

2.2. Interaction between category and function tags

Each potential verb argument in the Penn Treebank is annotated with a categorial tag (NP, PP, etc.), along with one or more function tags.

On the one hand, the role of a given function tag varies depending on which categorial tag it is associated with. For instance NP-CLR tends to denote a *fixed phrase* such as “took place” in (2), whereas PP-CLR tends to denote a compulsory argument (see (4) below).

```
(2) (S (NP-SBJ (PRP$ their)
          (NN meeting)
        (ADVP-TMP (RB never) )
        (VP (VBD took)
            (NP-CLR (NN place))))))
```

Hence it is not sufficient to consider function tags alone when retrieving the arguments of a verb. One must also examine the interaction between categorial and function tags.

Moreover, function tags interact among themselves. For example, the function tag LOC, for “location”, tends to denote an adjunct when combined to the tag for prepositional phrases (PP-LOC). But when the functional tag PRD (for “predicative”) is also present, then the judgement definitively switches to argumenthood, with the constituent marked PRD having a predicative syntactic function. So, in (3), “in the front row” is a predicative argument of “was”. In fact, the function PRD exhibits a rather simple interaction with other function tags. It can simply be viewed as dominating the whole sequence: when a tag PRD is present, the constituent is a predicative argument.

```
(3) (S (NP-SBJ (PRP$ his)
          (NN family))
      (VP (VBD was)
          (PP-LOC-PRD (IN in)
                      (NP (DT the)
                          (NN front)
                          (NN row))))))
```

But things are not always that simple. Let us consider for instance the interaction between the function tags CLR and LOC. The combination PP-CLR tends to denote an obligatory argument¹, whereas the sequence PP-LOC tends to denote an adjunct. Both cases appear in (4).

```
(4) (S (NP ``PS of New Hampshire shares``)
      (VP (VBD closed)
          (NP-TMP (NN yesterday) )
          (PP-CLR (IN at)
                  (NP (NP ($ $)
                      (CD 3.75)
                      (-NONE- *U*))
                  ...))
          (PP-LOC (IN in)
                  (NP (NNP New)
                      (NNP York)
                      (NNP Stock)
                      (NNP Exchange)
                      (JJ composite)
                      (NN trading))))))
```

¹But see Section 4.2 for PP introduced by the preposition “as”.

However, the combination PP-LOC-CLR, denotes neither a compulsory argument, nor an adjunct. Rather, PP-LOC-CLR tends to denote an optional argument, as in (5). So neither LOC, nor CLR seems to dominate a sequence of function tags. Rather, each contributes to some extent to the final judgment on argumenthood.

```
(5) (S (NP-SBJ-1 (NNP Avalon)
            (CC and)
            (NNS others))
      (VP (VBD invested)
          (NP (QP ($ $)
                (CD 14)
                (CD million))
            (-NONE- *U*))
          (PP-LOC-CLR (IN in)
                     (NP (NNP Athena)
                         (NNP Neurosciences)
                         (NNP Inc.))
                     ...))
```

Hence, in order to retrieve verb subcategorization frames, one must not only examine the interaction between categorial tags and function tags, but also the (sometimes complex) interaction between all the possible combinations of function tags.

2.3. Canonical subcategorization and syntactic alternations

Deciding whether a given sequence is a compulsory or an optional argument is delicate and often depends on the surface realization of a verb and of its arguments and on the potential syntactic alternations that the construction has undergone. For instance, “(PP-LOC-CLR in Ohio)”, in a clause in the active voice such as in (6), seems to be an optional argument, but it looks more like a compulsory argument when the clause is in the passive voice, as in (7), since (8) is ungrammatical. In what follows, the subcategorization of a verb is examined for the *canonical* realization of its arguments. The interaction between argumenthood constraints and syntactic alternations (e.g., passive) is beyond the scope of this work.

(6) The company based its headquarters in Ohio.

(7) The headquarters of the company are based in Ohio.

(8) * The headquarters of the company are based.

2.4. Optional argumenthood

Existing rules for determining argumenthood in the PTB ignore the notion of *optional argument*: in (9), “These Japanese companies” is clearly an optional argument of the verb “provide”.

(9) But the different business system would make it hard for Dentsu to ...
 (VP (VB provide)
 (NP-BNF these Japanese companies)
 (NP the same kind of services
 they do in Japan))

3. Overview of our Subcategorization Extraction tool

3.1. Methodology

In order to develop this tool, we first had to define which verbs in the PTB should receive a subcategorization: these are the verbs that are dominated by a verbal phrase (VP). This statement, which seems at first very trivial, allows one to avoid assigning a subcategorization frame to verbs which appear inside a noun phrase, such as in “(NP the/DT Dutch/NNP publishing/VBG group/NN)”. It also prevents one from assigning a subcategory to auxiliary verbs and to modals. A distinctive case are the gerunds and past participles when introducing a prepositional phrase, such as “including” in (10), generally regarded as prepositions, although they seem to preserve selectional properties from their usage as a verb we currently do not assign them a subcategorization frame either²

```
(10) (S (NP-SBJ Four ... surviving workers)
      (VP (VBP have)
          (NP asbestos-related diseases)
          (, ,)
          (PP (VBG including)
              (NP three with ... diagnosed cancer))))
```

For each of the verbs which receive a subcategory, subject extraction is easy since it is marked by the tag SBJ. As for other potential complements (i.e. all the right siblings of a verb inside a VP), in order to develop fine-grained rules, we examined by hand the 150+ different sequences of syntactic tags in the PTB that appear assigned to a right sibling of a verb inside a verbal phrase (e.g., “NP”, “NP-TMP”, “NP-TMP-CLR”, etc.).

For each of these sequences, we established whether the sequence was assigned to a modifier, an argument, or optional argument, occasionally using some additional structural contextual information to help disambiguate the particular sequence. When the tag was overwhelmingly assigned to an argument (optional or compulsory), we also mapped it to one of the following traditional syntactic functions:

- Subject: as in “(NP-SBJ John) eats an apple.”
- Predicative: “John is (NP-PRD a hockey player).”
- Direct-Object: “John sees (NP-DirObj Mary).”
- Second-Object: used for “dative objects” or “benefactives”, e.g., “John gave (NP-SecondObj Mary) a book.”
- LocDirObject: used for elements that are arguments, despite their “locational” or “directional” nature, as in “The plane goes (PP-LocDirObj from Denver) (PP-LocDirObj to Dallas).” (cf. the ungrammaticality of “* The plane goes”).
- Indirect-Object: used for other prepositional arguments, as in “John thinks (PP-IndirectObj about Mary).”

²Notice this is not dissimilar to other genuine cases of selectional relation across categories, e.g., nouns that take arguments often have the same or related selectional restrictions as their verb counterparts (cf. “account”).

Currently, we assign syntactic functions essentially to the constituents of the following categories: NP, PP, ADVP, ADJP. Syntactic functions are not assigned yet to sentential complements (S and SBAR), except for the functions **subject** and **predicative**. Lexical co-anchors are not assigned syntactic functions either (e.g. prepositions or particles selected by a verb, elements of a fixed phrase such as “hand” in “go hand in hand”).

The output we obtain may be partially or totally lexicalized: e.g., the program indicates which particular preposition introduces a PP complement.

The work we present just aims at adding information one would expect to find in the Penn Treebank. There seems to be no easy way to perform a serious evaluation of our result. One reason is the lack of annotated material to compare our output with. Existing subcategorization lexicons for English verbs tend not to establish the distinction between optional and compulsory arguments. Moreover, mapping subcategorization frames from one source (e.g. a dictionary) to another (e.g. our output) for evaluation purpose is a non trivial task if one wants to provide reliable and meaningful evaluation figures.³

The main problem with the evaluation, however, which was clear from a manual examination of the output we have performed, is that the boundary between adjuncthood and argumenthood is very fuzzy. Hence some constituent that our program has marked as obligatory argument could be judged as being optional argument or even adjunct, with different degrees of confidence. One such example is provided in (11), concerning the role of the PP “in its Micronite cigarette filters”, that our tool extracted as an optional argument. In general, we preferred to favor argumenthood over adjuncthood when in doubt. Our motivation is that it is potentially easier for users to get rid of information they judge superfluous, rather than to reconstruct information that is missing.

(11) Lorillard Inc. , the unit of New York-based Loews Corp. that makes Kent cigarettes , stopped using crocidolite in its Micronite cigarette filters in 1956.

Another source of errors in the output is due to original annotation errors. When possible, we have tried to have the program “bypass” some recurring annotation errors. For example, when the complementizer “that” is incorrectly tagged as a determiner, a sentential complement introduced by “that” is nonetheless generally correctly identified as long as it is inside an S or an SBAR constituent. Particles for particle verbs are generally correctly identified as such, even when they are not marked PRT. In a few cases, the original annotation errors could not be detected though. One such case is when a verb has been mistagged as an adjective. The program then quite obviously fails to identify it and assign a subcategorization frame to it.

3.2. Output samples

The program may output a verb lexicon where each verb is associated a subcategorization frame indicating ar-

³However, the extended treebank as well as the annotation tool itself will be freely available to the research community, and we would appreciate to have feedback.

(15) Newsweek ,trying to keep pace with rival Time magazine , announced new advertising rates for 1990 ...

(16) #####
 #VB: keep
 #Subj: NP-SBJ
 #Arguments: NP-CLR(pace)//PP-CLR#IndirectObj(with)//
 #####

4.2. Prepositional Phrases

Whenever a prepositional phrase is considered to be an argument of a verb, the preposition selected by the verb is also retrieved by scanning the treebank tree, and is noted as a “co-anchor”.

The following tag sequences denote compulsory arguments: PP-CLR-LOC, PP-DIR-CLR, PP-LOC-PRD, PP-PRD, PP-PRD-LOC, PP-TMP-PRD, PP-PUT.

The following tag sequences denote optional arguments: PP-BNF, PP-DTV, PP-DIR, PP-EXT, PP-LOC-CLR.

The following tag sequences denote adjuncts: PP-LOC, PP-MNR, PP-PRP, PP-TMP.

In addition, some PP had to be treated with more fine-grained structural rules. For instance, the sequence PP-CLR denotes an optional argument when it is introduced by the preposition *as*, and denotes a compulsory argument when it is introduced by another preposition.

Similarly, bare PPs, that is, those marked only “PP” without any other function tags, are hard to account for. After careful observation of the data, we concluded that:

- A PP occurring right after a comma is an adjunct.
- A PP introduced by the preposition “by” co-occurring with a verb marked passive is not argument.⁵
- A PP introduced by the preposition “by”, with a verb not marked passive, and whose immediate left sibling is marked PRD is an adjunct.
- All other bare PPs are compulsory arguments, and the preposition is retrieved as a co-anchor.

Examples: “*By today’s standards*” in (17) is introduced by *by* and preceded by a predicative, hence adjunct; “*By roughly half*” in (18) is a bare PP argument of *shrink*.

(17) (S (NP-SBJ (DT The) (NNS computers))
 (VP (VBD were)
 (ADJP-PRD (JJ crude))
 (PP (IN by)
 (NP (NP (NN today) (POS 's))
 (NNS standards))))))

(18) (VP (VBD shrank)
 (PP (IN by)
 (NP (RB roughly)
 (DT half)))

⁵It is not adjunct either, since it is the logical subject of a passive with Agent, for which the annotation LGS has been forgotten. So the extraction program simply ignores it.

4.3. Adjectival Phrases

The following sequences of tags denote a compulsory argument: ADJP-CLR, ADJP-PRD, ADJP-MNR.

Bare ADJPs, that is, ADJPs with no other function tags, are adjuncts when they are immediately preceded by a comma, and compulsory argument otherwise.

All the other sequences of tags occurring with an ADJP are adjuncts (e.g., ADJP-LOC, ADJP-ADV).

Example: “private” is an argument in (19).

(19) Of course it’s better...
 (S-1 (NP-SBJ (-NONE- *))
 (VP (TO to)
 (VP (VB sell)
 (ADJP-MNR (JJ private))))))

4.4. Adverbial Phrases

The following sequences of tags denote compulsory arguments: ADVP-LOC-CLR, ADVP-PRD-LOC, ADVP-LOC-PRD, ADVP-PRD, ADVP-TMP-PRD, ADVP-TMP-CLR, ADVP-MNR-CLR.

ADVP-EX denotes optional argument.

The following sequences of tags denote adjuncts: ADVP, ADVP-LOC, ADVP-MNR, ADVP-PRP, ADVP-TMP, ADVP-LOC-TMP, ADVP-LOC-TPC, ADVP-MNR-TMP, ADVP-PRD-TMP, ADVP-CLR-MNR.

Finally, for the following sequence of tags, one needs to scan the treebank to find a lexical co-anchor: ADVP-CLR, ADVP-DIR, ADVP-DIR-CLR, ADVP-PRD-LOC, ADVP-PUT, ADVP—PRT.

Example: “recklessly” is a compulsory argument of *behave* in (20).

(20) (SBAR (RB even)
 (IN if)
 (S (NP-SBJ (PRP they))
 (VP (VBD behaved)
 (ADVP-MNR-CLR (RB recklessly))))))
 ...

4.5. Particles

Particles are marked as lexical co-anchors. It is also worthy to note that a lot of annotation errors and incoherences are encountered with particles: some are correctly marked RP (e.g. “(VP (VBP show) (PRT (RP up)))”), but some are annotated IN, (e.g. “(VP (VBG lying) (PRT(IN down)))”), others are marked RB (e.g. “(VP (VB Sit) (PRT (RB down)))”). In some cases, the particle is marked RP, but the constituent PRT is not present. In other cases, neither the PRT, nor the RP are marked (e.g., “(VBN turned) (ADVP-DIR (RB away))”). Hence, particles have been retrieved with great care, not only with rules specifically aiming at indicators of particles (i.e. PRT and RP), but also with rules detecting co-anchors in adverbial phrases, etc. (21) shows an example of particle extraction which is not next to the verb.

(21) Filling out detailed forms about these individuals...
 (VP (MD would)
 (VP (VP (VB tip)
 (NP (DT the) (NNP IRS))
 (PRT (RP off)))

4.6. Sentential complements : S and SBAR

The following tag sequences denote adjuncts: S-ADV, SBAR-ADV, SBAR-LOC, SBAR-MNR, SBAR-PRP, SBAR-TMP, S-PRP, S-MNR.

The following tag sequences denote compulsory verb arguments: SBAR-PRD, S-PRD, SBAR-CLR, SBAR, SBAR-LOC-PRD, S-PRP-CLR, SBARQ, S-CLF, SQ.

A reasonable characterization of a sentential complement requires more information than for other categories, since the verb extends its selectional requirements to aspects of the internal structure of the sentential complement, such as the complementizer that introduces the clause (if any), or whether the complement is to be realized as a finite or non-finite clause, whether a non-finite clause is to be introduced by “to” or not, etc. In the following examples we summarize the kind of information provided by our tool. We recover the complementizer introducing the sentential complement: in (22), the sentential complement of “worry” is extracted as “*SBAR(whether)*”, indicating the selection for the complementizer “whether”. Selection for the complementizer “that” may occur in two forms: either the complementizer is overtly realized, as in (23), in which case the sentential complement of “argue” is extracted as “*SBAR-PRD(that)*”, indicating the selection for an overt complementizer “that”; or covert as in (24), i.e., the complementizer is not overtly realized (expressed in the annotation by the null element “(-NONE- 0)”). In the latter case, the complementizer “that” is enclosed in brackets in the reported subcategorization expression, indicating its optionality. So, the internal complement of “explained” in (24) is reported as “*SBAR([that])*”. In (25) the complement of “is” is given as “*S-PRD(VPinf(to))*”, indicating the selection for an “to”-infinitival sentential complement. An important case is depicted through the example in (26). The complement of “wait” is reported as “*SBAR(PP(for) VPinf(to))*”, which summarizes the selection for a structure under the SBAR that resembles a PP introduced by “for”, followed by a VP introduced by “to”.

- (22) You do n't even have
 (S (NP-SBJ (-NONE- *-1))
 (VP (TO to)
 (VP (VB worry)
 (SBAR (IN whether)
 (S (NP-SBJ (DT the) (NN ad))
 (VP (VBZ is)
 (ADJP-PRD (JJ truthful)))))))))
- (23) The finding probably will support those who ...
 (S (NP-SBJ (-NONE- *T*-6))
 (VP (VBP argue)
 (SBAR (IN that)
 (S (NP-SBJ (DT the) (NNP U.S.))
 (VP (MD should)
 (VP (VB regulate)
 (NP (NP (DT the) (NN class))
 (PP (IN of)
 (NP (NN asbestos))))
 ...)))))))))
- (24) (S (S-TPC-2 (NP-SBJ-1 More common chrysotile fibers)
 (VP (VP (VBP are)
 (ADJP-PRD (JJ curly)))
 (CC and)
 ...
 (NP-SBJ (NNP Dr.) (NNP Mossman))
 (VP (VBD explained)
 (SBAR (-NONE- 0)
 (S (-NONE- *T*-2))))

- (25) (S (NP-SBJ-1 (NP (NNP Mr.) (NNP Lane) (POS 's))
 (JJ final) (NN purpose))
 (VP (VBZ is) (RB n't)
 (S-PRD (NP-SBJ (-NONE- *-1))
 (VP (TO to)
 (VP (VB glamorize)
 (NP (NP (DT the)
 (NNP Artist) (POS 's))
 (NN vagabond)
 (NN existence)))))))))
- (26) (S (NP-SBJ (DT That))
 (VP (VBZ permits)
 (S (NP-SBJ (PRP them))
 (VP (TO to)
 (VP (VB wait)
 (SBAR (IN for)
 (S (NP-SBJ (NNS prices))
 (VP (TO to)
 (VP (VB rise))))))))))

4.7. A word on coordination

Standard coordination is simply marked by the embedding of two constituents of the same category. When VPs are coordinated, as illustrated in (27), they share the same subject but each verb will have its own list of internal arguments. In the example both verbs would have “*NP-SBJ*” as subject, whereas for internal arguments, “vary” would take “*NP-DirObj*”, and “go” would take “*PP-CLR(after)*”. V coordination may appear annotated in two ways: it may be annotated as VP coordination, with argument labels dominating traces, in which case the process is as mentioned for VP coordination; or it may appear as in (28). In this case we can safely assign the same subcategorization verb to all coordinated verbs.

- (27) (S (NP-SBJ (NN portfolio) (NNS managers))
 (VP (MD can)
 (VP (VP (VB vary)
 (NP (NNS maturities))
 (CC and)
 (VP (VB go)
 (PP-CLR (IN after)
 (NP the highest rates))))))
- (28) They could ...
 (VP (VB accept)
 (CC and)
 (VB hire)
 (NP (JJR more)
 (NN labor))
 (PP-DIR-CLR (IN from)
 (PP (JJ outside)
 (NP (NNP Japan))))))

When one deals with the coordination not of VPs but of some internal argument of the verb, one may have to scan each member of the coordination for potential co-anchors, even in the case where the members are of the same category. (29) illustrates that point : a coordination of PPs is argument of the verb. Each PP has to be scanned independently in order to find the prepositions introducing each of them i.e., resulting in two reported frames: one in which the PP argument is reported as “*PP-IndirectObj(around)*”, the other in which it is reported as “*PP-IndirectObj(at)*”.

- (29) (VP (VBZ puts)
 (NP (DT the) (NN dollar))
 (PP-PUT (PP (IN around)
 (NP (CD 1.8200)
 (NNS marks))
 (PP-TMP at the end of November)
 (CC and)
 (PP (IN at)
 (NP (CD 141.33) (NNS yen))))))

The problem is more complex when the coordinated constituents are of different category (marked UCP in the PTB terminology, for “Unlike Coordinated Phrase”), as in (30), where the two alternative internal arguments would be “NP-DirObj” and “SBAR(that)”.

```
(30) (S (NP-SBJ (NN Compound) (NNS yields))
      (VP (VBP assume)
          (UCP (NP (NP (NN reinvestment) )
                  (PP (IN of)
                      (NP (NNS dividends))))
              (CC and)
              (SBAR (IN that)
                    (S (NP-SBJ (DT the)
                              (JJ current)
                              (NN yield))
                      (VP continues for a year))))))
```

One last problem is that coordinations, while mostly perceived as distributive, may occasionally have to be interpreted in a cumulative way. For instance in a sentence such as “John gives used appliances and to local charities” the correct analysis would assign two internal arguments to “gives” under a single subcategorization frame (as opposed to two frames as the previous examples). We do not handle such cumulative analyses. (31) shows a rare such example from the PTB, for which our current analysis fails.

```
(31) (S (NP-SBJ (DT The) (NN turf) )
      (ADVP-TMP (RB recently) )
      (VP (VBZ has)
          (VP (VBN ranged)
              (PP-CLR (PP (IN from)
                        (NP (NNP Chile) ))
                    (PP (TO to)
                        (NP (NNP Austria) ))
                    (PP (TO to)
                        (NP (NNP Portugal) ))))))))
```

4.8. Assigning a syntactic function

The function **subject** is assigned to any sequence, marked SBJ in the treebank ⁶.

The function **predicative** is assigned to any sequence marked PRD (NP-PRD, ADVP-PRD, S-PRD, SBAR-PRD, NP-TTL-PRD, etc.). It is also assigned to the following tag sequences: ADJP-CLR (as in “It ended (ADJP-CLR-Predicative unchanged)”), ADVP-MNR-CLR (“behave recklessly”), ADVP-TMP-CLR, ADVP-CLR-MNR and NP-ADV.

The function **direct object** is assigned essentially to NPs, more specifically to: NP, NP-TTL (“Have you read (NP-TTL-DirObj ‘killing a mockingbird’)”). It is also assigned to ADJP (“This means (ADJP-DirObj confrontational)”) and ADJP-MNR (“to sell (ADJP-MNR-DirObj private)”).

The function **LocDirObj** is assigned essentially to PPs, more specifically to: PP-DIR (“go from X to Y”), PP-EXT (“X grew by 12%”), PP-CLR-LOC (“X arrived (PP-CLR-LOC-LocDirObj at the station)”), PP-DIR-CLR, and PP-LOC-CLR (“locate X (PP-LOC-CLR-LocDirObj in the station)”). It is also assigned to ADJP-ADV (“This machine operates (ADJP-ADV-LocDirObj fully loaded)”), ADVP-EXT (“The average number of options traded has surged (ADVP-EXT-LocDirObj nearly tenfolds)”), ADVP-LOC-CLR, NP-TMP-CLR (“It took (NP-TMP-CLR 10

years) to ...”)⁷, NP-EXT (“X grew (NP-EXT-LocDirObj 12 %)”), NP-LOC-CLR (“X lived (NP-LOC-CLR-LocDirObj there)”).

The function **Second Object** is assigned to any tag sequence bearing a tag -BNF- or -DTV- (i.e. PP-BNF, PP-DTV, NP-BNF).

The function **Indirect Object** is essentially assigned to PPs, more specifically: PP-CLR (“talk (PP-CLR about X)”), PP-PUT and PP.

The following non sentential tag sequences are not assigned a syntactic function because they denote either a lexical co-anchor or part of a fixed phrase: ADVP-CLR (“looking ADVP-CLR(forward) PP-CLR(to)”), ADVP-DIR, ADVP-DIR-CLR, ADVP-PUT, ADVP—PRT, NP-CLR (“take effect”), and NP-MNR-CLR.

These assignment of syntactic functions work fairly well in practice, but could:

1. be further refined;
2. be extended to sentential complements.

5. Conclusion

We have presented a set of sufficiently general and fine-grained rules as well as an implementation which will be reusable and freely available to the research community to extract verb arguments from the Penn-Treebank. We hope that these rules will be enriched over time, and that more rules will be added in order to:

- Extend the assignment of syntactic functions to sentential complements
- Extract non verbal arguments (e.g. noun arguments)

This implementation is currently used in work on grammar extraction. In addition to this rule-based approach to annotate verb arguments in the PTB, we are also currently investigating the use of a decision-tree approach (which could prove useful to functionally annotate unseen sequences such as the output of a parser).

6. References

- J. Chen and K. Vijay-Shanker. 2000. Automated extraction of TAGs from the Penn Treebank. In *IWPT-00*.
- M. Collins. 1997. Three generative, lexicalised models for statistical parsing. In *ACL-97*, Madrid.
- M. Marcus, B. Santorini, and M.A. Marcinkiewicz. 1993. Building a large annotated corpus of english: the Penn Treebank. In *Computational Linguistics*, 19:2.
- M. Marcus, G. Kim, M.A. Marcinkiewicz, R. MacIntyre, A. Bies, M. Ferguson, K. Katz, and B. Schusberger. 1994. The Penn Treebank: Annotating predicate-argument structure. In *Proceedings of the DARPA Human Language Technology Workshop*.
- F. Xia. 2001. *Automatic grammar generation from two perspectives*. Ph.D. thesis, Univ. of Pennsylvania.

⁶Only one subject is marked for each verb, so subject coordinations, traces etc. are ignored if the verb in question has already been assigned a subject

⁷These constructions are not direct object because they do not passivize.