Reflections
MMT strategy
Lab 7 phenomena

Ling 567
May 14, 2019

Overview

- General reflections
- MMT strategy
- Coordination
- Wh questions
- Definiteness/cognitive status

Reflections

- How does working with this system differ from what we were doing in 566?
- What have you learned about grammar engineering?
- What have you learned about linguistic structure (morphology/syntax/semantics)?
- What have you learned about incremental development?
- What would you say are best practices for developing these grammars?

Overview

- General reflections
- MMT strategy
- Coordination
- Wh questions
- Definiteness/cognitive status

MMT strategy

- Sentences use closed vocabulary in order to focus on broad range of phenomena
- ‘Fake it’ with a lexical interlingua
- If you don’t have a lexical item (e.g. *dog*) pick one that’s close (e.g. *kangaroo*) but give it _dog_n_rel as the PRED value. (And then leave a comment documenting this.)
- English doesn’t do much with aspect; if your language does, pick any convenient aspect value.
- Similarly, you don’t have to worry about definiteness

MMT sentences

Dogs sleep
Dogs chase cars
I chase you
Dogs eat
The dogs chase cars
The dogs dont chase cars
I think that you know that dogs chase cars
I ask whether you know that dogs chase cars
Cats and dogs chase cars
Dogs chase cars and cats chase dogs
Cats chase dogs and sleep
Do cats chase dogs
Hungry dogs eat
Dogs in the park eat
Dogs eat in the park
The dogs are hungry
The dogs are in the park
The dogs are the cats
I asked what the dogs chased
The dog s car sleeps
My dogs sleep
The dog sleeps because the cat sleeps
The dog sleeps after the cat sleeps
Who sleeps
What do the dogs chase
What do you think the dogs chase
Who asked what the dogs chase
I can eat glass
It doesnt hurt me

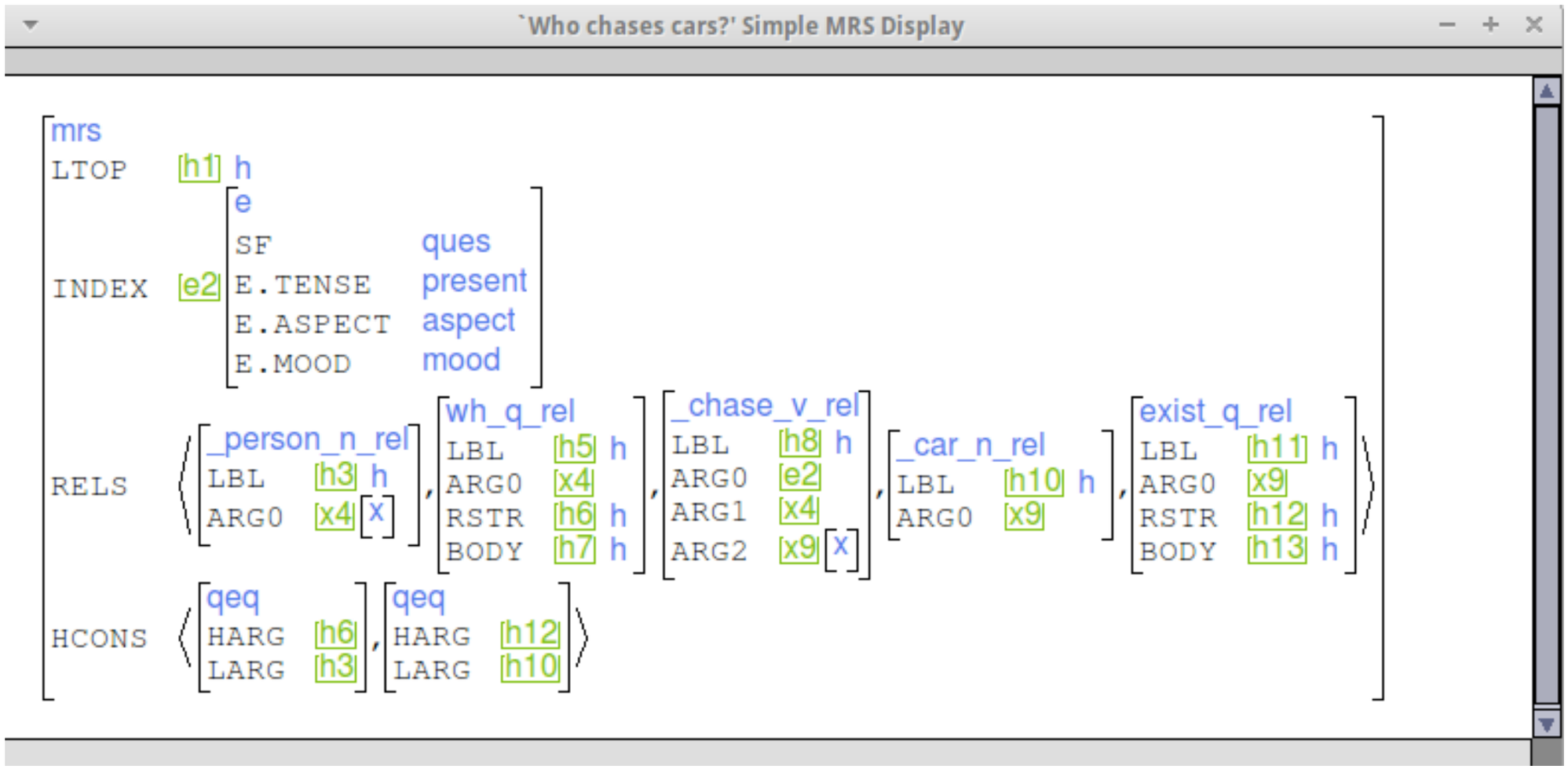
Overview

- General reflections
- MMT strategy
- Coordination
- Wh questions
- Definiteness/cognitive status

Overview

- General reflections
- MMT strategy
- Coordination — visit customization system page
- Wh questions
- Definiteness/cognitive status

Wh-questions: Target representations

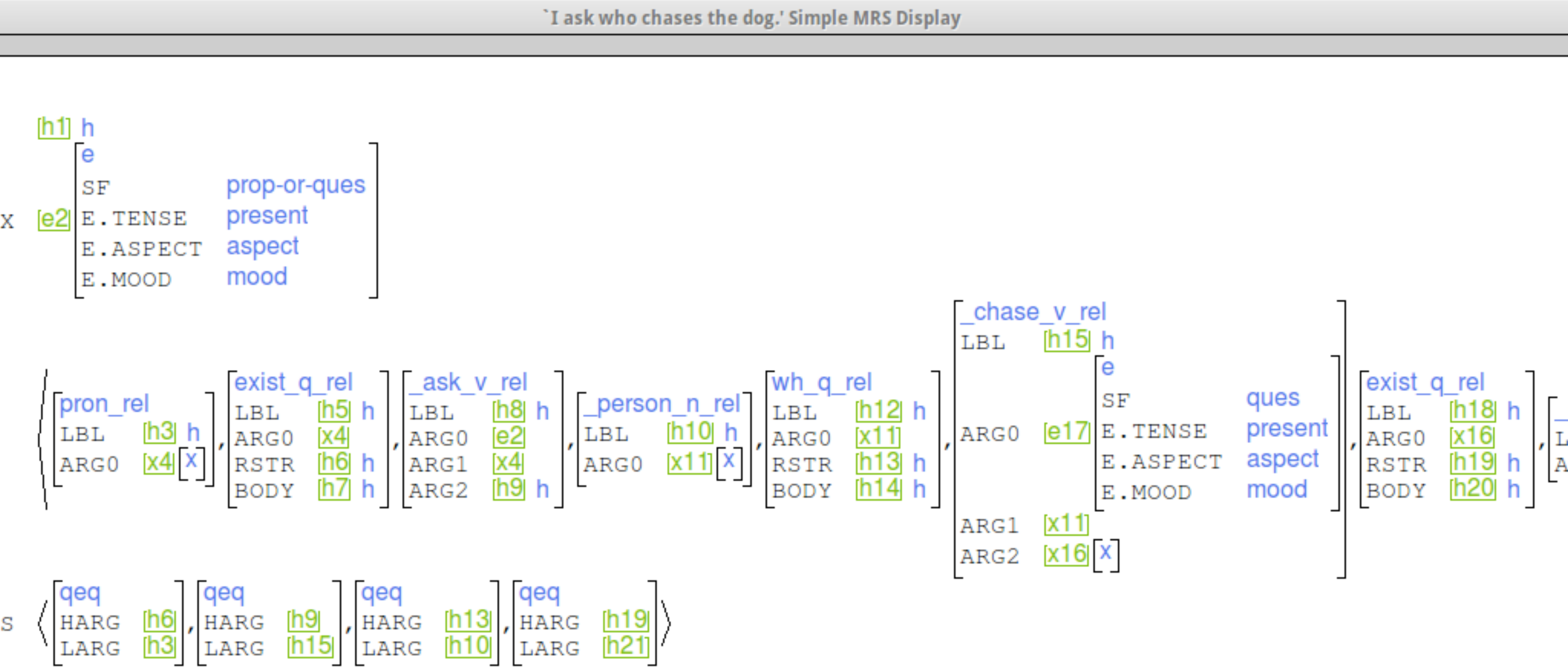


Wh-questions: Target representations

`What do you think the dog chases?' Simple MRS Display

```
mrs
LTOP   [h1] h
INDEX  [e2] [ e
              SF        ques
              E.TENSE   present
              E.ASPECT  aspect
              E.MOOD    mood ]
RELS   <[ _thing_n_rel
          LBL   [h3] h
          ARG0  [x4] [x] ],
        [ wh_q_rel
          LBL   [h5] h
          ARG0  [x4]
          RSTR  [h6] h
          BODY  [h7] h ],
        [ pron_rel
          LBL   [h8] h
          ARG0  [x9] [x] ],
        [ exist_q_rel
          LBL   [h10] h
          ARG0  [x9]
          RSTR  [h11] h
          BODY  [h12] h ],
        [ _think_v_rel
          LBL   [h13] h
          ARG0  [e2]
          ARG1  [x9]
          ARG2  [h14] h ],
        [ exist_q_rel
          LBL   [h15] h
          ARG0  [x16] [x]
          RSTR  [h17] h
          BODY  [h18] h ],
        [ _dog_n_rel
          LBL   [h19] h
          ARG0  [x16] ],
        [ _chase_v_rel
          LBL   [h20] h
          ARG0  [e21] [e]
          ARG1  [x16]
          ARG2  [x4] ]>
HCONS  <[ qeq
          HARG  [h6]
          LARG  [h3] ],
        [ qeq
          HARG  [h11]
          LARG  [h8] ],
        [ qeq
          HARG  [h14]
          LARG  [h20] ],
        [ qeq
          HARG  [h17]
          LARG  [h19] ]>
```

Wh-questions: Target representations



`I ask who chases the dog.' Simple MRS Display

```
h1 h
     [e
      SF          prop-or-ques
x e2  E.TENSE     present
      E.ASPECT    aspect
      E.MOOD      mood        ]

                                                                                                [_chase_v_rel
                                                                                                 LBL   h15 h
                                                                                                           [e
 ⟨[pron_rel     ],[exist_q_rel ],[_ask_v_rel ],[_person_n_rel],[wh_q_rel    ],                  SF         ques     ],[exist_q_rel  ],
   LBL   h3 h     LBL   h5 h     LBL   h8 h     LBL   h10 h     LBL   h12 h                    ARG0  e17 E.TENSE    present    LBL   h18 h
   ARG0  x4 [x]   ARG0  x4       ARG0  e2       ARG0  x11 [x]   ARG0  x11                            E.ASPECT   aspect     ARG0  x16
                  RSTR  h6 h     ARG1  x4                       RSTR  h13 h                          E.MOOD     mood  ]    RSTR  h19 h
                  BODY  h7 h     ARG2  h9 h                     BODY  h14 h                    ARG1  x11                   BODY  h20 h
                                                                                                 ARG2  x16 [x]                     ]

s ⟨[qeq      ],[qeq      ],[qeq      ],[qeq      ]⟩
    HARG h6     HARG h9     HARG h13    HARG h19
    LARG h3     LARG h15    LARG h10    LARG h21
```

The problem

- Wh questions involve long distance dependencies
  - Connecting sentence-initial wh-words to gaps
  - Detecting the presence of in-situ wh-words, and correlating [SF ques]

Solution for English

- SLASH (<= GAP) to encode the absence of an element
  - LDD Top: Head-filler rule
  - LDD Bottom: Comp & subj extraction rules (unary phrase structure rules)
  - LDD Middle: Lexical threading/amalgamation of SLASH; constraints on rules other than head-argument
- QUE to encode the wh property
  - Controls eligibility to be filler in wh question rule
- Initial symbol should require SLASH <! !>

Top: Head-Filler Rule

```
wh-ques-phrase := basic-head-filler-phrase & interrogative-clause &
            head-final &
   [ SYNSEM.LOCAL.CAT [ MC bool,
                 VAL #val,
                 HEAD verb & [ FORM finite ] ],
  HEAD-DTR.SYNSEM.LOCAL.CAT [ MC na,
                 VAL #val & [ SUBJ < >,
                              COMPS < > ] ],
  NON-HEAD-DTR.SYNSEM.NON-LOCAL.QUE <! ref-ind !> ].
```

Bottom: Subject & complement extraction

extracted-comp-phrase := basic-extracted-comp-phrase &
[SYNSEM.LOCAL.CAT.HEAD verb,
HEAD-DTR.SYNSEM.LOCAL.CAT.VAL.SUBJ cons].

extracted-subj-phrase := basic-extracted-subj-phrase &
[SYNSEM.LOCAL.CAT.HEAD verb,
HEAD-DTR.SYNSEM.LOCAL.CAT.VAL.COMPS < >].

Bottom: Subject & complement extraction

```
basic-extracted-comp-phrase := basic-extracted-arg-phrase & head-compositional &
  [ SYNSEM canonical-synsem &
        [ LOCAL.CAT [ VAL [ SUBJ #subj,
                            SPR #spr,
                            COMPS #comps ],
                      MC #mc ] ],
    HEAD-DTR [ SYNSEM
          [ LOCAL.CAT [ VAL [ SUBJ #subj,
                              SPR #spr,
                              COMPS < gap &
                                    [ NON-LOCAL.SLASH #slash ]
                                    . #comps > ],
                        MC #mc ],
           NON-LOCAL.SLASH #slash ] ],
    C-CONT [ RELS <! !>,
                HCONS <! !>,
                ICONS <! !> ] ].
```

Bottom: Subject & complement extraction

```
gap := expressed-non-canonical &
  [ LOCAL #local,
    NON-LOCAL [ REL 0-dlist,
                QUE 0-dlist,
                SLASH 1-dlist &
                    [ LIST < #local > ] ] ].
```

Middle: Lexical threading of SLASH values (cf: GAP Principle, Ling 566)

- Inspired by Bouma, Malouf & Sag 2001
- SLASH on mother comes from head daughter (except in head-adj rules)
- Elements with non-empty ARG-ST build their own SLASH value by appending SLASH of their arguments (basic-one-arg et al)
- Unexpressed arguments are stamped with the type unexpressed:

```
unexpressed := synsem-min &
  [ NON-LOCAL [ SLASH 0-dlist,
                REL 0-dlist,
                QUE 0-dlist ] ].
```

Middle: head-mod phrases

```
basic-head-mod-phrase-simple := head-mod-phrase & binary-headed-phrase &
  [ SYNSEM [ LOCAL.CAT.MKG #mkg,
             NON-LOCAL [ SLASH [ LIST #first,
                                 LAST #last ],
                                 REL 0-dlist ] ],
    HEAD-DTR.SYNSEM
          [ ... ,
            NON-LOCAL #nonloc &
                [ SLASH [ LIST #middle,
                          LAST #last ] ],
    NON-HEAD-DTR.SYNSEM
          [ ... ,
            NON-LOCAL [ SLASH [ LIST #first,
                                LAST #middle ],
                        QUE 0-dlist & [ LIST null ] ] ],
    C-CONT [ RELS <! !>, ICONS <! !> ] ].
```

The wh words

- english.tdl

```
wh-pronoun-noun-lex := norm-hook-lex-item & basic-icons-lex-item &
  [ SYNSEM [ LOCAL [ CAT [ HEAD noun,
                           VAL [ SPR < >,
                                 SUBJ < >,
                                 COMPS < >,
                                 SPEC < > ] ],
                     CONT [ RELS <! [ LBL #larg,
                                      ARG0 #ind & ref-ind ],
                                    [ PRED "wh_q_rel",
                                      ARG0 #ind,
                                      RSTR #harg ] !>,
                           HCONS <! [ HARG #harg,
                                      LARG #larg ] !> ] ],
                 NON-LOCAL.QUE <! #ind !> ] ].
```

- lexicon.tdl

```
what := wh-pronoun-noun-lex &
  [ STEM < "what" >,
    SYNSEM.LKEYS.KEYREL.PRED "_thing_n_rel" ].

who := wh-pronoun-noun-lex &
  [ STEM < "who" >,
    SYNSEM.LKEYS.KEYREL.PRED "_person_n_rel" ].
```

Solution for English

- SLASH (<= GAP) to encode the absence of an element
  - LDD Top: Head-filler rule
  - LDD Bottom: Comp & subj extraction rules (unary phrase structure rules)
  - LDD Middle: Lexical threading/amalgamation of SLASH; constraints on rules other than head-argument
- QUE to encode the wh property
  - Controls eligibility to be filler in wh question rule
- Initial symbol should require SLASH <! !>

Solution for pseudo-English

- Psuedo-English mimics a wh-in-situ language; the wh words stay put
- Still need to be able to tell that they’re there
- Don’t need SLASH, will use QUE to detect presence of the wh words
  - => Still need the lexical threading stuff working
  - => Same entries for wh pronouns
- Special rule at the top of the tree to license wh interrogatives
- Initial symbol requires QUE <! !>, to force the wh interrogative rule to apply

wh-int-cl

```
wh-int-cl := clause & head-compositional &  head-only &
  [ SYNSEM [ LOCAL.CAT [ VAL #val,
                         MC bool ],
             NON-LOCAL non-local-none ],
    C-CONT [ RELS <! !>,
             HCONS <! !>,
             HOOK.INDEX.SF ques ],
    HEAD-DTR.SYNSEM [ LOCAL.CAT [ HEAD verb & [ FORM finite ],
                                  VAL #val &
                                    [ SUBJ < >,
                                      COMPS < > ] ],
                      NON-LOCAL [ SLASH <! !>,
                                  REL <! !>,
                                  QUE <! ref-ind !> ] ] ].
```

Ancillary changes

- Constrained the value of NON-LOCAL on the root symbol in roots.tdl to non-local-none.
- Added norm-zero-arg as supertype for adverb-lex.
- Modified bare-np-phrase to inherit from head-valence-phrase as an additional supertype (this copies up the NON-LOCAL value).
- Modified bare-np-phrase to constrain the element of the daughter's non-empty SPR list to be unexpressed

Ancillary changes

- Modified subj-raise-aux to constrain the element of the non-empty SUBJ list of its complement to be unexpressed

- Made the following modifications to the analysis of coordination (as type addenda):

topormid-coord-phrase :+ [SYNSEM.NON-LOCAL #nl,
LCOORD-DTR.SYNSEM.NON-LOCAL #nl,
RCOORD-DTR.SYNSEM.NON-LOCAL #nl].

bottom-coord-phrase :+ [SYNSEM.NON-LOCAL #nl,
NONCONJ-DTR.SYNSEM.NON-LOCAL #nl].

Wh-questions further thoughts

- English does different things for matrix v. embedded, subject v. non-subject questions:
    - Who sleeps?
    - Who did they think sleeps?
    - What does the dog chase?
    - I ask who sleeps.
    - I ask who the dog chases.

Wh-questions further thoughts

- I think our semantic representations aren’t quite capturing all the info yet:
  - Who did you ask if Sandy saw?
  - => Two clauses with [SF ques], but which one does ‘who’ belong to as a question?

Overview

- General reflections
- MMT strategy
- Coordination
- Wh questions
- Definiteness/cognitive status

Discourse status: What's that?

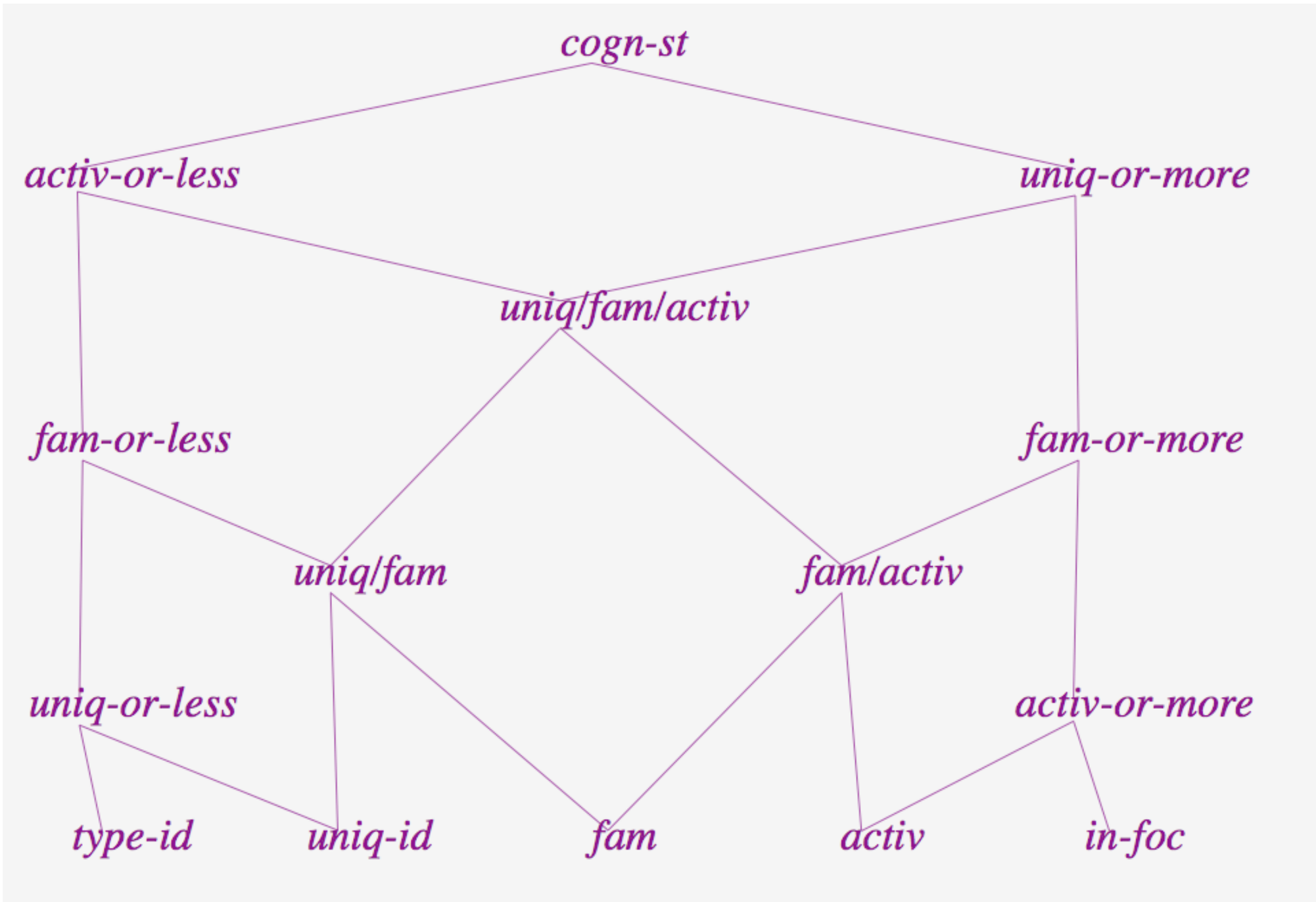
- A property of referents, describing their relationship to the common ground of a conversation
- Tends to be reflected syntactically in markers of "definiteness" as well as demonstratives and constraints on the availability of types of NPs in particular constructions.
- Closely related to (but distinct from) information structure
- The binary distinction "definite"/"indefinite" is not sufficient
- Furthermore, discourse status can be broken down into hearer-oriented "cognitive status" and speaker-oriented "specificity"

Givenness hierarchy (Gundel et al 1993, Prince 1981)

Type id <	Referential <	Uniq. id. <	Familiar <	Activated <	In focus
a N	indefinite *this N*	*the N*	*that N*	*that, this* *this N*	*it*

NB: “In focus” != focus

Borthen & Haugereid's proposal



Borthen & Haugereid's proposal

$$\begin{bmatrix} \text{SYNSEM.LOC.CONT.REF-PROP} & \begin{bmatrix} \textit{ref-prop} & \\ \text{INDEX} & \begin{bmatrix} \textit{ref} & \\ \text{PER} & \textit{per} \\ \text{NUM} & \textit{num} \\ \text{GEND} & \textit{gend} \end{bmatrix} \\ \text{COGN-ST} & \textit{cogn-st} \\ \text{SPECI} & \textit{bool} \\ \text{PART} & \textit{bool} \\ \text{UNIV} & \textit{bool} \end{bmatrix} \end{bmatrix}$$

Borthen & Haugereid's proposal

- SPECI indicates specificity (speaker-oriented)
- Compatible with both "definite" and "indefinite" NPs:
  - *The fastest runner won*.
  - *The next customer will receive a reward*.
  - *I'm looking for a book*.
- Corresponds to overt syntactic phenomena in at least Norwegian (specificity adjectives) and Turkish (accusative case precludes specific interpretation)

Matrix-based proposal

$$\begin{bmatrix} \text{HOOK.INDEX} & \begin{bmatrix} \text{PNG} & \begin{bmatrix} \text{PER} & \textit{person} \\ \text{NUM} & \textit{number} \\ \text{GEND} & \textit{gender} \end{bmatrix} \\ \text{COG-ST} & \textit{cog-st} \\ \text{SPECI} & \textit{bool} \end{bmatrix} \end{bmatrix}$$

Your tasks

- Determine what overt marking of cognitive status and/or specificity occurs in your language
- Constrain the COG-ST value of pronouns, demonstratives, articles (if applicable), and any morphology that relates to cognitive status
- Constrain the COG-ST value of dropped arguments (see next section)