

Wh-Questions Tgrep2 Corpus Annotation Guide

Morgan Moyer

October 17, 2020

1 Macros

1. **Disfluencies @ DISFL:**
/EDITED|UH|PRN|-UNF/
2. **Wh-Node @ WH:**

```
/WP|WRB|WDT/  
!< / ^t| ^T/  
!>> @DISFL  
>> / ^S/  
!> / ^N| ^A/
```

2 Categorical Variables

1. **WhPhraseType:**
Does the Wh-head have a sister? (e.g., which children); ‘monomorphemic’ otherwise (e.g., who).

(a) Complex: @WH \$ / ^N| ^D| ^J| ^RB| ^RP| ^PP/
(b) Monomorphemic: @WH !\$ / ^N| ^J| ^RB| ^PP/
2. **ModalPresent:** (yes/no): is there a modal auxiliary (can, could, shall, should, may, might, must) in the question?

(a) Yes: @WH . /MD/
(b) No: @WH !. /MD/
3. **Finite** (yes/no): Does the question have a tensed (non-modal) verb? If yes = Finite, if not= Infinitive.

(a) Finite: @WH [!. /MD/ | [!. (/ ^-NONE-/ . /TO/)]
(b) Infinitival: @WH . (/ ^-NONE-/ . /TO/)
4. **QuestionType:** what kind of syntactic construction is the question appearing in? The most important ones are “Embedded” and “Root”

(a) Embedded: @WH
>> (/ ^SBAR/ [\$ / ^VB| ^JJ/] | [> / ^PP/ \$ / ^JJ| ^VB/])
!>> (/SBAR-NOM/ [!\$ / ^V |JJ/] | [!> (/ ^PP/ \$ / ^VB| ^JJ/)])

- (b) Root: @WH
 >> (/^SQ|^SBARQ/ [!\$ / ^VB/ | !> / ^N/])
 !>> /SBAR-NOM/
- (c) Relative: @WH
 >> (/^S/ [> / ^N/ | > (/^P/ > / ^N/)])
 !>> (/SBAR-NOM/ [\$ / ^VB|^JJ/] | [> (/^PP/ \$ / ^VB|^JJ/)])
- (d) embadjunct: @WH
 >> (/^SBAR/ \$ / ^VP|^AD/ !\$ / ^VB|^JJ/)
- (e) adjunct: @WH
 [>> (/^SBAR/ [> (/^PP/ !\$ / ^VB|^JJ/)] | [!\$ / ^VB|^JJ/)])
- (f) nominalized: @WH
 >> (/SBAR-NOM/ [!\$ / ^VB|^JJ/] | [!> (/^PP/ \$ / ^VB|^JJ/)])
- (g) subject: @WH
 [>> /SBAR-SBJ/ | >> (/^S/ >> /TOP/)]
- (h) fragment: @WH
 >> (/^SBAR/ > /FRAG/)
- (i) exclam: @WH
 >> (/^S/ !<< / ^V/)
 !>> / ^SBAR-NOM/

5. DegreeQ

- (a) Yes: @WH \$ / ^JJ|^RB/
- (b) No: @WH !\$ / ^JJ|^RB/

6. WhAll

- (a) Yes: @WH >> (/^WH/ << (/^NP/ < / ^DT/ !< / ^N/))

7. SubjectAux:

- (a) Yes: @WH >> (/^SBAR/ << / ^SQ/)

3 String Variables

1. MatVerb

@WH >> (/^S/ \$ / ^VB/=print)

2. Wh

@WH < / ^w|^h/=print

3. Modal

@WH . /MD/=print

4. Verb1

@WH >> (/^S/ [<< / ^VP/ << / ^VB/=print])

5. Verb2

@WH >> (/^S/ [<< / ^VP/ [< / ^VP/ < / ^VB/=print]])

6. **DeterminerSubject**
@WH >> (/^S/ << (/DT/=print >> /^NP-SBJ/))
7. **DeterminerNonSubject**
@WH >> (/^S/ << (/DT/=print >> /^VP/))
8. **FullWhPhrase**
@WH > /^W/=print \$ /^N|^J|^RB/
9. **WhParse**
10. **Sentence**
@WH >> (*=print !_i *)
11. **Question**
@WH >> /^S/=print
12. **QuantifiedPredicate**
@WH
>> (/^WH/ > (/^S/ << (/^VP/ << /all|every|some|one|any|most|least|more|most|much/=print)))
13. **QuantifiedSubject**
@WH
>> (/^WH/ > (/^S/ << (/^NP-SBJ/ << /all|every|some|one|any|most|least|more|most|much/=print)))