|you IT|push-ed
%syn: <ip SNP Vf DNP (<cp **[DWHREL]** <ip SNP Vf)

*XXX: I visited the place where you were staying
%mor: IPROS|I IT|visit-ed DART|the place **CREL|where** PROS|you IAUX|be-s-d stay-ing
%syn: <ip SNP V-ed DNP (<cp **[AWHREL]** <ip SNP AUXB V-ing)

*XXX: I saw the man that was leaving
%mor: IPROS|I IT|see-d DART|the man **CREL|that** IAUX|be-s-d leave-ing

%syn: <ip SNP Vf DNP (<cp **[SREL]** AUXB V-ing)

*XXX: that is the kind of person that she is
%mor: DPRO|that IAUX|be-s DART|the kind P|of person **CREL|that** IPROS|she IAUX|be-s
%syn: <ip SNP AUXB PNP (<cp **[PREL]** <ip SNP AUXB)

*XXX: you are sitting on the chair that I was looking for
%mor: PROS|you IAUX|be-s sit-ing P|on DART|the chair **CREL|that** IPROS|I IAUX|be-s-d look-ing P|for
%syn: <ip SNP AUXB V-ing PP (<cp **[OREL]** <ip SNP AUXB V-ing P)

For stranding and long distance extraction, see section 3.

For null relativizers (or relative pronoun deletion), see section 12.1.

10 Extraction

10.1 P-stranding

Wh-movement (in a question or relative clause) that strands a preposition is identified with a stranded preposition, P, and a fronted Wh-phrase prefixed with an O (oblique).

*XXX: Who did he talk to?
%mor: **CWH|who** CAUX|do-d IPROS|he talk **P|to**
%syn: <cp **OWH** AUXD <ip SNP V **P**

10.2 Long Distance Extraction

Wh-phrases moved long distance;i.e., from a subordinate clause to the main clause, are suffixed with an E for extracted.

The suffix E does not affect the S/D/I/A/P/O prefixes (see section 13 for *that*-deletion).

*XXX: who do you think left early ?
%mor: **CWH|who** CAUX|do PROS|you think CCOM|0that IT|leave-d early
%syn: <cp **SWHE** AUXD <ip SNP V (<cp [0CCOM] <ip Vf ADVP)

*XXX: what do you think I should eat ?
%mor: **CWH|what** CAUX|do PROS|you think CCOM|0that IPROS|I IAUX|should eat
%syn: <cp **DWHE** AUXD <ip SNP V (<cp [0CCOM] <ip SNP AUXM V)

Short distance adjunct extraction:

*XXX: why do you think he left ?
%mor: **CWH|why** CAUX|do PROS|you think CCOM|0that IPROS|he IT|leave-d
%syn: <cp **AWH** CAUXD <ip SNP V (<cp [0CCOM] <ip SNP Vf)

Long distance adjunct extraction:

*XXX: why do you think he left ?
%mor: **CWH|why** CAUX|do PROS|you think CCOM|0that IPROS|he IT|leave-d
%syn: <cp **AWHE** CAUXD <ip SNP V ([0CCOM] <ip SNP Vf)

Caution 1: In cases of ambiguity, look for the most natural reading in the discourse. If both WH and WHE are equally possible, indicate this in the %com line.

Caution 2: In cases with three levels of embedding, insert a %com line to clarify which

clause the WHE comes from.

10.3 Topicalization

There is no special way of coding topicalization. Only the placement of the moved constituent is coded for..

10.4 Dislocation

The dislocated constituent is prefixed with a T (for topic) in addition to its grammatical function (S/D/I/A/O). The resumptive pronoun is marked as if it were an ordinary argument pronoun, i.e, TSNP, TDNP, TINP, TPP, TONP, etc. The TXNP is placed to the left In left dislocation and to the right in right dislocation.

*XXX: casper the friendly ghost, last night I had him in my dream
%mor: **casperthefriendlyghost** last night IPROS|I IT|have-d **PROO|him** P|in DPOSD|my dream
%syn: **TDNP** ADVP <ip SNP Vf **DNP** PP

*XXX: he doesn't like to scare anyone, casper the friendly ghost.
%mor: **IPROS|he** IAUX|do-s~NEG|not like I|to scare DQUA|any+PRO|one **casperthefriendlyghost**
%syn: <ip **SNP** AUXD NEG V ([TO] V DNP) **TSNP**

- Long distance dislocation can be marked by suffixing an E.

*XXX: my best friend bill I am not sure that you know him well
%mor: **DPOSD|my good-est friend bill** IPROS|I IAUX|be-s NEG|not sure CCOM|that PROS|you know **PROO|him** well
%syn: **TDNPE** <ip SNP AUXB NEG PAP (<cp [CCOM] <ip SNP V **DNP** ADVP)

- Indicate dislocation from inside an NP (from the specifier position) on the %com line.

*XXX: my best friend bill his mother does not let us play in the living room
%mor: **DPOSD|my good-est friend bill** DPOSD|his mother IAUX|do-s~NEG|not let PROO|us play P|in DART|the living+room
%syn: **TNP** <ip **SNP** AUXD NEG V ([SC] SNP V PP)
%com: the dislocated NP corresponds to the specifier of the SNP.

11 Complementation

11.1 Infinitivals

Infinitival clauses are introduced with a [TO] in the clause initial position on the %syn tier, and *to* is prefixed with an I|xxx on the %mor tier.

*XXX: he tried to leave
%mor: IPROS|he IT|try-ed **I|to** leave
%syn: <ip SNP Vf (**[TO]** V)

If there is a *for* complementizer in the sentence, it is coded as a CCOM at both %mor and %syn. Place the subject between CCOM and TO on the %syn:

*XXX: he wants very much for me to leave
%mor: IPROS|he IA|want-s very much **CCOM|for** PRO|me I|to leave
%syn: <ip SNP Vf ADVP (<cp **[CCOM]** <ip SNP [TO] V)

- *for*-complementizers are never specified if the subject of the clause is missing, i.e., with PRO subjects. They are indicated only if the clause is in principle a control clause that has an overt (accusative) subject, with or without an overt *for*, e.g., with verbs like *want*, *would like*, *need* (verbs that would typically take subjunctive clauses in many languages).

*XXX: I need to be more careful
%mor: IPROS|I need I|to AUX|be more care-ful
%syn: <ip SNP V ([TO] AUXB PAP)

*XXX: I need you to be more careful
%mor: IPROS|I need **CCOM|0for PRO|you** I|to AUX|be more care-ful
%syn: <ip SNP V (<cp **[0CCOM]** <ip **SNP** [TO] AUXB PAP)

- *for*-deleted verbs can be identified by two properties:
  (a) They do not allow the lower subject to passivize, e.g., **He was wanted to go*, **You are needed to be more careful*.
  (b) The infinitival portion cannot be separated from the accusative argument as a distinct argument: **She wants me my departure*, **I need you your attention*.

- Exceptional Case Marking (raising-to-object) verbs take infinitival complements with accusative subects, but they are not in principle control verbs. They are CP-deletion verbs that have no CCOM, e..g., *believe*, *consider*, *assume*.

*XXX: I considered him to be a loser
%mor: IPROS|I IT|consider-ed **PROO|him I|to** AUXB|be DART|a lose-ag

%syn: <ip SNP Vf (**<ip SNP [TO]** AUXB PNP)

CP-deletion verbs can be distinguished from other verbs by two properties:

(a) They allow the accusative argument to passivize, e.g., *He is believed to be hiding in his basement*, *He is considered to be a loser*.

(b) The infinitival portion cannot be separated from the accusative phrase as an independent argument: **I believe him the story*, **I consider him the truthfullness* (with *consider*, a small clause structure is possible).

- Verbs that take three arguments appear to have infinitival complements with accusative subjects, but the infinitivals are actually the DNPs (or INPs) of the main clause, e.g., *promise*, *convince*, *tell*, *ask*. They thus do not have CCOMs either. These verbs are control verbs.

*XXX: she told me to sit still
%mor: IPROS|she IT|tell-d **PROO|me I|to** sit still
%syn: <ip SNP Vf **DNP** (**[TO]** V ADVP)

Three argument verbs or *control* verbs are differentiated form the others as follows:

(a) They allow passivization of the accusative argument, e.g., *I was told to leave quickly*, *I was promised to be given some award*.

(b) They allow the infinitival portion to be a nominal expression, as in *She told me her story*, *I promised him an award*.

- Verbs that cannot take overt subjects (or *for*-complementizers) are not ever specified for a CCOM on either the %mor or the %syn tier, e.g., *try*, *hope*, *intend*, etc.

*XXX: he hopes to be a farmer when he grows up
%mor: IPROS|he IA|hope-s **I|to** AUXB|be DAR|a farm-ag CADJ|when IPROS|he IA|grow-s PRT|up
%syn: <ip SNP Vf (**[TO]** AUXB PNP (<cp [CADJ] <ip SNP Vf PRT))

Caution: See pg. 19 for a description of how to code quasi-modals or semi-auxiliaries like *going to, have to*, etc.
ADVP

11.2 Gerunds

Gerunds are marked as participial clauses, PC, on the %syn tier. If the gerund is a complement of a preposition, the PC domain is preceded by a P.

*XXX: I started running
%mor: IPROS|I IT|start-ed **run-ing**
%syn: <ip SNP Vf (**[PC] V-ing**)

*XXX: I thought about eating dinner
%mor: IPROS|I IT|think-d P|about **eat-ing** dinner
%syn: <ip SNP Vf P (**[PC] V-ing** DNP)

Caution 1: the V that follows the [PC] domain label is marked as a V-ing at both %mor and %syn.

Caution 2: a gerund is a complement clause, it is an argument of the higher verb. Participial modifiers (section 11.4) are secondary predicates that make further assertions about the state of the subject or the object.

11.3 Small Clauses

Small clauses are XP complements with minimal inflectional projections (maybe only accusative Case position) that essentially provide secondary predication. A small clause is headed by a lexical category that is typically in its bare form, not tensed or infinitival:

- VP complements of causatives (*He made me leave early*).
- VP complements of perception verbs (*He saw me leave early*).
- AP or PP resultatives (*I painted the house red*, *I want that man out*).
- AP depictives (*We ate the meat raw*)

The small clause domain is labeled as an [SC] at %syn, and the head of the small clause is coded only for its category, V, ADJ, P, etc. If there is a main clause object that also looks like the small clause subject, code it as the small clause subject.

Caution 1: the small clause domain is not an IP, the SNP of the small clause does not have an <ip preceding it -- hence, the accusative case on the SNP.

Caution 2: The term 'small clause' is a nonspecific label that refers to a variety of categories, VP, AP, PP, etc. Therefore the head of this projection needs to be coded separately. For example, the PP complement of *want* in *I want him off my ship* contains *him off my ship*, so the small clause would have to be ([SC] SNP P ONP). Likewise, the AP complement of *clean* in *I wiped the table clean* contains *the table clean*, so the small clause would be ([SC] SNP ADJ) with no complement for the adjective.

*XXX: he made me eat my carrots
%mor: IPROS|he IT|make-ed **PROO|me** eat DPOSD|my D|carrot-pl
%syn: <ip SNP Vf (**[SC] SNP V DNP**)

*XXX: she needs me in that room
%mor: IPROS|she IA|need-s **PROO|me** P|in DDEM|that room
%syn: <ip SNP Vf (**[SC] SNP P ONP**)

*XXX: we wiped the table clean

%mor: IPROS|we IT|wipe-ed DART|the table clean
%syn: <ip SNP Vf (**[SC] SNP ADJ**)

*XXX: we stored the milk cold
%mor: IPROS|we IT|store-ed DART|the milk cold
%syn: <ip SNP Vf (**[SC] SNP ADJ**)

Note: Resultatives describe the state of the object as a result of the main predicate, e.g., the table is clean as a result of the wiping. Depictives describe the state of the object vis a vis the main predicate, e.g., the milk is cold when it is being stored.

Note: Various forms of *let me...* imperatives are coded as main verb *let* in the imperative form, taking an SC complement:

*XXX: let me look at those pictures
%mor: let PROO|me look P|at DDEM|that-p D|picture-pl
%syn: **V-imp** ([SC] SNP V PP)

Note: Frozen forms like *let me go* or *let go of me* are not coded.

11.4 Participial modification

When the -en and -ing participial forms act as modifiers, the domain is marked at %syn as a MDF. The argument that the MDF is modifying is indicated by prefixing an S, D, I, or O. The form of the verb is marked the usual way with -en and -ing on the %mor and %syn tiers.

*XXX: she picked up the phone exhausted
%mor: IPROS|she IT|pick-ed PRT|up DART|the phone exhaust-ed
%syn: <ip **SNP** Vf PRT DNP (**[SMDF]** V-en)

*XXX: I watched the bird eating the worm
%mor: IPROS|I IT|watch-ed DART|the bird eat-ing DART|the worm
%syn: <ip SNP Vf **DNP** (**[DMDF]** V-ing DNP)

*XXX: I gave it to the man running up and down the hallway
%mor: IPROS|I IT|give-d PRO|it P|to DART|the man run-ing P|upanddown DART|the hall+way
%syn: <ip SNP Vf DNP **INP** (**[IMDF]** V-ing PRT PP)

Caution: The modifier of the postverbal subject in the existential construction is prefixed with an S, not a P.

*XXX: there was a ghost named mannekin

%mor: PROE|there IAUX|be-s-d DART|a ghost name-en mannekin
%syn: <ip EXP AUXB **PNP** (**[SMDF]** V-en PNP)

*XXX: I am working with a man overwhelmed with work
%mor: IPROS|I IAUX|be-s work-ing P|with DART|a man over+whelm-en P|with work
%syn: <ip SNP AUXB V-ing **PP** (**[OMDF]** V-en PP)

Caution: It is often difficult to decide whether or not the participial form is truly a V-en or just a derived adjective, ADJ. The implication of assuming a V-en form is that an ADJ is learned, stored, and used as a single adjectival item. Since the truly participial V-en forms tend to behave like adjectives anyway, the distinction between a V-en and an ADJ is based more on a theoretical bias than any solid empirical fact. One can argue that *tired* and *bored* are lexicalized as adjectives (rather than *tire-en* and *bore-en*), but *amazed* and *terrified* are participial, i.e., *amaze-en* and *terrify-en* . There is no sure way of knowing which is the case, and we will in general assume compositionality

12 Deletion

All cases of deletion are coded with a zero '**0**' before the deleted element at both %mor and %syn.

At %mor: **0** is placed right before the morphological content of what is missing. It may follow a capital code and a pipe, or it could follow a hyphen that indicates a morphological boundary.

At %syn: **0** is placed before the capital code of the missing element.

12.1 Grammatical deletion

There are cases of deletion that are triggered or allowed by the syntactic context, e.g., relative pronoun deletion, complementizer deletion, preposition deletion. The deleted material is usually a syntactic head, an X^0, and the deleted material is always coded at %mor. It is also coded at %syn if it contains significant structural information, which is typically the case for deletions at the C-level.

*XXX: I read the book you were talking about
%mor: IPROS|I IT|read DART|the book **CREL|0that** PROS|you IAUX|be-s-d talk-ing P|about
%syn: <ip SNP Vf DNP (<cp **[0OREL]** <ip SNP AUXB V-ing P)

*XXX: I know he is home
%mor: IPROS|I know **CCOM|0that** IPROS|he IAUX|be-s **P|0at** home
%syn: <ip SNP V (**[0CCOM]** <ip SNP AUXB **PPP)**

*XXX: she wants me to talk
%mor: IPROS|she IA|want-s **CCOM|0for** PROO|me I|to talk
%syn: <ip SNP Vf (**[0CCOM]** <ip SNP [TO] V)

For ungrammatical deletion, see section 13.1, Placement/Deletion errors.

12.2 Elliptical deletion

Ellipsis is a case of deletion that is triggered or allowed by discourse context. The parts that are deleted are typically much larger than those in grammatical deletion, i.e., XPs. The deleted material is omitted from the %mor tier and only coded at %syn in cases of topic or diary drop, gapping, and VP deletion.

*YYY: why was he scared ?
*XXX: because he saw a ghost
%mor: CADJ|because IPROS|he IT|see-d DART|a ghost
%syn: **0CP** (<cp [CADJ] <ip SNP Vf DNP)

*YYY: when was he scared ?
*XXX: when he saw the ghost
%mor: CADJ|when IPROS|he IT|saw DART|the ghost
%syn: **0CP** (<cp [CADJ] <ip SNP Vf DNP)

*XXX: he said he would come back, and he did
%mor: IPROS|he IT|say-d CCOM|0that IPROS|he IAUX|will-d come PRT|back ONJ|and IPROS|he IAUX|do-d
%syn: <ip SNP Vf (<cp [0CCOM] <ip SNP AUXM V PRT [ONJ] <ip SNP AUXD **0VP**)

*YYY: did you write that letter?
*XXX: did that yesterday
%mor: IAUX|do-d DPRO|that yesterday
%syn: **0SNP** AUXD DNP ADVP

*YYY: did you write that letter?
*XXX: I will now
%mor: IPROS|I IAUX|will now
%syn: <ip SNP AUXM **0VP** ADVP

Do not assign grammatical functions S, D, I, O, and P if the verb is ellipsed because in these cases, the underlying structure is only conjectured.

*YYY: what did he eat?
*XXX: apples
%mor: D|apple-pl
%syn: **NP**

*YYY: what is your name?
*XXX: mannekin
%mor: mannekin
%syn: **NP**

13 Errors

There are three types of errors:

- omission/deletion errors refer to the absence of obligatory grammatical elements,; i.e., elements or constituents mandated by the syntax.
- placement errors refer simply to errors with respect to the position of a given element, but not its form
- form errors refer only to form, but not position.

Omission or deletion errors can be viewed as a kind of placement error and will be discussed below together with placement errors. In addition, sometimes more than one type of error co-occurs. In such cases, each error type is coded independently.

Caution: If the transcript makes clear that a particular form is grammatical in the speaker's dialect, even if though it would be ungrammatical in Standard English, the structure in question is not treated as an error.

13.1 Placement/Deletion errors

Placement errors occur when:

- There is an element that should not have been there.
- There is a missing element that should have been there.
- There is an element that is placed lower than it should have been.
- There is an element that is placed higher than it should have been.

Added, extra elements that should not have been there are marked with an equal sign before it, e.g., **=**xxx. If it is the entire category that should have been deleted, '**=**' goes before the category both at %mor and %syn.

*XXX: I want for to go now
%mor: IPROS|I want **CCOM|=for** I|to go now
%syn: <ip SNP V (**[=CCOM]** [TO] V ADVP)

*XXX: I can to play now
%mor: IPROS|I IAUX|can **I|=to** play now
%syn: <ip SNP AUXM **=[TO]** V ADVP

*XXX: who did leave ?
%mor: CWH|who **CAUX|=do-d** leave
%syn: <cp SWH **=AUXD** leave

*XXX: who does he think that left ?
%mor: CWH|who CAUX|do-s IPROS|he think **CCOM|=that** IT|leave-d

%syn: <cp SWH AUXD <ip SNP V (**[=CCOM]** Vf)

If it is a morpheme that should not have been there, the '**=**' is placed before the morpheme at %mor:

*XXX: I sees you
%mor: IPROS|I **IA|see-=s** PRO|you
%syn: <ip SNP V DNP

*XXX: she is sitting on the furnitures
%mor: IPROS|she IAUX|be-s sit-ing P|on DART|the **D|furniture-=pl**
%syn: <ip SNP AUXB V-ing PP

*XXX: there are a lot of peoples
%mor: PROE|there IAUX|be-s DART|a DCAR||lot P|of **D|person-p-=pl**
%syn: <ip EXP AUXB PNP

Missing elements that should have been there are marked by placing an **=0** before it. Again, if what is missing is the entire category, the **=0** is placed before it both at %mor and %syn.

*XXX: I am scared monsters
%mor: IPROS|I IAUX|be-s scared **P|=0of** D|monster-pl
%syn: <ip SNP AUXB V-en **=0P** ONP

However, where context does not make explicit whether a pronoun or a lexically specified NP would have been in the target sentence, a missing DP/NP is marked only on the syntactic tier.

*XXX: she put on the table
%mor: IPROS|she put P|on DART|the table
%syn: <ip SNP V **=0DNP** PP

*XXX: not tell mother
%mor: NEG|not tell mother
%syn: **=0SNP =0AUXD** NEG V INP

Where there is no evidence that the missing verb would bear any inflection, it is coded as a bare verb (only as *be*).

*XXX: I sleeping
%mor: IPROS|I **IAUX|=0be** sleep-ing
%syn: <ip SNP **=0AUXB** V-ing

If the missing material is simply an affixal morpheme, **=0** is placed before it at %mor, and if the missing affix is inflectional morphology on the verb, the verb is coded as a **Vn** at %syn.

*XXX: he see me
%mor: IPROS|he **IT|see-=0d** PROO|me
%syn: <ip SNP **Vn** DNP

*XXX: I ate two sandwich
%mor: IPROS|I IT|eat-d DCAR|two **D|sandwich-=0pl**
%syn: <ip SNP Vf DNP

Material that is ungrammatically low (on the right) in the sentence is prefixed with an **=L** at %syn. It is coded without reference to its ungrammatical placement at %mor.

*XXX: what he can say ?
%mor: CWH|what IPROS|he **IAUX|can** say
%syn: <cp DWH <ip SNP **=LAUXM** V

*XXX: fall down tim
%mor: fall PRT|down **tim**
%syn: V PRT **=LSNP**

*XXX: I took down it
%mor: IPROS|I IT|take-d PRT|down **PRO|it**
%syn; <ip SNP Vf PRT **=LDNP**

Material that is ungrammatically high (on the left) is prefixed with an **=H** at %syn. It is coded without reference to its ungrammatical placement at %mor.

Caution: **=H** is used sparingly. All else being equal =L is preferred over =H. The distribution of =H and =L is as follows:

- If a single =L captures the data as well as a single =H, =L is used.
- If multiple =L's are needed to code data that can otherwise be coded with a single =H, the =H is used.
- Most importantly, the overall sense of what the person is doing in the entirety of the collected sample is the ultimate determining factor.

*XXX: no he can leave
%mor: **NEG|no** IPROS|he IAUX|can leave
%syn: **=HNEG** <ip SNP AUXM V

*XXX: she can it be reading
%mor: IPROS|she IAUX|can **PRO|it** AUX|be read-ing

%syn: <ip SNP AUXM **=HDNP** AUX V-ing

Compare the last example with the following one, in which the error is captured equally with a single =HDNP or a single =LV. =LV is the preferred form.

*XXX: she can it read
%mor: IPROS|she IAUX|can **PRO|it read**
%syn: <ip SNP AUXM **DNP =LV**

13.2 Form errors

By definition, form errors are morphological and marked at %mor only. The error is ignored at %syn.

All form errors in which a morpheme is used that is simply the wrong morpheme are prefixed with an equal sign '**=**', which follows the hyphen '-'.

*XXX: I have listening to story
%mor: IPROS|I **IAUX|=have** V-ing P|to DART|=0the story
%syn: <ip SNP **AUXH** V-ing PP

*XXX: she were sleeping
%mor: IPROS|she **IAUX|be-=s-d** sleep-ing
%syn: <ip SNP **AUXB** V-ing

*XXX: he is my bestest friend
%mor: IPROS|he IAUX|be-s DPOSD|my **good-est-=est** friend
%syn: SNP AUXB PNP

All form errors that are "misselection" errors, in that a morpheme from the wrong cell of the paradigm is selected or the wrong allomorph is used, are coded with an exclamation sign '**!**' preceding the ungrammatical form. These are typically errors that involve regular vs. irregular forms.

*XXX: I goed to school that day
%mor: IPROS|I **IT|go-!ed** P|to school DDEM|that day
%syn: <ip SNP **Vf** PP ADVP

*XXX: all the childs are coming
%mor: DQUA|all DART|the **D|child-!s** IAUX|be-s come-ing
%syn: <ip SNP AUXB V-ing

*XXX: I should have went home
%mor: IPROS|I IAUX|should AUX|have go-!ed P|0to home
%syn: <ip SNP AUXM AUXH **Vf** 0P NP

(The participial/past mistake is treated as a misanalysis of morphological content based on regular and many irregular verb forms. However, if the transcript indicates that such a form is grammatical in the speaker's dialect, it is not treated as an error.)

14 Comparatives

There are two comparative types: (a) those that compare unequally, e.g., *more/-er...than*, and (b) those that compare equally, e.g., *as...as*.

Both types of comparatives may compare adjectives or quantities.

Both comparative types consist of two parts: (a) what is compared, e.g., *tall-er* and *as…*, and (b) what it is being compared with/to e.g., *than* and *as*.

With both types of comparatives, the part that refers to what it is compared with/to can be a PP or a reduced, elliptical sentence: *He is taller **than me*** (PP), cf. *Who is he taller than? He is as tall **as me*** (PP), cf. *Who is he as tall as?* as opposed to, *He is taller **than I (am)*** (CP) and *He is as tall **as I (am)*** (CP).

In every one of these cases, the compared (first) portion of the comparative, is coded as a morphologically complex form:

- If an adjective takes *-er*, it is coded as such at %mor: *tall-er*
- If an adjective takes *more*, it is coded as a compound at %mor: *more+intelligent*.
- If an adjective is being "equalled", it is coded as a compound at %mor: *as+tall*
- If a quantity is of the *more* type, it is coded as DCAR|more.
- If the quantity is of the *as many* or *as much* type, it is coded as DCAR|*as+many* or as DCAR|*as+much*

14.1 Prepositional form

In these forms, the portion that refers to what is being compared with/to is a PP, with *than* and *as* as prepositions. These cases are treated as prepositions on %mor, and as PPs on %syn:

*XXX: he is taller than bill
%mor: IPROS|he IAUX|be-s **tall-er P|than bill**
%syn: <ip SNP AUXB **PAP PP**

*XXX: he is more intelligent than bill
%mor: IPROS|he IAUX|be-s **more+intelligent P|than bill**
%syn: <ip SNP AUXB **PAP PP**

*XXX: he ran faster than bill
%mor: IPROS|he IT|run-d **fast-er P|than bill**

%syn: <ip SNP Vf **AP PP**

*XXX: he is as tall as bill
%mor: IPROS|he IAUX|be-s **as+tall P|as bill**
%syn: <ip SNP AUXB **PAP PP**

*XXX: he ran as fast as bill
%mor: IPROS|he IT|run-d **as+fast P|as bill**
%syn: <ip SNP Vf **AP PP**

*XXX: he ate more apples than me
%mor: IPROS|he IT|eat-d **DCAR|more D|apple-pl P|than PROO|me**
%syn: <ip SNP Vf **DNP PP**

*XXX: he ate as many apples as me
%mor: IPROS|he IT|eat-d **DCAR|as+many D|apple-s P|as PROO|me**
%syn: <ip SNP Vf **DNP PP**

Caution: Almost any case of comparatives with an accusative expression after *than* or *as* is coded as a prepositional comparative. A comparative is marked as the elliptical form only if there is good evidence to do so, nominative Case, some auxiliary, or other reason to assume a hidden (ellipsed) clause there.

14.2 Elliptical form

With the elliptical forms, the second parts, the parts that are being compared to/with, are reduced sentences with major chunks deleted. The ellipitical sentence is introduced by CPR (CP Reduced), which stands for the Ps, *than* and *as*. All the removed parts are coded as deleted at %syn, thus are prefixed with '**0**'. The ellipsed parts are not coded at %mor.

*XXX: he is taller than bill is
%mor : IPROS|he IAUX|be-s **tall-er CPR|than bill IAUX|be-s**
%syn: <ip SNP AUXB **PAP (<cp [CPR] <ip SNP AUXB 0AP)**

*XXX: he is as tall as bill is
%mor: IPROS|he IAUX|be-s **as+tall CPR|as bill IAUX|be-s**
%syn: <ip SNP AUXB **PAP (<cp [CPR] <ip SNP AUXB 0AP)**

*XXX: I ate more apples than she did
%mor: IPROS|I IT|eat-d **DCAR|more D|apple-pl CPR|than IPROS|she IAUX|do-d**
%syn: <ip SNP Vf **DNP (<cp [CPR] <ip SNP AUXD 0VP)**

*XXX: I ate as many apples as she did
%mor: IPROS|I IT|eat-d **DCAR|as+many D|apple-pl CPR|as IPROS|she IAUX|do-d**
%syn: <ip SNP Vf **DNP (<cp [CPR] <ip SNP AUXD 0VP)**

Cases where the comparative form modifies a verb are treated as if there were a further ellipsis inside the comparative: *sneeze more* means *sneeze more times*, *sleep more* means *sleep more time*, etc. The deleted N is not mentioned at %mor but is coded as an NP on the %syn tier.

*XXX: I ran more than she did
%mor: IPROS|I **IT|run-d DCAR|more CPR|than IPROS|she IAUX|do-d**
%syn: <ip SNP **Vf NP (<cp [CPR] <ip SNP AUXD 0VP)**

*XXX: I ran as much as she did
%mor: IPROS|I **IT|run-d DCAR|as+much CPR|as IPROS|she IAUX|do-d**
%syn: <ip SNP **Vf NP (<cp [CPR] <ip SNP AUXD 0VP)**

15. Codes

Here is the list of codes, various combinations are not listed (for example: IAUX|be-s-d for *was* is not listed).

Table 1: Codes for the morphological tier of the I-system , Verbs, and tag questions

Code	*Definition of code*
IA|xxx-s	overt agreement marked on the verb
IT|xxx-ed IT|xxx-d	tense marked on the verb
xxx-en xxx-n xxx-ing	participles (xxx stands for a verb)
IAUX|	Auxiliary, *do* and Modals
TAUX|	auxiliary, modal or *do* in question tags
IPROS|	subject pronoun with overt nominative case
TPROS|	subject pronoun with overt nominative case in tag quesitons
I|to	infinitival 'to'

Tabel 2: Other codes on the morphological trier

Code	*Definition of code*
PRO|	pronoun without overt case, in subject, object position or in tag quesitons
PROO|	object pronoun with overt case

PROE|	expletives
ANAPH|	anaphors
NEG|	negation

Table 3: Codes for the morphological tier of the D-system

Code	*Definition of code*
DART|	article
DPOSD|	possessive determiner (my, his/her, etc.)
DPOS|	independent possessive pronoun (mine, yours, etc.)
DDEM|	demonstrative determiner
D|xxx-pl	(regular) plural noun
D|xxx-p	(irregular) plural noun
D|xxx-'s	possessive affix
DDEM|xxx-p	demonstrative determiner marked for plural
DPRO|xxx-p	plural pronouns ('ones', 'those' etc.; but not 'we' etc.)
DQUA|	basic quantifiers
DCAR|	cardinal quanitifiers

Table 4: Codes for the morphological tier of the C-system

Code	*Definition of code*
CCOM|	embedding complementizer
CWH|	WH-words in quesitons
CADJ|	adjunct complementizer (e.g. *because*)
CREL|	relativizer
CAUX|	inverted auxiliary (e.g. in questions)
CPR|	*than*, *as* in comparative elliptical sentences

Table 5: Codes for the syntactic tier

Code	*Definition of code*
SNP	subject noun phrase
DNP	Direct or oblique object noun phrase
INP	indirect object noun phrase[MK1]
ONP	object of a preposition in stranding or in case preposition is left out (by mistake)
PNP	Predicative nominal
PPP	Predicative prepositional phrase
PAP	Predicative adjective phrase
TXNP	Dislocated NP. X stands for the grammatical function (S, D, I, A, P or O)
XWHE	Extracted WH-phrases (long distance extraction)
EXP	expletives
AUX	auxiliary
TAG	auxiliary in tag questions
V (V-en, V-n, Vf, V-ing)	main verb
PRT	particle
AP	adjectival phrase

PP	prepositional phrase
ADVP	adverbial phrase
NEG	negation
CCOM	complementizer introducing complement clause
CADJ	complementizer introducing adjunct clause
XWH	WH-word (X is S, D, I, A, P or O,) in main clauses (questions)
XWHREL	WH relativizer (X is S, D, I, A, P or O,)
XREL	relativezer *that* (X is S, D, I, A, P or O,)
XWHE	extracted phrase (X is S, D, I, A, P or O,)
XMDF	modifier (X is S, D, I , Por O)
SC	small clause
TO	infinitival complements
PC	gerunds
CPR	*than, as* in comparative elliptical sentences

Table 6: Codes for the lexical tier

Code	*Definition of code*
N|	noun
V|	verb
ADV|	adverb
ADJ|	adjective

Table 7: Codes for mistakes on the morphological trier

Code	*Definition of code*
!	misselection
=	added extra element
=0	missing element
Vn	missing morphology on the verb

Table 6 Codes for mistakes on the syntactical trier

Code	*Definition of code*
=	added XP
=0	missing XP

=L	material that is ungrammatically low
+H	material that is ungrammatically high

Page: 62

[MK1]In my notes, I also see an O, as in ONP and PROO, which handles the object of a preposition in stranding or cases where P is removed (typically an error) but the NP remains.