

# UCCA's Foundational Layer: Annotation Guidelines v2.0

<http://www.cs.huji.ac.il/~oabend/ucca.html>

May 30, 2018

## Contents

<b>1</b>	<b>General Principles for Annotation</b>	<b>4</b>
<b>2</b>	<b>A Bird's Eye View of the Categories</b>	<b>4</b>
<b>3</b>	<b>Technical Notes and Guidelines</b>	<b>10</b>
<b>4</b>	<b>Classification of Prepositions</b>	<b>11</b>
<b>5</b>	<b>Participant-Adverbial Distinction</b>	<b>12</b>
<b>6</b>	<b>Analyzability</b>	<b>12</b>
<b>7</b>	<b>Detailed Guidelines</b>	<b>13</b>
7.1	Scenes . . . . .	13
	Annotating Scenes within Scenes. . . . .	13
	Dependent Scenes. . . . .	13
	Distinguishing Ground and Participant Scenes. . . . .	13
	Static Scenes. . . . .	14
	Scene or not a Scene. . . . .	14
	One Scene or two. . . . .	15
	Secondary Verb or Participant Scene. . . . .	15
	Scene within Scene or two Parallel Scenes. . . . .	15
	Verbs that can be primary or secondary. . . . .	16
	Secondary predicates. . . . .	16
	Seem/look/appear/sound/feel. . . . .	16
	Similes. . . . .	17
	Cognitive events . . . . .	17
	Results of Scenes. . . . .	17
	Scene Agents. . . . .	17
	Noun Scenes. . . . .	17
	Scenes with neither a P nor an S. . . . .	17
	Adjective Scenes without a noun. . . . .	18
	Imperatives. . . . .	18
	Fragments. . . . .	18
	Expletive it. . . . .	18
	Cooperating participants. . . . .	18

7.2	Remotes. . . . .	18
	Two Types of verbs that take a participant Scene. . . . .	18
	Prominent Cases of Remote Units. . . . .	19
7.3	Secondary Relations in Scenes. . . . .	19
	Possession. . . . .	20
	Framing of Scenes. . . . .	20
7.4	Non-Scene Units. . . . .	20
	Determiners. . . . .	20
	Distinguishing between Determiners and Prepositions. . . . .	20
	Appositions. . . . .	20
	Extraposition. . . . .	20
	Fused E Scenes. . . . .	21
	Numbers and Quantifiers. . . . .	21
	Single-word Scenes within Elaborators. . . . .	21
	Comparatives/Superlatives. . . . .	21
	Directions. . . . .	21
	Passive “by”. . . . .	22
	Preposition Stranding. . . . .	22
7.5	Processes/States. . . . .	22
	Modals and Auxiliaries. . . . .	22
	Distinguishing identity and other static Scenes. . . . .	22
	Infinitive “to”. . . . .	23
	Light Verbs. . . . .	23
	Possessive “have”: . . . . .	23
	Adjective followed by a Scene: . . . . .	23
	Causatives. . . . .	23
	Secondary Verbs with an additional role. . . . .	24
	Polite Forms. . . . .	24
7.6	Other Relations. . . . .	24
	Punctuation. . . . .	24
	Linkers with a single argument. . . . .	24
	Elaboration of/by a Coordination. . . . .	24
	Vocatives. . . . .	24
	Titles. . . . .	24
	Focus Constructions. . . . .	25
	Question Words. . . . .	25
	Non-contiguous Linkers. . . . .	25
	Dates and Names. . . . .	25
	Reflexives. . . . .	26
	Complex Prepositions. . . . .	26
	Frame of reference. . . . .	26
	Several Coordinated Predicates. . . . .	26
7.7	Morphology. . . . .	26
	Inflectional and Derivational Morphology. . . . .	26
	Coarsed Word/Phrase. . . . .	27

<b>8 Chapter E: Criteria for compound splitting in German</b>	<b>27</b>
Criterion 1: . . . . .	27
Criterion 2: . . . . .	27
Criterion 3: . . . . .	28
<b>9 Chapter F: Possible Post-processing Notes</b>	<b>28</b>
<b>10 Chapter G: Plain Text Notation</b>	<b>28</b>
Non-contiguity: . . . . .	29
Remote Units: . . . . .	29
Implicit Units: . . . . .	29
<b>11 Caveats and Future Work</b>	<b>29</b>

# 1 General Principles for Annotation

NSS: what does UCCA stand for? :)

1. A UCCA annotation task consists of the annotation of multiple sentences, usually a paragraph or several paragraphs long. When you receive a task, take a few minutes to read the entire text, in order to understand the context.
2. UCCA divides the text into units (stretches of text; not necessarily contiguous), each referring to a relation, a participant in a relation or a relation along with its participants. The types of relations we annotate are listed below.
3. The units must cover all the tokens, except punctuation tokens which are not annotated.
4. Units may contain other sub-units, giving rise to a hierarchical structure.
5. Each unit is assigned a category, reflecting its role in a super-ordinate relation in which it participates. The category does not necessarily reflect the meaning of the unit taken in itself. For instance, all the units in **boldface** have the same category, as they all describe “horse” in finer detail:
  - “A **beautiful** horse”
  - “A **police** horse”
  - “A horse **with no name**”
  - “The horse **that won the race**”
  - “A **winning** horse”
6. UCCA does not annotate ambiguity. When reading ambiguous text, decide on the most likely interpretation in your opinion and use it for annotating the entire passage.

## 2 A Bird's Eye View of the Categories

Units may be analyzed according to **one** of the following models:

### Model #1: Scenes

1. A Scene is some description of an action, movement or state (including abstract actions or states). It generally has a time when it happened, a location, and a ground (the circumstances in which the description was uttered or written). In concrete cases, a Scene can be imagined as a single mental image or a short script.
  - “Woody walked in the park” (1 Scene)
  - “I got home and took a shower” (2 Scenes)
2. A Scene has a main relation (exactly one), which determines the type of the Scene and what happened in it. This main relation can be either a STATE (S) – if the Scene is constant in time; or a PROCESS (P) – an action, movement or some other relation that evolves in time.
3. Each Scene is considered a unit, and is therefore, like all units, also assigned a category as a whole. The category of the Scene unit reflects the role of that unit in the super-ordinate relation it participates in (see below).
4. Scenes may contain any number of PARTICIPANTS (A). These are the principal participants in the main relation of the Scene (including locations). Participants may refer either to physical or abstract entities.

- “John<sub>A</sub> boiled [an egg]<sub>A</sub>”
- “Programming<sub>A</sub> is widely taught nowadays”

5. In static Scenes, the main relation is annotated as a STATE. The State unit should not include its auxiliary verbs if present.

- “John<sub>A</sub> is tall<sub>S</sub>”
- “[The apple tree]<sub>A</sub> is in<sub>S</sub> [the garden]<sub>A</sub>”
- “[An apple]<sub>A</sub> weighs<sub>S</sub> [200 g]<sub>A</sub>”
- “[This apple]<sub>A</sub> does weigh<sub>S</sub> [200 g]<sub>A</sub>”
- “John<sub>A</sub> is [a writer]<sub>S</sub>”
- “[Big<sub>S</sub> (dogs)<sub>A</sub>]<sub>E</sub> dogs<sub>C</sub>” **NSS: why is “big” both S and E? OA: it’s an E-Scene (the “dogs” was omitted, so I put it back**

6. In dynamic Scenes, the main relation is marked as a PROCESS (P). The Process unit should not include its auxiliary verbs if present.

- “John<sub>A</sub> kicked<sub>P</sub> [the ball]<sub>A</sub>”
- “John<sub>A</sub> has been kicking<sub>P</sub> [the ball]<sub>A</sub> since breakfast”

7. ADVERBIALS (D) are relations that do not introduce another Scene, but semantically modify the Scene or its PROCESS or STATE. Common cases of Ds are modal relations (like “can”, “may” or “want”), manner relations (like “quickly” or “patiently”) and relations that specify a sub-event (like “begin” or “finish”).

- “John<sub>A</sub> began<sub>D</sub> swimming<sub>P</sub>”
- “John<sub>A</sub> may<sub>D</sub> come<sub>P</sub> [to the party]<sub>A</sub>”
- “[His workers]<sub>A</sub> treat<sub>P</sub> him<sub>A</sub> [with disrespect]<sub>D</sub>”
- “John<sub>A</sub> cleverly<sub>D</sub> answered<sub>P</sub> [the manager’s question]<sub>A</sub>”

See Chapter 5 for how to distinguish Ds and As in edge cases.

8. Units whose primary purpose is to specify the time in which the Scene occurred should be marked as TIME (T). However, if time is described by introducing another Participant or another Scene, it should receive a Scene or Participant category instead. Frequency and duration are also considered TIME.

- “John<sub>A</sub> may<sub>D</sub> come<sub>P</sub> later<sub>T</sub>”
- “John<sub>A</sub> may<sub>D</sub> come<sub>P</sub> [at around eight]<sub>T</sub>”
- “I<sub>A</sub> get<sub>F</sub> treated<sub>P</sub> regularly<sub>T</sub>”
- “[John<sub>A</sub> [showed up]<sub>P</sub>]<sub>H</sub> during<sub>L</sub> [[the<sub>E</sub> filming<sub>C</sub>]<sub>P</sub>]<sub>H</sub>” (two Scenes! see below)

## Model #2: Non-Scene Units

Some relations do not evoke a Scene on their own, and give rise to *Non-Scene units*. We distinguish between several types of such units.

1. CONNECTORS (N) relate two or more entities (annotated as Centers) in a way that highlights the fact that they have a similar type or role. They are usually conjuncts such as the English “and”, “or”, “instead of” or “except” or the German “sowie”, “ebenso” and “genauso wie”.
  - “[John<sub>C</sub> and<sub>N</sub> Mary<sub>C</sub>]<sub>A</sub> went<sub>P</sub> [to school]<sub>A</sub> together<sub>D</sub>”
  - “I<sub>A</sub> ’ll have<sub>P</sub> [coffee<sub>C</sub> and<sub>N</sub> cookies<sub>C</sub>]<sub>A</sub>” **NSS: this is one of many places where ] is preceded by a space. if that's not intentional, should be easy to fix with a find-replace.**
2. ELABORATORS (E) add some information about one specific entity or relation, without changing its semantic type. These either include inherent attributes (attributes that cannot change because they define the entity), or one of several types of relations specified below. The unit which is elaborated on, and which determines the type of the unit, is marked as a CENTER (C).
  - “Queen<sub>C</sub> of<sub>R</sub> England<sub>E</sub>” (describes a type of a queen; the fact that she is the queen of England is inherent to her being a queen)
  - “His<sub>E</sub> father<sub>C</sub>”
  - “Chocolate<sub>E</sub> cookies<sub>C</sub>”, “Metal<sub>E</sub> belt<sub>C</sub>” (the substance something is made of)

Other types of relations that are considered Es:

- Determiners (“the<sub>E</sub> last king<sub>C</sub> of Scotland”)
  - Class descriptor: Units comprised of a sub-unit that specifies the name of the entity in question, and another unit specifies which category it belongs to. In these cases, the specific unit is marked with *C* and the class descriptor is marked as *E*.
    - “[the name]<sub>E</sub> [John]<sub>C</sub>”
    - “the state<sub>E</sub> of Washington<sub>C</sub>”
    - “[the year]<sub>E</sub> [1966]<sub>C</sub>”
3. If it is hard to say which of the sub-units adds information to which, both units should be marked as Cs. That is, if there is no one unit that determines the type of entity, all units that determine its type should be marked as Cs.

A frequent example of that is part-whole relations: Units comprised of a sub-unit that specifies the whole, and one that specifies a sub-part of it.

- “bottom<sub>C</sub> of<sub>R</sub> [the<sub>E</sub> sea]<sub>C</sub>”
- “tip<sub>C</sub> of<sub>R</sub> [the<sub>E</sub> iceberg]<sub>C</sub>”
- “See-<sub>C</sub> boden<sub>C</sub>” *de*

### Model #3: Inter-Scene relations

1. Linkage is the term for inter-Scene relations in UCCA. There are four major types of relations in which Scenes may participate, and therefore four types of categories Scene units may receive. The next three items describe these types.
2. **Elaborator Scenes:** an E-Scene adds information to a previously established unit. Usually answers a “which X” or “what kind of X” question. Es should place the C they are elaborating as a *remote* A (see below). A way to check where a Scene is an E-Scene is to ask whether the Scene along with the C it relates to are of the same type as the C itself.
  - “[The<sub>E</sub> dog<sub>C</sub> [that ate my homework (*dog*)<sub>A</sub>]<sub>E</sub>]<sub>A</sub> is brown” (“dog” is a remote A in “that ate my homework”)
  - “The<sub>E</sub> person<sub>C</sub> [whom<sub>F</sub> I<sub>A</sub> gave<sub>P</sub> [the present]<sub>A</sub> [to<sub>R</sub> (*person*)<sub>C</sub>]<sub>A</sub>]<sub>E</sub>”
  - “Brad played [an<sub>E</sub> American<sub>C</sub> [taken to the Adriatic (*American*)<sub>A</sub>]<sub>E</sub>]<sub>A</sub>”
3. **Participant Scenes:** an A-Scene is a participant in a larger Scene. It does not add information to some specific participant in it, and if you remove it, it doesn't retain the same type. Usually answers a “what” question about the Scene.
  - “[Talking to strangers]<sub>A</sub> is<sub>F</sub> ill-advised<sub>S</sub>” (answers “what is ill-advised?”)
  - “John<sub>A</sub> said<sub>P</sub> [he's hungry]<sub>A</sub>” (answers “what did John say?”)
  - “[John<sub>C</sub> 's<sub>R</sub>]<sub>A</sub> accurate<sub>D</sub> kick<sub>P</sub>]<sub>A</sub> saved<sub>P</sub> [the game]<sub>A</sub>” (answers “what saved the game?”)
4. **Parallel Scenes:** any other Scene receives the category Parallel Scene (H). Sometimes there is an accompanying relation word and sometimes not. If so, it is a Linker (L). Note that there are no Adverbial (D) Scenes. Except for Ground (see below), if a Scene is not an A (Participant) or an E (Elaborator), it's an H.
  - “[John managed to amuse himself]<sub>H</sub> while<sub>L</sub> [waiting in line (*John*)<sub>A</sub>]<sub>H</sub>”
  - “[My house feels fresh]<sub>H</sub> [thanks to]<sub>L</sub> [[the Battery Park Pest]<sub>A</sub> (IMPLICIT)<sub>P</sub>]<sub>H</sub>.”
  - “[The minute]<sub>L</sub> [I got home]<sub>H</sub> [I noticed the new painting]<sub>H</sub>”
  - “[If]<sub>L</sub> [you build it]<sub>H</sub> [they will come (IMPLICIT)<sub>A</sub>]<sub>H</sub>” (where they come to is implicit)
  - “[I'd done some research]<sub>H</sub>, [asked a couple of questions]<sub>H</sub> and<sub>L</sub> [found myself thinking]<sub>H</sub>” **NSS: isn't there an implicit argument: the topic of all of these? OA: implicit units are admittedly one of the weaker points of the scheme, as in the absence of a strong lexical model, defining what is implicit and what is simply not there is difficult. The reasons we added them still, is that it would be more confusing imo not to include them. Do you think we need to be more upfront about them? NSS: Yeah I think there should be some discussion, even if there are borderline cases.**
  - “[You're only saying this]<sub>H</sub> because<sub>L</sub> [John told you to (*say this*)<sub>A</sub>]<sub>H</sub>”
  - “Nach<sub>L</sub> [einer Rolle in einem Thriller]<sub>H</sub> [spielte sie in einem Actionfilm mit]<sub>H</sub>.”<sup>de</sup>

Specific cases of Parallel Scenes include (examples of relevant Linkers in brackets): purposive (“in order to” or “to”, ‘um + zu-Infinitiv’<sup>de</sup>), logical (“if ... then ...”), temporal (“when X, Y”, “before X, Y”), coordination (“and”, “but”), and contrastive linkages (“however”, “still”, “jedoch”<sup>de</sup>).

Scenes that are not related to any other units and are therefore in the top level of organization in the text are also Hs (Parallel Scenes).

Linkers do not necessarily appear between the Scenes they are linking (see example #2 above).

5. A unit is marked as GROUND (G), if its primary purpose is to relate some unit to its the speech event; either the speaker, the hearer or the general context in which the text was uttered/written/conceived.<sup>1</sup>

Gs are similar to Ls, except that they don't relate the Scene to some other Scene in the text, but rather to the speech act of the text (the speaker, the hearer or their opinions). By convention, Ground units should be positioned within the Scene they relate to.

- “[Surprisingly]<sub>G</sub> , [our flight]<sub>A</sub> arrived<sub>P</sub> [on time]<sub>T</sub>]<sub>H</sub>”
- “[In my opinion]<sub>G</sub>, John<sub>A</sub> is<sub>F</sub> coming<sub>P</sub> home<sub>A</sub> ]<sub>H</sub>”

Note that a complete Scene that refers to the ground (with As and Ds etc.) should be annotated as a Scene and not as a G. That is, if a unit alludes to the speech event, but is missing almost all its elements save for one word or expression, it should be a G. If the speech event is mentioned more elaborately, it should be annotated as a Scene.

- “[I<sub>A</sub> was<sub>F</sub> surprised<sub>S</sub> ]<sub>H</sub> when<sub>L</sub> [[our flight]<sub>A</sub> arrived<sub>P</sub> [on time]<sub>T</sub>]<sub>H</sub>”
- But: “[Surprisingly]<sub>G</sub>, [our flight]<sub>A</sub> arrived<sub>P</sub> [on time]<sub>T</sub>]<sub>H</sub>”
- “I<sub>A</sub> told<sub>P</sub> you<sub>A</sub> already<sub>D</sub> [that John can't make it]<sub>A</sub>”

## Relations that Appear in All Models

There are three types of categories that may appear anywhere in the text: Functions (F), Relators (R) and Quantifiers (Q).

1. Relators are relations that relate between two or more entities within Scene units as well as non-Scene units. Rs in English are usually prepositions (see Section 4 below for a more elaborate discussion).

- Relators within Scene units. When a Relator connects between the main event (P/S) and another Scene element (A,T,D) then it should be included inside the Scene element (A,T,D) it pertains to.
  - “John said [that<sub>R</sub> he<sub>A</sub> 's<sub>F</sub> going<sub>P</sub> home<sub>A</sub>]<sub>A</sub>”
  - “I<sub>A</sub> referred<sub>P</sub> [to<sub>R</sub> John<sub>C</sub> and<sub>N</sub> Mary<sub>C</sub>]<sub>A</sub>”
  - “I<sub>A</sub> referred<sub>P</sub> [to<sub>R</sub> John<sub>C</sub> and<sub>N</sub> to<sub>R</sub> Mary<sub>C</sub>]<sub>A</sub>”
- Within non-Scene units By convention, we place the Rs in non-Scene units as siblings of the Es, Qs and Cs they relate (on the same level with them).
  - “[a<sub>E</sub> group<sub>C</sub>]<sub>Q</sub> of<sub>R</sub> journalists<sub>C</sub>”
  - “bottom<sub>C</sub> of<sub>R</sub> [the<sub>E</sub> sea<sub>C</sub>]<sub>C</sub>”
  - “Queen<sub>C</sub> of<sub>R</sub> England<sub>E</sub>”
- When will we not use Relators?: - To link between Parallel Scenes (for that see Linkers). - To connect between Centers that have the same parent unit and carry a similar type or role (for that see Connectors).

2. Functions (F) are units that do not introduce a new participant or relation. They can only be interpreted as part of a larger construction in which they are situated, or convey some aspect of meaning which is not covered by the foundational layer (e.g., tense or focus). Usually in these cases, they cannot be substituted with any other word.

- “I<sub>A</sub> want<sub>D</sub> to<sub>F</sub> run<sub>P</sub> [a<sub>E</sub> marathon<sub>C</sub>]<sub>A</sub>”
- “I<sub>A</sub> am<sub>F</sub> going<sub>P</sub> [to<sub>R</sub> the<sub>E</sub> supermarket<sub>C</sub>]<sub>A</sub>”
- “It<sub>F</sub> is likely<sub>S</sub> [that he will make it]<sub>A</sub>”

---

<sup>1</sup>The speech event is called Ground following R. Langacker.



- “Let<sub>F</sub> me<sub>A</sub> introduce<sub>P</sub> John<sub>A</sub>”

Fillers (e.g. "ummm", "like", "okay", "so") should be marked also as Fs. (“ummm<sub>F</sub> I<sub>A</sub> heard<sub>P</sub> you<sub>A</sub> say<sub>P</sub> that<sub>A</sub>”)

NSS: should “you say that” be an A of ‘heard’? OA: well, if that what he heard then yes; it depends on context. NSS: What is the other interpretation?

3. We use QUANTIFIERS (Q) to mark expressions that:

- Describe the quantity or magnitude of an entity:<sup>2</sup>
  - “three<sub>Q</sub> apples<sub>C</sub>”
  - “several<sub>Q</sub> apples<sub>C</sub>”
  - “I<sub>A</sub> bought<sub>P</sub> [[three<sub>E</sub> kilos<sub>C</sub>]<sub>Q</sub> of<sub>R</sub> apples<sub>C</sub>]<sub>A</sub>”
- Any expression that set defines that an entity is a group or a set (e.g., “group of ...”, “hundreds of ...”).
  - “[a<sub>E</sub> group<sub>C</sub>]<sub>Q</sub> of<sub>R</sub> journalists<sub>C</sub>”
  - “[a<sub>E</sub> swarm<sub>C</sub>]<sub>Q</sub> of<sub>R</sub> bees<sub>C</sub>”
  - “[a<sub>E</sub> variety<sub>C</sub>]<sub>Q</sub> of<sub>E</sub> colors<sub>C</sub>”
- In Scene units, for any expression that indicates the amount of occurrences of a single event:
  - “We<sub>A</sub> had<sub>F</sub> talked<sub>P</sub> [three<sub>E</sub> times<sub>C</sub>]<sub>Q</sub> [over the last week]<sub>T</sub>”
- Ordinals are also marked as Qs:
  - “[My<sub>A</sub> first<sub>Q</sub> kick<sub>P</sub>]<sub>A</sub> saved<sub>P</sub> [the game]<sub>A</sub>”
  - “The<sub>E</sub> first<sub>Q</sub> king<sub>C</sub> of<sub>R</sub> Scotland<sub>E</sub>”
  - “I<sub>A</sub> got<sub>P</sub> here<sub>A</sub> first<sub>Q</sub>”
  - “I<sub>A</sub> was<sub>F</sub> [the first]<sub>Q</sub> to<sub>F</sub> arrive<sub>P</sub>”

## Remote and Implicit Units

1. There are instances where a sub-unit in a given unit is not explicitly mentioned. We can indicate the missing sub-unit in two ways:

- (a) Add a reference of the missing unit from another place in the text, as a Remote unit.
- (b) When it does not appear explicitly in any place in the text, add an Implicit unit to stand for the missing sub-unit.

We add a Remote/Implicit unit whenever there is a element or relation which we think to be strongly present in the conceptualization of the Scene, but is not explicitly mentioned.

NSS: this may be quite subjective. I haven't read the detailed guidelines yet, so maybe that has more precise criteria, but a thought is that ellipsis constructions and what FrameNet calls Definite Null Instantiations (where a specific referent is omitted in context) and Constructional Null Instantiations (e.g., imperative, passive without by-phrase) should trigger a remote/implicit unit. Indefinite Null Instantiations (e.g., “I ate.”—what is eaten is unimportant) probably should NOT trigger remote/implicit units. OA: agreed. see my comment above. Let's discuss this on our next meeting.

Remote and Implicit units should be assigned relevant categories like any other unit.

Examples: (target relations (X) underlined, remote units (Y) boldfaced)

<sup>2</sup>Expressions of distance will be marked as As and then internally, Quantity to mark the actual measurement: “He ran [100<sub>Q</sub> meters<sub>C</sub>]<sub>A</sub>”.

- “[John got home]<sub>H</sub> and<sub>L</sub> [took a shower (John)<sub>A</sub>]<sub>H</sub>”
- “[The dog [I saw last night (dog)<sub>A</sub>]<sub>E</sub> ]<sub>A</sub> was<sub>F</sub> brown<sub>S</sub>”
- “[We<sub>A</sub> just<sub>T</sub> opened<sub>P</sub> (IMPLICIT)<sub>A</sub>]<sub>H</sub>” (the thing opened is implicit) **NSS: I think “we” is metonymic for the institution that opened, such as a show or business. I don’t think “we” is agentive in the sense of causing something independent to open—if referring to a bottle you’d need to say “We just opened it.” OA: OK. Let’s remove this example.**
- “[Sure<sub>G</sub> John<sub>A</sub> is<sub>F</sub> (IMPLICIT)<sub>P</sub>]<sub>H</sub>” (John is what? if it appears explicitly in the previous sentence, it’s a remote unit; otherwise it’s an implicit unit)
- “[John<sub>A</sub> is<sub>F</sub> tall<sub>S</sub>]<sub>H</sub>, [Mary<sub>A</sub> is<sub>F</sub> n’t<sub>D</sub> (tall)<sub>S</sub> ]<sub>H</sub>”

### 3 Technical Notes and Guidelines

1. With any problem or question, contact the administrator of the project. If there is uncertainty, mark your guess and add “uncertain”.
2. When annotating a remote unit, select the minimal possible relevant unit, and not its ancestors.
3. Top-level annotation (i.e., of units directly below the passage level) should be annotated, wherever possible, according to the Scene model. The only exceptions are cases that do not describe a Scene in any way (such as section titles).
4. Prefer Ls over Ds, where possible.
5. Prefer Ls over Gs where possible.
6. Prefer Ls over Ts where possible.
7. Prefer Ds over Ts where possible.
8. Prefer annotating A-Scenes and E-Scenes over Parallel Scenes where possible.
9. Prefer separating participants from their relations where possible.
10. Prefer Ds over a longer P/S with an E inside it. More generally, try to avoid complex or long P/S.
11. Use Implicit units sparingly and prefer Remote units where possible.
12. Do not create units only to be used later as a Remote unit. Use existing units instead.
13. Since morphology in English is very impoverished, we take a pragmatic approach and in our primary layer do not annotate parts of words, but only sets of complete words. In further layers, a sub-unit may cover a part of a word, as long as that part refers to a relation covered by UCCA or to a participant in it.
14. Function units (Fs) do not refer to a participant or relation and, since the UCCA annotation reflects participation in relations, it is often not clear in what level of the hierarchy an F unit should be placed in. When this occurs, include the F in the deepest unit that stands to reason.
15. Single words are often Scenes as well. This will usually happen where none of the participants is explicitly mentioned.
  - “[The<sub>E</sub> [negotiations<sub>P</sub>]<sub>C</sub> ]<sub>A</sub> [took place]<sub>P</sub> [in<sub>R</sub> Rome<sub>C</sub>]<sub>A</sub>”

- “The<sub>E</sub> [available<sub>S</sub>]<sub>E</sub> options<sub>C</sub>”
- “[Crying<sub>P</sub> (you)<sub>A</sub>]<sub>A</sub> makes<sub>D</sub> you<sub>A</sub> stronger<sub>P</sub>”
- “[I<sub>A</sub> went<sub>P</sub> [to<sub>R</sub> the<sub>E</sub> store<sub>C</sub>]<sub>A</sub>]<sub>H</sub> for<sub>L</sub> [eggs<sub>A</sub> (I)<sub>A</sub> (IMPLICIT)<sub>P</sub>]<sub>H</sub>” (“for” is a purposive linker. The implicit P in the second Scene is for the buying action)

## 4 Classification of Prepositions

Prepositions are in frequent use in English. They include words such as “in”, “on”, “after”, “with” and “under” or “nach”, “in” and “auf” in German<sup>de</sup>. Some prepositions are multi-worded, in which case they are internally annotated as unanalyzable. Examples include “thanks to” and “on top of”.

### 1. Prepositions as Relators:

(a) In Scene units, Relators are included inside the Scene element they pertain to:

- “John<sub>A</sub> put<sub>P</sub> [the<sub>E</sub> hat<sub>C</sub>]<sub>A</sub> [on<sub>R</sub> the<sub>E</sub> shelf<sub>C</sub>]<sub>A</sub>”
- “John<sub>A</sub> relied<sub>P</sub> [on<sub>R</sub> his<sub>E</sub> father<sub>C</sub>]<sub>A</sub>”
- “John<sub>A</sub> just<sub>T</sub> heard<sub>P</sub> [of<sub>R</sub> his<sub>A</sub> re-election<sub>P</sub>]<sub>A</sub>”
- “John<sub>A</sub> referred<sub>P</sub> [to<sub>R</sub> Mary<sub>C</sub>]<sub>A</sub> [in<sub>R</sub> his<sub>A</sub> dissertation<sub>P</sub>]<sub>A</sub>”

NSS: include some examples where PP is D or T rather than A? OA: Yes. Dotan, could you add such?

(b) In non-Scene units, they are placed on the same level with the Es, Qs and Cs they relate:

- “President<sub>C</sub> of<sub>R</sub> [the<sub>E</sub> USA<sub>C</sub>]<sub>E</sub>”
- “The<sub>E</sub> finest<sub>E</sub> hotels<sub>C</sub> of<sub>R</sub> [the<sub>E</sub> world<sub>C</sub>]<sub>E</sub>”
- “bottom<sub>C</sub> of<sub>R</sub> [the sea]<sub>C</sub>”
- “[a period]<sub>C</sub> of<sub>R</sub> time<sub>C</sub>”
- “[a group]<sub>Q</sub> of<sub>R</sub> journalists<sub>C</sub>”
- “millions<sub>Q</sub> of<sub>R</sub> dollars<sub>C</sub>”
- “plenty<sub>Q</sub> of<sub>R</sub> fish<sub>C</sub>”
- “[four<sub>Q</sub> episodes<sub>C</sub>]<sub>E</sub> of<sub>R</sub> Dallas<sub>C</sub>”

NSS: include some non-“of” examples, like “people with red hair”, “the writing on the wall”? OA: good idea. Dotan, could you add some?

2. **Phrasal verbs:** the preposition changes the semantics of the verb in an unpredictable way. In that case the preposition is considered to be a part of the S or P. The P/S together form an unanalyzable unit (as it does not have sub-parts with significant semantic input).

- “John<sub>A</sub> [gave up]<sub>P</sub> [his pension]<sub>A</sub>”
- “John<sub>A</sub> let<sub>P-</sub> Mary<sub>A</sub> down<sub>-P</sub>”
- “John<sub>A</sub> [took]<sub>P-</sub> Mary<sub>A</sub> [up on]<sub>-P</sub> [her<sub>A</sub> promise<sub>P</sub>]<sub>A</sub>”

Note that this case does not cover cases where the preposition doesn’t change the semantics of the main relation, but is mandatory (“inherent preposition”), such as in “John is next [to<sub>R</sub> Mary<sub>C</sub>]<sub>A</sub>”, “John relies [on<sub>R</sub> Mary<sub>C</sub>]<sub>A</sub>”.

3. **Main relations:** If the preposition is the main relation in the Scene, then it is a P or S. NSS: can it ever be P? OA: good question. not sure we ever ran into one...

- “[The apple tree]<sub>A</sub> is<sub>F</sub> in<sub>S</sub> [the garden]<sub>A</sub>”
- “John<sub>A</sub> is<sub>F</sub> into<sub>S</sub> Mary<sub>A</sub>”

## 5 Participant-Adverbial Distinction

A basic issue in almost any grammatical theory is to determine when a unit is a participant and when it is a secondary relation. In UCCA, this is the distinction between Participants and Adverbials.

- Any unit that introduces a new participant is an A. Subjects, objects, instruments, locations, destinations are therefore invariably As.
- Adverbs and any other units that introduce another relation (without introducing a participant) into the Scene are Ds. Manner adverbs (e.g., “quickly”, “politely”) are invariably Ds.
- Prepositional phrases constitute most of the borderline cases.

### Examples:

1. “John<sub>A</sub> suffered<sub>P</sub> [for<sub>R</sub> the<sub>E</sub> team<sub>C</sub>]<sub>A</sub>”
2. “Woody<sub>A</sub> walked<sub>P</sub> [in<sub>R</sub> the<sub>E</sub> park<sub>C</sub>]<sub>A</sub> yesterday<sub>T</sub>”
3. “John<sub>A</sub> cut<sub>P</sub> [the cake]<sub>A</sub> [with<sub>R</sub> a<sub>E</sub> knife<sub>C</sub>]<sub>A</sub>”
4. “John<sub>A</sub> behaved<sub>P</sub> recklessly<sub>D</sub>”
5. “Woody<sub>A</sub> treated<sub>P</sub> him<sub>A</sub> [with<sub>R</sub> disrespect<sub>C</sub>]<sub>D</sub>”
6. “Texas<sub>A</sub> won<sub>P</sub> [in<sub>R</sub> its<sub>E</sub> home<sub>E</sub> court<sub>C</sub>]<sub>A</sub>”
7. “John<sub>A</sub> bought<sub>P</sub> milk<sub>A</sub> [next door]<sub>A</sub> [for<sub>R</sub> 50<sub>Q</sub> p<sub>C</sub>]<sub>A</sub>” (“next door” is a location, albeit a vague one)

## 6 Analyzability

By default, analyze all cases down to the word level. The only cases which should not be analyzed are:

- Where the internal structure cannot be analyzed using any of the models: Scene, E+C, multiple Cs (possibly with N), inter-Scene relations.
- This usually happens where it’s not clear what the meanings of the individual words in this context are.
- Names should not be internally analyzed.

### Examples:

1. “The<sub>E</sub> October<sub>E</sub> [Revolution<sub>P</sub>]<sub>C</sub>”: analyzable although it is not simply a revolution that happened in October, but rather a specific one.
2. “Chief<sub>E</sub> executive<sub>E</sub> officer<sub>C</sub>”: analyzable.
3. “University<sub>C</sub> of<sub>R</sub> Texas<sub>E</sub>”: analyzable.
4. “[The<sub>E</sub>]<sub>P</sub> – real<sub>D</sub> [deal<sub>C</sub>]<sub>P</sub>”: analyzable although it’s an idiomatic expression since the sub-parts do convey relevant meaning.
5. “as well as”: unanalyzable since it’s not really clear which categories to assign to the individual words.
6. “give up”: unanalyzable as it is not clear what meaning “give” and “up” contribute to the expression.
7. “I saw Tom Cruise in Top Gun”: “Tom Cruise” and “Top Gun” are unanalyzable (names). **NSS: What about compositional names (“The Curious Incident of the Dog in the Night-time”, “Food and Drug Administration”)?** **OA: Names of movies and such we still treat as unanalyzable. We should probably say that.**

## 7 Detailed Guidelines

### 7.1 Scenes

**Annotating Scenes within Scenes.** In order to analyze a Scene within a Scene we have two options:

1. Analyze it first with Center-Elaborator relations (see Section 2 for elaboration on non-Scene units).
2. Analyze it first with Process/State-Participant relations (see Section 2 for elaboration on Scene units).

To determine this we ask ourselves what would we mark as the Center? If it's a concrete entity then we begin with Center-Elaborator relations, but if it's some kind of action or state then we annotate it directly as a Scene.

- Analysis of a Scene within Scene first with Center-Elaborator relations:
  - “[The<sub>E</sub> dog<sub>C</sub> [that<sub>R</sub> ate<sub>P</sub> [my homework]<sub>A</sub> (dog)<sub>A</sub>]<sub>E</sub> ]<sub>A</sub> is<sub>F</sub> brown<sub>S</sub>”
  - “I<sub>A</sub> like<sub>S</sub> [ burned<sub>P</sub> (coffee)<sub>A</sub>]<sub>E</sub> coffee<sub>C</sub> ]<sub>A</sub>”
  - “Brad<sub>A</sub> played<sub>P</sub> [an<sub>E</sub> American<sub>C</sub> [(American)<sub>A</sub> going<sub>P</sub> [to<sub>R</sub> the<sub>E</sub> Adriatic<sub>C</sub> ]<sub>A</sub> ]<sub>E</sub> ]<sub>A</sub>”
- Analysis of a Scene within Scene directly as a Scene:
  - “[ [John<sub>C</sub> 's<sub>R</sub> ]<sub>A</sub> kick<sub>P</sub> ] saved<sub>P</sub> [the<sub>E</sub> game<sub>C</sub> ]<sub>A</sub>”
  - “John<sub>A</sub> said<sub>P</sub> [he<sub>A</sub> invented<sub>P</sub> [skating<sub>P</sub> ]]<sub>A</sub>”

**Dependent Scenes.** A Scene is not necessarily something that can stand on its own. It may require a larger construction to rely on, but it is still considered a Scene:

1. “[he<sub>A</sub> retired<sub>P</sub> ]<sub>H</sub> [with]<sub>L</sub> [ [a<sub>E</sub> rank<sub>C</sub> ]<sub>S</sub> [of<sub>R</sub> major<sub>C</sub> ]<sub>A</sub> (he)<sub>A</sub> ]<sub>H</sub>”
2. “[Mary<sub>A</sub> cuts<sub>P</sub> [her<sub>E</sub> hair<sub>C</sub> ]<sub>A</sub> ]<sub>H</sub> [like]<sub>L</sub> [ [a<sub>E</sub> boy<sub>C</sub> ]<sub>A</sub> (cuts)<sub>P</sub> ]<sub>H</sub>”
3. “[once<sub>T</sub> poor<sub>S</sub> (he)<sub>A</sub> ]<sub>H</sub>, [he<sub>A</sub> now<sub>T</sub> owns<sub>S</sub> [a<sub>E</sub> spacious<sub>E</sub> apartment<sub>C</sub> ]<sub>A</sub> ]<sub>H</sub>”

**Distinguishing Ground and Participant Scenes.** A Ground unit relates to the speech event itself or some aspect of it. It does not introduce a new Scene above and beyond evoking the speech event. We do not analyze the internal structure of Gs.

1. “[ [The truth is that]<sub>G</sub> John<sub>A</sub> is<sub>F</sub> [a<sub>E</sub> conservative<sub>C</sub> ]<sub>S</sub> ]<sub>H</sub>”
2. “[ [Surprisingly]<sub>G</sub> , [I]<sub>A</sub> [saw]<sub>P</sub> [John]<sub>A</sub> [in the park]<sub>A</sub> ]<sub>H</sub>”
3. “[ [To my surprise]<sub>G</sub> [I]<sub>A</sub> [saw]<sub>P</sub> [John]<sub>A</sub> [in the park]<sub>A</sub> ]<sub>H</sub>”
4. “[you<sub>A</sub> can<sub>D</sub> go<sub>P</sub> home<sub>A</sub>, [for all I care]<sub>G</sub> ]<sub>H</sub>”

Contrast with participant Scenes. Both “I think” and “Mary saw” introduce a new Scene, with a new P. They are therefore participant Scenes.

1. “I<sub>A</sub> think<sub>P</sub> [that<sub>R</sub> John<sub>A</sub> is<sub>F</sub> [a<sub>E</sub> conservative<sub>C</sub> ]<sub>S</sub> ]<sub>A</sub>”
2. “Mary<sub>A</sub> saw<sub>P</sub> [John<sub>A</sub> running<sub>P</sub> [in<sub>R</sub> the<sub>E</sub> park<sub>C</sub> ]<sub>A</sub> ]<sub>A</sub>”

**Static Scenes.** Static Scenes are Scenes which can be fully described by a single picture, that does not change throughout the duration of the state. Following are several examples of static Scenes.

1. Identity. Expressing the identity between two entities.<sup>3</sup> Identity is the case where there are two well-defined entities (not a set or a relation, but two well-defined entities) and optionally a copula (e.g., “be”). This is the only case where a copula serves as the main relation of a Scene.

(a) “[The morning star]<sub>A</sub> is<sub>S</sub> [the evening star]<sub>A</sub>”

(b) “[That<sub>E</sub> person<sub>C</sub>]<sub>A</sub> is<sub>S</sub> John<sub>A</sub>”

(c) But: “John<sub>A</sub> is<sub>F</sub> [a<sub>E</sub> member<sub>C</sub>]<sub>S</sub> [of<sub>R</sub> the<sub>E</sub> NRA<sub>C</sub>]<sub>A</sub>” (since here “a member of the NRA” is not one specific entity, but a role that could apply to multiple people). **NSS: “set” is confusing since “a member” is singular. How about: “role that could apply to multiple people”? OA: changed**

2. Attribution/Benefaction/Location. Specifying a quality, a benefactor or a location of an argument.

(a) “[This<sub>E</sub> man<sub>C</sub>]<sub>A</sub> is<sub>F</sub> clever<sub>S</sub>”

(b) “[This<sub>E</sub> present<sub>C</sub>]<sub>A</sub> is<sub>F</sub> for<sub>S</sub> [ [John<sub>C</sub> ’s<sub>R</sub>]<sub>A</sub> birthday<sub>P</sub>]<sub>A</sub>”

(c) “[The<sub>E</sub> apple<sub>E</sub> tree<sub>C</sub>]<sub>A</sub> is<sub>F</sub> in<sub>S</sub> [the<sub>E</sub> garden<sub>C</sub>]<sub>A</sub>”

3. Alienable Possession (i.e., except for cases of possession used to express a body part, e.g., “my hand”, or a relative, e.g., “my father”, which is not a Scene).

(a) “[This<sub>E</sub> book<sub>C</sub>]<sub>A</sub> is<sub>F</sub> John<sub>A</sub> ’s<sub>S</sub>”

(b) “[This<sub>E</sub> book<sub>C</sub>]<sub>A</sub> is<sub>F</sub> mine<sub>S+A</sub>” (S+A: both an S and an A)

**NSS: What about ‘my book’, ‘John’s book’?**

4. Existential Scenes: This is a special type of a static Scene. Since “there are” determines the relation here (namely existence), it is marked as S.

(a) “There<sub>S</sub> are<sub>F</sub> [thousands<sub>Q</sub> of<sub>R</sub> us<sub>C</sub>]<sub>A</sub>”

**Note:** the category is not defined by the words comprising the unit, but by the function the unit has in the unit it is placed in. Consider these pairs of examples:

(a) “John<sub>A</sub> is<sub>F</sub> sitting<sub>P</sub> [in<sub>R</sub> the<sub>E</sub> garden<sub>C</sub>]<sub>A</sub>” / “[The<sub>E</sub> apple<sub>E</sub> tree<sub>C</sub>]<sub>A</sub> is<sub>F</sub> in<sub>S</sub> [the<sub>E</sub> garden<sub>C</sub>]<sub>A</sub>”

(b) “[John<sub>A</sub> bought<sub>P</sub> wine<sub>A</sub>]<sub>H</sub> for<sub>L</sub> [[Mary<sub>C</sub> ’s<sub>R</sub>]<sub>A</sub> birthday<sub>P</sub>]<sub>H</sub>” / “[This present]<sub>A</sub> is<sub>F</sub> for<sub>S</sub> [[John<sub>C</sub> ’s<sub>R</sub>]<sub>A</sub> birthday<sub>P</sub>]<sub>A</sub>”

**Scene or not a Scene.** One of the most important decisions in UCCA annotation is to determine whether a relation is an S/P (and evokes a Scene) or not. Processes are usually easier to spot – they describe an event that evolves in time, usually some action or movement. As for States, they differ from non-Scenes in not being inherent properties of the Center, but something that may have been different in the past or will be different in the future.

1. “[The<sub>E</sub> outbreak<sub>C</sub>]<sub>D</sub> [of<sub>R</sub> the<sub>E</sub> War<sub>C</sub>]<sub>P</sub>” – a Scene.
2. “Oscillating<sub>P</sub> [between<sub>R</sub> atheism<sub>C</sub> and<sub>N</sub> agnosticism<sub>C</sub>]<sub>A</sub>” – a Scene.
3. “[John<sub>C</sub> ’s<sub>R</sub>]<sub>A</sub> accurate<sub>D</sub> kick<sub>P</sub>” – a Scene.
4. “[broken<sub>S</sub> (glass)<sub>A</sub>]<sub>E</sub> glass<sub>C</sub>]<sub>A</sub> is<sub>F</sub> dangerous<sub>S</sub>” – a Scene.

<sup>3</sup>Do not confuse identity with “Noun as a P/S” (see below).

5. “John<sub>A</sub> always<sub>D</sub> wanted<sub>P</sub> [a<sub>E</sub> garden<sub>C</sub> [ with<sub>S</sub> trees<sub>A</sub> (garden)<sub>A</sub> ]<sub>E</sub>]<sub>A</sub>” – “with trees” is a Scene, since it is some property of the garden that could potentially change. **NSS: In “The garden with trees was beautiful”, would “the garden with trees” be analyzed differently? What about “The garden was beautiful with trees”?**
6. “The trees are<sub>F</sub> in<sub>S</sub> [the<sub>E</sub> garden<sub>C</sub>]<sub>A</sub>” – a Scene, since being in the garden is not an inherent property of the trees.

**One Scene or two.** Where two potentially Scene-evoking relations appear in proximity to one another, the question of whether to consider them one complex Scene or two separate ones arises. It should be one Scene if the two relations are conceptually hard to separate and are similar in their participants, time, location and ground. It should be two Scenes if this is not the case.

1. “[I got home]<sub>H</sub> and<sub>L</sub> [took a shower]<sub>H</sub>” (2 Scenes with a temporal relation)
2. “[it took a lot of effort]<sub>H</sub> to<sub>L</sub> [win this fight]<sub>H</sub>” (2 Scenes, with a purposive relation)
3. “[he is on vacation]<sub>H</sub>, [sailing a yacht near Greece]<sub>H</sub>” (2 Scenes)
4. “[John<sub>A</sub> eats<sub>P</sub> (enthusiastically)<sub>D</sub>]<sub>H</sub> and [drinks<sub>P</sub> enthusiastically<sub>D</sub> (John)<sub>A</sub>]<sub>H</sub>” (2 Scenes, “eating” and “drinking” are two conceptually different actions)
5. “[She<sub>A</sub> [went away]<sub>P</sub>]<sub>H</sub> [angry<sub>P</sub> (She)<sub>A</sub>]<sub>H</sub>” (borderline case; could be 1 Scene or 2; the two original Scenes, of her going away and of her being angry are fused into one)

**Secondary Verb or Participant Scene.** Distinguishing between secondary verb constructions and Participant Scene constructions is done by determining whether the sentence in question refers to one or two Scenes. Participant Scenes correspond to cases where there are two separate Scenes, while secondary verbs correspond to the cases where there are two relations, one dependent (secondary, not describing a Scene in its own right, could not by itself be the P/S of a Scene) and one independent (the main relation) within the same Scene.

1. “He<sub>A</sub> demanded<sub>P</sub> [to<sub>R</sub> see<sub>P</sub> [the<sub>E</sub> manager<sub>C</sub>]<sub>A</sub> (He)<sub>A</sub>]<sub>A</sub>” (2 Scenes, since the demanding and the seeing are two separate Scenes which can take place in different times and locations)
2. “He<sub>A</sub> began<sub>D</sub> kicking<sub>P</sub> [the ball]<sub>A</sub>” (one Scene, since “began” does not describe an action in its own right, but is dependent on the “kicking”)
3. “He<sub>A</sub> wants<sub>D</sub> to<sub>F</sub> kick<sub>P</sub> [the ball]<sub>A</sub>” (one Scene, since “wants” does not describe an action in its own right, but is dependent on the “kicking”)
4. “He<sub>A</sub> became<sub>D</sub> [a<sub>E</sub> doctor<sub>C</sub>]<sub>S</sub>” (one Scene; the becoming and him being a doctor are the same conceptual event)
5. “He<sub>A</sub> is<sub>F</sub> known<sub>P</sub> [as<sub>R</sub> [a<sub>E</sub> doctor<sub>C</sub>]<sub>S</sub> (He)<sub>A</sub>]<sub>A</sub>” (two Scenes; him being known to be something and him being a doctor)
6. “[John]<sub>A</sub> said<sub>P</sub> [he<sub>A</sub> is<sub>F</sub> [a<sub>E</sub> doctor<sub>C</sub>]<sub>S</sub>]<sub>A</sub>” (two Scenes; John saying and him being a doctor are easy to conceptualize as two different scenes, the sentence just places them together)

**Scene within Scene or two Parallel Scenes.** In order to decide whether a Scene should be included within a larger Scene we first need to ask what role it will be assigned. If we think it is an A then we can indeed include it as an A-Scene in the larger Scene. But If we think it should be a D or T then we instead mark it separately as an H since Ds and Ts cannot be Scenes.

- “John said [that two men were fighting in the street]<sub>A</sub>” Scene within Scene (“men fighting in the street” is an A Scene in the larger Scene)
- “[John usually plays soccer]<sub>H</sub> after<sub>L</sub> [he finishes his homework]<sub>H</sub>” two Parallel Scenes (If we replace “he ... homework” with a simple non-Scene unit, e.g., “John usually plays soccer after 16:00”, then it’s clear that the relation between the units is Time, but since T can’t be a Scene, we mark it as an H instead).
- “[You didn’t do it]<sub>H</sub> [the way]<sub>L</sub> [you should have (*do*)<sub>P</sub> (*it*)<sub>A</sub>]<sub>H</sub>” two Parallel Scenes (“The way you should have” relates to the manner in which “you didn’t do it”, and therefore can theoretically be referred to as a D, but since Ds can’t be Scenes we mark it as an H.)

**Verbs that can be primary or secondary.** There are certain verbs that in some cases will function as secondary verbs (and therefore as Ds) and in other cases as primary verbs and this depends on the context in the specific scene under question.

1. “John<sub>A</sub> remembered<sub>D</sub> to<sub>F</sub> take<sub>P</sub> [the keys]<sub>A</sub>” (context-dependent, but it’s very likely that the “remembered” here is mostly for emphasis and therefore secondary)
2. “John<sub>A</sub> remembered<sub>P</sub> [ [the<sub>E</sub> hike]<sub>C</sub> ]<sub>P</sub> [with all his friends]<sub>A</sub> ]<sub>A</sub>”
3. “John<sub>A</sub> forgot<sub>P</sub> [ how<sub>D</sub> to<sub>F</sub> [ride]<sub>P</sub> [his bicycle]<sub>A</sub> ]<sub>A</sub>” (clearly the forgetting and the riding are not in the same time)

**Secondary predicates.** A depictive or resultative should be marked separately from the main predicate as an independent parallel Scene.

1. Depictives:

- “[John<sub>A</sub> left<sub>P</sub> home]<sub>A</sub> ]<sub>H</sub> [young<sub>S</sub> (John)<sub>A</sub> ]<sub>H</sub>”
- “[John<sub>A</sub> ate<sub>P</sub> [the food]<sub>A</sub> ]<sub>H</sub> [cold<sub>S</sub> (food)<sub>A</sub> ]<sub>H</sub>”
- “[He<sub>A</sub> left<sub>P</sub> [the party]<sub>A</sub> ]<sub>H</sub> [angry<sub>S</sub> (he)<sub>A</sub> ]<sub>H</sub>”

2. Resultatives:

- “[Mary<sub>A</sub> painted<sub>P</sub> [the fence]<sub>A</sub> ]<sub>H</sub> [blue<sub>S</sub> (fence)<sub>A</sub> ]<sub>H</sub>”
- “[He<sub>A</sub> [cried himself]<sub>P</sub> ]<sub>H</sub> to<sub>L</sub> [sleep<sub>P</sub> (He)<sub>A</sub> ]<sub>H</sub>”

**Seem/look/appear/sound/feel.** These verbs are often used without specifying the experiencer of the feeling/perception. In these cases they should be treated as a G. In case where the experiencer is stated, they should be a separate Scene.

1. “[The<sub>E</sub> coffee]<sub>C</sub> ]<sub>A</sub> seems<sub>G</sub> to<sub>F</sub> be<sub>F</sub> hot<sub>S</sub>”
2. “[The<sub>E</sub> car]<sub>C</sub> ]<sub>A</sub> looks<sub>G</sub> good<sub>S</sub>”
3. “It<sub>F</sub> seemed<sub>P</sub> [to<sub>R</sub> Mary]<sub>C</sub> ]<sub>A</sub> [that<sub>R</sub> [the<sub>E</sub> coffee]<sub>C</sub> ]<sub>A</sub> is<sub>F</sub> hot<sub>S</sub> ]<sub>A</sub>”
4. “It<sub>F</sub> appears<sub>G</sub> that<sub>F</sub> he<sub>A</sub> had<sub>F</sub> left<sub>P</sub> [the country]<sub>A</sub>”
5. “[Your dog]<sub>A</sub> looks<sub>G</sub> like<sub>F</sub> [a<sub>E</sub> cat]<sub>C</sub> ]<sub>S</sub>”



**Similes.** In most cases similes should be treated as separate Scenes:

1. “[He<sub>A</sub> eats<sub>P</sub>]<sub>H</sub> like<sub>L</sub> [a horse (*eats*)<sub>P</sub>]<sub>H</sub>”

An exception would be when the verb does not evoke a Scene of its own (e.g. “looks”, “seems”. See section 7.1 above) and is therefore considered a G. Then the whole phrase should be marked as one Scene:

1. “He<sub>A</sub> looks<sub>G</sub> like<sub>F</sub> [a horse]<sub>S</sub>”
2. “He<sub>A</sub> looks<sub>G</sub> like<sub>F</sub> he<sub>A</sub> just<sub>T</sub> saw<sub>P</sub> [a dinosaur]<sub>A</sub>”

**Cognitive events** Cognitive events (e.g. think, see, wonder) should be marked as Processes.

1. “I<sub>A</sub> see<sub>P</sub> [that you both are getting along]<sub>A</sub>”
2. “I<sub>A</sub> think<sub>P</sub> [it’s OK]<sub>A</sub>”
3. “I<sub>A</sub> wonder<sub>P</sub> [whether we’re doing a mistake]<sub>A</sub>”

**Results of Scenes.** Results of Scenes can be Scenes in their own right.

1. “[the<sub>E</sub> outcome<sub>C</sub>]<sub>P</sub> [[of<sub>R</sub> the<sub>E</sub> trial<sub>C</sub>]<sub>P</sub>]<sub>A</sub>”

**Scene Agents.** The agents of Scenes are Scenes in their own right. *NSS: This statement doesn’t make sense to me. I think you mean that deverbal nouns can simultaneously evoke a scene and be a participant in a larger scene? I don’t know that it’s necessarily about the agent role, because “the car drivers are old” would have the same analysis, right?*

1. “[car<sub>A</sub> drivers<sub>P</sub>]<sub>A</sub> are<sub>F</sub> diligent<sub>S</sub>” *NSS: Is it also possible to analyze this as saying that diligence is a manner of driving, thus one scene?*

**Noun Scenes.** A noun Scene is a case when a noun-phrase serves as a Scene and the noun itself is the main relation in the Scene (the P or S). They should be internally analyzed as Scenes, with a P/S, As, Ds and Ts. In general, deverbal nouns are cases of noun Scenes, although not all noun Scenes are formed by deverbal nouns.

1. “[John<sub>C</sub> ’s<sub>R</sub>]<sub>A</sub> accurate<sub>D</sub> kick<sub>P</sub>]<sub>A</sub> saved<sub>P</sub> [the game]<sub>A</sub>”
2. “[Him<sub>A</sub> destroying<sub>P</sub> [the city]<sub>A</sub> ]<sub>A</sub> was<sub>F</sub> [a<sub>E</sub> disaster<sub>C</sub>]<sub>S</sub>”
3. “[The<sub>E</sub> destruction<sub>C</sub>]<sub>P</sub> [of the city]<sub>A</sub> ]<sub>A</sub> was<sub>F</sub> [a<sub>E</sub> disaster<sub>C</sub>]<sub>S</sub>”
4. “[His]<sub>A</sub> destruction<sub>P</sub> of<sub>F</sub> [the city]<sub>A</sub> ]<sub>A</sub> was<sub>F</sub> [a disaster]<sub>S</sub>”
5. “[Gone with the Wind]<sub>A</sub> is<sub>F</sub> [one<sub>C</sub> of<sub>R</sub>]<sub>E</sub>]<sub>S</sub>– [Selznick<sub>C</sub> ’s<sub>R</sub>]<sub>A</sub> productions–<sub>S</sub>” *NSS: bracket mismatch*
6. “[War<sub>P</sub>]<sub>A</sub> is<sub>F</sub> imminent<sub>S</sub>’

**Scenes with neither a P nor an S.** Some Scenes have no P or S, since it is omitted or implied. In this case, we should add them as remote units.

1. “[John bought eggs]<sub>H</sub> and<sub>L</sub> [Mary<sub>A</sub> [chewing gum]<sub>A</sub> (*bought*)<sub>P</sub> ]<sub>H</sub>”
2. “[John<sub>A</sub> wanted<sub>P</sub> [a real life]<sub>A</sub> ]<sub>H</sub>, [not<sub>D</sub> [life in a caravan]<sub>A</sub> (*John*)<sub>A</sub> (*wanted*)<sub>P</sub>]<sub>H</sub>”
3. “[how about]<sub>S</sub> coffee<sub>A</sub>?”

**Adjective Scenes without a noun.** When we encounter an adjective Scene without a noun (e.g. ‘Amazing’, ‘Great’, ‘Splendid’) we have two options:

1. If we can detect the missing unit from the surrounding text we add it as a Remote unit.
2. If we are not sure what the missing Implicit unit is, we add an Implicit unit and mark the adjective as S.

**NSS:** Is there a role for Ground here/for interjections generally? If someone says “Amazing!” they are conveying their attitude.

- “[Amazing<sub>S</sub> ! (IMPLICIT) A]<sub>H</sub>”

**Imperatives.** Imperative clauses should be marked as a Scene, with an implicit A.

1. “Stop<sub>P</sub> (IMPLICIT) A !”
2. “Please<sub>F</sub> [take care]<sub>P</sub> [of your brother]<sub>A</sub> (IMPLICIT) A”

**Fragments.** Occasionally, a fragment of text does not describe a Scene in its own right, nor belongs to any other specific Scene. The category of such a unit, as always, is determined by its role in a super-ordinate relation it participates in (if any). By default, it’s an H.

1. “[No]<sub>H</sub>, [this will not stand]<sub>H</sub>”
2. “Thanks<sub>H</sub>”

**Expletive it.** Sometimes “it” is used to take the place of the subject when there is no other A which does so. In this case it should be marked as an E

1. “It<sub>F</sub> is<sub>F</sub> strange<sub>S</sub> [that<sub>R</sub> I<sub>A</sub> met<sub>P</sub> him<sub>A</sub> here<sub>A</sub>]<sub>A</sub>”
2. “It<sub>F</sub> is<sub>F</sub> likely<sub>S</sub> [to<sub>R</sub> rain<sub>P</sub>]<sub>A</sub>”

**Cooperating participants.** If two participants cooperatively participate in the same Process or perform it in an identical manner then they should be united in one A with two Cs. **NSS:** This only applies if they are coordinated, right? “John and Mary played tennis”, not “John played tennis with Mary.”

- “[John<sub>C</sub> and<sub>N</sub> Mary<sub>C</sub>]<sub>A</sub> went to the park“
- “A conversation was held [between<sub>R</sub> [[the Prime Minister]<sub>C</sub> and<sub>N</sub> [the Queen]<sub>C</sub>]<sub>C</sub>]<sub>A</sub>“

## 7.2 Remotes.

**Two Types of verbs that take a participant Scene.** Note that some verbs with a participant Scene have a remote unit taken from the participant Scene or vice versa. Other verbs do not exhibit such behavior.

1. “I<sub>A</sub> expected<sub>P</sub> [John<sub>A</sub> to<sub>F</sub> come<sub>P</sub>]<sub>A</sub>”
2. “We<sub>A</sub> agreed<sub>P</sub> [for John to give the funeral oration]<sub>A</sub>”
3. “I<sub>A</sub> persuaded<sub>P</sub> [John<sub>A</sub> to<sub>F</sub> come<sub>P</sub>]<sub>A</sub>”
4. “John<sub>A</sub> promised<sub>P</sub> [to<sub>F</sub> be<sub>F</sub> better<sub>S</sub> (John)<sub>A</sub>]<sub>A</sub>”

**Prominent Cases of Remote Units.** A remote unit is a unit that is referenced in a Scene in which it is not contained. By convention, the remote unit should be selected to be the minimal unit that refers to the target entity (for instance, “table” and not “the red table”). Several prominent cases of remote participants:

1. Coordination. The subject is often omitted.

(a) “[John<sub>A</sub> [had<sub>F</sub> dinner<sub>C</sub>] <sub>P</sub>]<sub>H</sub> and<sub>L</sub> [went<sub>P</sub> [to<sub>R</sub> bed<sub>C</sub>] <sub>A</sub> (John)<sub>A</sub>]<sub>H</sub>”

2. Relative Scenes. The head of the elaboration is not contained in the relative Scene.

(a) “[The<sub>E</sub> table<sub>C</sub> [I<sub>A</sub> ’m<sub>F</sub> using<sub>P</sub> (table)<sub>A</sub>]<sub>E</sub>]<sub>A</sub> is<sub>F</sub> too<sub>D</sub> short<sub>S</sub>”

3. Infinitives. The subject is omitted.

(a) “[Driving<sub>P</sub> [to<sub>F</sub> school<sub>C</sub>] <sub>A</sub> (John)<sub>A</sub>]<sub>A</sub> upsets<sub>P</sub> John”

(b) “[to<sub>F</sub> be<sub>F</sub> expected<sub>P</sub> [to<sub>F</sub> wash<sub>P</sub> [the<sub>E</sub> car<sub>C</sub>] <sub>A</sub> (Mary)<sub>A</sub>]<sub>A</sub> (IMPLICIT)<sub>A</sub>]<sub>A</sub> infuriates<sub>P</sub> Mary<sub>A</sub>”

### 7.3 Secondary Relations in Scenes.

1. Quantity Adverbs. Adverbs of quantity such as “just” and “only” should be annotated as Ds whenever possible.

(a) “There<sub>S</sub> is<sub>F</sub> only<sub>D</sub> [one piece of cake]<sub>A</sub>”

(b) “[The supermarket]<sub>A</sub> is<sub>F</sub> just<sub>D</sub> around<sub>S</sub> [the corner]<sub>A</sub>”

2. Negation. Negation is considered an adverbial.

(a) “John<sub>A</sub> did<sub>F</sub> n’t<sub>D</sub> touch<sub>P</sub> [the piano]<sub>A</sub>”

(b) “[John]<sub>A</sub> is<sub>F</sub> [no]<sub>D</sub> [joker]<sub>P</sub>”

Some pronouns and linkers express negation on a Scene. In this case, they also serve as Ds in that scene.

(a) “Nobody<sub>A+D</sub> came<sub>P</sub> [to<sub>R</sub> [the<sub>E</sub> party<sub>C</sub>] <sub>P</sub>]<sub>A</sub>”

(b) “[I<sub>A</sub> left<sub>P</sub>]<sub>H</sub> without<sub>L</sub> [eating<sub>P</sub> [my<sub>S</sub> banana<sub>A</sub>]<sub>A</sub> (I)<sub>A</sub> (without)<sub>D</sub>]<sub>H</sub>”

3. D in coordination. Occasionally, several entities are connected by an N, where there is a D (usually a frequency, probability or temporal relation) which relates specifically to one of them. In this case, the proper annotation is to annotate it as a D.

(a) “He<sub>A</sub> appeared<sub>P</sub> [in<sub>R</sub> [Head of the Class]<sub>C</sub>, [Freddy ’s Nightmares]<sub>C</sub>, [Thirtysomething]<sub>C</sub>, and<sub>N</sub> [( [for a second time]<sub>D</sub> ) [Growing Pains]<sub>C</sub> ]<sub>C</sub> ]<sub>A</sub>.”

(b) “John<sub>A</sub> is<sub>F</sub> intending<sub>D</sub> to<sub>F</sub> go<sub>P</sub> [to<sub>R</sub> [Rome<sub>C</sub>, Paris<sub>C</sub> and<sub>N</sub> [perhaps<sub>D</sub> London<sub>C</sub>]<sub>C</sub> ]<sub>C</sub> ]<sub>A</sub>”.

4. Secondary main verbs: Sometimes the Process appears as the subject of the sentence, where the main verb is the secondary verb. In these cases, we still mark the secondary verb as D, and the subject as the main relation.

(a) “[John<sub>C</sub> ’s<sub>R</sub>]<sub>A</sub> career<sub>P</sub> ended<sub>D</sub> abruptly<sub>D</sub>”

(b) “[The<sub>E</sub> race<sub>C</sub>]<sub>P</sub> began<sub>D</sub> [early<sub>E</sub> in<sub>R</sub> the<sub>E</sub> morning<sub>C</sub>]<sub>T</sub>”

(c) “His<sub>A</sub> service<sub>P</sub> was slow<sub>D</sub>”

5. Copulas: Some verbs in English may serve as copulas, mostly taking an adjective or adjectival phrases as arguments. These are also considered Ds:

(a) “John<sub>A</sub> stayed<sub>D</sub> awake<sub>S</sub> [all evening]<sub>T</sub>”

(b) “Mary<sub>A</sub> turned<sub>D</sub> ill<sub>S</sub>”

**Possession.** We distinguish between two types of possession: inalienable (body parts and relatives) and alienable (owning something). Inalienable possession does not evoke a Scene, while alienable possession does. Possession should not be confused with the relations of “Corresponding Profiles” and “Part-Whole Relations” (see below).

1. “John<sub>A</sub> ’s<sub>S</sub> car<sub>A</sub>” (alienable)
2. “[John<sub>C</sub> ’s<sub>R</sub>]<sub>E</sub> hand<sub>C</sub>” (inalienable)
3. “[John<sub>C</sub> ’s<sub>R</sub>]<sub>E</sub> brother<sub>C</sub>” (inalienable)

### NSS: possessive pronouns?

Possession constructions can also be used to express other relations, the range of which is language-dependent.

1. “This<sub>E</sub> seat<sub>C</sub> [of<sub>R</sub> the<sub>E</sub> Knesset<sub>C</sub>]<sub>E</sub>” (not a Scene)

In the case of non-familial relations between two animates (unlike familial ones), we do consider the unit as a Scene:

1. “[John<sub>A</sub> ’s<sub>S</sub> (*employee*)<sub>A</sub>]<sub>E</sub> employee<sub>C</sub>”
2. “[John<sub>A</sub> ’s<sub>S</sub> (*friend*)<sub>A</sub>]<sub>E</sub> friend<sub>C</sub>”

**Framing of Scenes.** Some Scenes are wrapped in a complex preposition that frames them (e.g., “story of”, “rumor of”, “belief that”). In this case, the framing noun serves a separate Scene, which takes the framed Scene as a Participant.

1. “[the<sub>E</sub> story<sub>C</sub>]<sub>P</sub> [of<sub>R</sub> [a<sub>E</sub> young<sub>E</sub> girl<sub>C</sub>]<sub>A</sub> sentenced<sub>P</sub> [to<sub>R</sub> death<sub>P</sub> (*girl*)<sub>A</sub>]<sub>A</sub>]<sub>A</sub>”
2. “[the<sub>E</sub> rumor<sub>C</sub>]<sub>P</sub> [of<sub>R</sub> his<sub>A</sub> retirement<sub>P</sub>]<sub>A</sub>”
3. “[the<sub>E</sub>]<sub>P-</sub> strange<sub>D</sub> [belief]<sub>-P</sub> [that<sub>R</sub> chickens<sub>A</sub> are<sub>F</sub> immortal<sub>S</sub>]<sub>A</sub>”

## 7.4 Non-Scene Units.

**Determiners.** Determiners should be annotated as elaborators of the noun.

1. “The<sub>E</sub> Knesset<sub>C</sub>”
2. “A<sub>E</sub> big brown dog<sub>C</sub>”

**Distinguishing between Determiners and Prepositions.** Prepositions (which usually function as Relators) relate between two entities while Determiners (which usually function as Elaborators) merely mark the noun phrase.

1. It is rare to find [people with<sub>R</sub> such<sub>E</sub> wonderful personality]

**Appositions.** Appositions are cases where two consecutive units are semantically parallel and refer to the same entity. If one is a proper name and the other isn’t, the first is the C, and the second is the E.

1. “John<sub>C</sub>, [my history teacher]<sub>E</sub>” (apposition; **This should be an E-Scene; move**)

**Extraposition.** Cases where an E does not create a contiguous stretch of text with its center. In this case, they should be marked together as a non-contiguous unit.

1. “He saw [that painting]<sub>A-</sub> before, [[that lovely magnificent painting]<sub>E</sub>]<sub>-A</sub>”
2. “I met [the guy]<sub>A-</sub> yesterday, [[whom I first saw in the park]<sub>E</sub>]<sub>-A</sub>”

**Fused E Scenes.** There are many constructions that resemble an E Scene construction, but have a somewhat different form (they don't have a clear Center). Their internal structure should look like that of a Scene:

1. “[What<sub>A</sub> I<sub>A</sub> meant<sub>P</sub>]<sub>A</sub> was<sub>S</sub> [I want to have dinner]<sub>A</sub>”
2. “[Any<sub>E</sub> recipes<sub>C</sub> [she<sub>A</sub> used<sub>P</sub> (recipes)<sub>A</sub>]<sub>E</sub>]<sub>A</sub> are<sub>F</sub> marked<sub>P</sub> [in<sub>R</sub> red]<sub>C</sub>]<sub>D</sub>”
3. “you<sub>A</sub> are<sub>F</sub> playing<sub>P</sub> [with<sub>R</sub> somebody<sub>A</sub> better<sub>S</sub> [than<sub>R</sub> you<sub>C</sub>]<sub>A</sub>]<sub>A</sub>”

**Numbers and Quantifiers.** They are considered Qs. The question of their scope is not addressed in the current layer of the annotation. Therefore they are considered a part of the unit adjacent to it.

1. “[All<sub>Q</sub> Greeks]<sub>C</sub>]<sub>A</sub> are<sub>F</sub> mortals<sub>S</sub>”
2. “[Two<sub>Q</sub> bananas]<sub>C</sub>]<sub>A</sub> are<sub>F</sub> lying<sub>P</sub> [on<sub>R</sub> the<sub>E</sub> table]<sub>C</sub>]<sub>A</sub>”
3. “Millions<sub>Q</sub> of<sub>R</sub> homes<sub>C</sub>”

**Single-word Scenes within Elaborators.** In some cases, we first analyze a noun-phrase with Center-Elaborator relations, and then internally annotate one of its units as a Scene. If that Scene is a one-word Scene which has a remote unit that is part of the same noun-phrase, we are not required to add it (it will be filled in automatically) **OA: this is something we may want to rethink. OK, should say that it really helps in annotation**

- “[Great<sub>S</sub>]<sub>E</sub> [man]<sub>C</sub>” (in principle, “man” should have been added as a Remote unit to the E unit, but in practice for convenience we don't add it)
- “[Brown<sub>S</sub>]<sub>E</sub> [dog]<sub>C</sub>” (dog should have been added as a Remote A, but can be omitted for brevity)

**Comparatives/Superlatives.** Comparatives/superlatives generally evoke a static Scene. If the domain of application is explicitly mentioned (namely the set of entities the comparison applies to), it should be marked as a participant.

1. “[Jordan]<sub>A</sub> was<sub>F</sub> better<sub>S</sub> [than<sub>R</sub> James]<sub>C</sub>”
2. “[Jordan]<sub>A</sub> was<sub>F</sub> [more<sub>E</sub> beautiful]<sub>C</sub>]<sub>S</sub> [than<sub>R</sub> James]<sub>C</sub>”
3. “[China]<sub>A</sub> is<sub>F</sub> [the<sub>E</sub> greatest]<sub>C</sub>]<sub>S</sub>”
4. “[China]<sub>A</sub> is<sub>F</sub> [the<sub>E</sub> greatest]<sub>C</sub>]<sub>S</sub> [place on earth]<sub>A</sub>”

**Directions.** Directions should be considered as As, as they can be said to refer to an abstract location. This applies to both absolute directions (like “north”) and relative directions (like “away”). **NSS: not sure I find this intuitive; they feel more adverbial to me. sometimes there's a clear location that can be inferred in context (“she came in/out”—presumably we have a reference point from context) but not always. E.g., “the bird flew up” can simply mean the bird is ascending, without an implicit source or goal.**

1. “John<sub>A</sub> told<sub>P</sub> Mary<sub>A</sub> [to<sub>F</sub> come<sub>P</sub> in<sub>A</sub> ]<sub>A</sub>”
2. “John<sub>A</sub> walked<sub>P</sub> away<sub>A</sub>”
3. “They danced<sub>P</sub> [the night]<sub>T</sub> away<sub>D</sub>” (a non-literal use)

**Passive “by”.** The “by” of the passive should be annotated as R.

1. “He<sub>A</sub> is<sub>F</sub> scolded<sub>S</sub> [by<sub>R</sub> many<sub>C</sub>]<sub>A</sub>”

**NSS: What if there's no by-phrase: is it implicit?**

**Preposition Stranding.** In some cases, an A is missing but its preposition is in place. We mark the preposition as an A, with an R inside of it, and add a remote C:

1. “The<sub>E</sub> book<sub>C</sub> [I<sub>A</sub> 'm<sub>F</sub> looking<sub>P</sub> [for<sub>R</sub> (book)<sub>C</sub> ]<sub>A</sub>]<sub>E</sub>”
2. “The<sub>E</sub> work<sub>C</sub> [I<sub>A</sub> [pay<sub>F</sub>]<sub>P</sub>– most<sub>D</sub> [attention<sub>C</sub>]<sub>–P</sub> [to<sub>R</sub> (work)<sub>C</sub>]<sub>A</sub>]<sub>E</sub>”

## 7.5 Processes/States.

**Modals and Auxiliaries.** Modals should invariably be annotated as secondary verbs (and therefore as Ds). This applies to “would” as well. Auxiliary verbs (“be”, “have”, “will” and “do”), which do not have significant semantic input in their own right<sup>4</sup> are considered Fs.

1. “John<sub>A</sub> will<sub>F</sub> come<sub>P</sub>”
2. “Mary<sub>A</sub> should<sub>D</sub> come<sub>P</sub>”
3. “Mary<sub>A</sub> is<sub>F</sub> coming<sub>P</sub>”
4. “John<sub>A</sub> [has to]<sub>D</sub> come<sub>P</sub>”
5. “I<sub>A</sub> have<sub>F</sub> done<sub>P</sub> it<sub>A</sub>”
6. “John<sub>A</sub> does<sub>F</sub> n't<sub>D</sub> know<sub>P</sub> him<sub>A</sub>”
7. “se disposait” *fr*

**Distinguishing identity and other static Scenes.** Occasionally nouns are used as Ps or Ss, accompanied by some inflection of the verb “be”. UCCA distinguishes between two cases:

1. Identity: where there are two separate, defined entities. **NSS: they're not separate entities in the world. separate referring expressions that are asserted to have the same referent?**
  - (a) “[The morning star]<sub>A</sub> is<sub>S</sub> [the evening star]<sub>A</sub>”
  - (b) “[That person over there]<sub>A</sub> is<sub>S</sub> John<sub>A</sub>”
2. Attribution: where there is one A, and the noun is used to describe some set of elements to which that A belongs to. In this case, we include the set-denoting noun in the P or S.
  - (a) “[his speech]<sub>A</sub> was<sub>F</sub> [an embarrassment]<sub>S</sub>”
  - (b) “[Brad Pitt]<sub>A</sub> was<sub>F</sub> [a slimeball]<sub>S</sub> [as the boyfriend in the soap opera Dallas]<sub>A</sub>”
  - (c) “John<sub>A</sub> is<sub>F</sub> [[six<sub>E</sub> years]<sub>C</sub>]<sub>E</sub> old<sub>C</sub>]<sub>S</sub>”

**NSS: might be good to have a section on various uses of the copula in English. identity, attribution as well as focus constructions, existentials, ...**

---

<sup>4</sup>UCCA in its foundational layer does not annotate tense. Even if it did, the tense would not be considered a feature encoded on the auxiliaries, but rather in the combination of the auxiliary and the inflection.

**Infinitive “to”.** By convention, when “to” is used as an F (same for “zu” in German), it should not be included within the process/state.

1. “He wanted<sub>D</sub> to<sub>F</sub> come<sub>P</sub> home<sub>A</sub>”
2. “[to<sub>F</sub> kick<sub>P</sub> [a penalty shot]<sub>A</sub> [in soccer]<sub>A</sub> (IMPLICIT)<sub>A</sub>] <sub>A</sub> is<sub>F</sub> exciting<sub>S</sub>”

**Light Verbs.** Cases where the verb is almost void of meaning, and most of the meaning is determined by the object. The verb is usually “have”, “give”, “take” or “make” (although there are other examples). Annotation: both the the light verb and the following object should be included inside the P/S. The light verb as an F and the object as a C.

1. “John<sub>A</sub> [took<sub>F</sub> a<sub>E</sub> shower<sub>C</sub>]<sub>P</sub>”
2. “Mary<sub>A</sub> [gave<sub>F</sub>]<sub>P-</sub> John<sub>A</sub> [a<sub>E</sub> smile<sub>C</sub>]<sub>-P</sub>”
3. “Brad<sub>A</sub> [made<sub>F</sub> a<sub>E</sub> guest<sub>E</sub> appearance<sub>C</sub>]<sub>P</sub> [on<sub>R</sub> ABC<sub>C</sub>]<sub>A</sub>”

**Possessive “have”:** Whenever “have” carries the semantic meaning of ownership and precedes a concrete object (e.g. book, pen), it should be marked an S.

1. “John<sub>A</sub> has<sub>S</sub> [a book]<sub>A</sub>”

But whenever “have” is part of a phrase describing an action or event, then it should be marked as a “light verb”.

1. John<sub>A</sub> [had<sub>F</sub> a<sub>E</sub> walk<sub>C</sub>]<sub>P</sub> yesterday<sub>T</sub>”
2. John<sub>A</sub> [has<sub>F</sub> problems<sub>C</sub>]<sub>S</sub>
3. John<sub>A</sub> [has<sub>F</sub> hobbies<sub>C</sub>]<sub>S</sub>

**Adjective followed by a Scene:** Analyzed as a D+P construction.

1. “John<sub>A</sub> is<sub>F</sub> easy<sub>D</sub> to<sub>F</sub> please<sub>P</sub>”
2. “John<sub>A</sub> is<sub>F</sub> likely<sub>D</sub> to<sub>F</sub> leave<sub>P</sub>”
3. “John<sub>A</sub> is<sub>F</sub> ready<sub>D</sub> to<sub>F</sub> come<sub>P</sub>”
4. “London<sub>A</sub> is<sub>F</sub> great<sub>D</sub> for<sub>F</sub> music<sub>P</sub>”

**Causatives.** We view the causation word (often “make” or “cause”) construction as a secondary verb.

1. “John<sub>A</sub> makes<sub>D</sub> Mary<sub>A</sub> happy<sub>S</sub>”
2. “John<sub>A</sub> inspires<sub>D</sub> interest<sub>P</sub> [in<sub>R</sub> Mary<sub>C</sub>]<sub>A</sub>”
3. “We just got<sub>D</sub> [our sunroom]<sub>A</sub> built<sub>P</sub> by Patio World”
4. “Mary had<sub>D</sub> [her hair]<sub>A</sub> done<sub>P</sub>”

**Secondary Verbs with an additional role.** Some secondary verbs introduce another role beside the roles of the main verb. An example is “help”, “force” and “permit”. Like all secondary verbs, such verbs are considered Ds. The additional participant is marked as an A in the Scene.

1. “John<sub>A</sub> helped<sub>D</sub> Mary<sub>A</sub> climb<sub>P</sub> [the ladder]<sub>A</sub>”
2. “John<sub>A</sub> forced<sub>D</sub> [Mary]<sub>A</sub> to<sub>F</sub> climb<sub>P</sub> [the ladder]<sub>A</sub>”
3. “he<sub>A</sub> is<sub>F</sub> guilty<sub>D</sub> of<sub>F</sub> not<sub>D</sub> cleaning<sub>P</sub> [the dishes]<sub>A</sub>”

**Polite Forms.** Words that only serve as part of a construction for politely addressing someone are Fs.

1. “Gehen<sub>P</sub> Sie<sub>F</sub> raus<sub>D</sub> !”
2. “[[Sie und Ihr komischer Vogel]<sub>G+A</sub>, raus<sub>P</sub>] !” [“you and your funny bird, out!”] (here “Sie” is part of the vocative)
3. “Gehen<sub>P</sub> [Sie<sub>C</sub> und<sub>N</sub> Hans<sub>C</sub>] raus<sub>D</sub> !”<sup>de</sup>

## 7.6 Other Relations.

**Punctuation.** Not annotated in the current layer of UCCA (even commas).

**Linkers with a single argument.** We also allow Ls with a single argument. This usually happens if an L relates one Scene with everything that follows/precedes it, without there being any particular unit that the Scene relates to. Another case where we use a single argument linker is when one of its arguments is omitted. An example would be a paragraph that starts with “However” that contrasts with everything that was written in the previous paragraph.

**Elaboration of/by a Coordination.** When a certain unit relates to multiple units that carry an identical role, we unify all the multiple units under one parent unit.

- “I have [10<sub>Q</sub> [brothers<sub>C</sub> and<sub>N</sub> sisters<sub>C</sub>]<sub>C</sub>]<sub>A</sub>”
- “Queen<sub>C</sub> of<sub>R</sub> [England<sub>C</sub> and<sub>N</sub> Canada<sub>C</sub>]<sub>E</sub>”
- “I may have forgotten my keys [on<sub>R</sub> [[the table]<sub>C</sub> or<sub>N</sub> couch<sub>C</sub>]<sub>C</sub>]<sub>A</sub>”

**Vocatives.** Vocatives should be considered as Ground, as they are exclusively part of the speech event Scene. If the Participant is not mentioned otherwise, add it as a remote

1. “[John<sub>G</sub>, who<sub>A</sub> is<sub>F</sub> this<sub>A</sub> ?]<sub>H</sub>”
2. “[John<sub>G</sub>, go<sub>P</sub> outside<sub>A</sub> (John)<sub>A</sub>]<sub>H</sub>” (In this case, John is also a Participant in the Scene and therefore is added as a Remote)
3. “[Nein<sub>G</sub>, Herr Kapitan]<sub>H</sub>”<sup>de</sup>

**Titles.** By convention, titles of names are considered Elaborators of the proper name.

1. “I<sub>A</sub> can<sub>D</sub> ’t<sub>D</sub> find<sub>P</sub> [Captain<sub>E</sub> Nemo<sub>C</sub>]<sub>A</sub>”
2. “[Queen<sub>E</sub> Mary<sub>C</sub>]<sub>A</sub> went<sub>D</sub> to<sub>F</sub> sleep<sub>P</sub>”



**Focus Constructions.** Some constructions are used to emphasize one specific entity. These distinctions are generally not treated in this layer of annotation and are therefore Fs. The difference between the examples below and existential “there” sentences is that here the main relation is not merely the existence of some entity.

1. “There<sub>F</sub> are<sub>F</sub> [lots<sub>Q</sub> of<sub>R</sub> them<sub>C</sub>]<sub>A</sub> lurking<sub>P</sub> [in<sub>R</sub> the<sub>E</sub> bushes<sub>C</sub>]<sub>A</sub>”
2. “It<sub>F</sub> was<sub>F</sub> John<sub>A</sub> who<sub>F</sub> wrote<sub>P</sub> [this<sub>E</sub> novel<sub>C</sub>]<sub>A</sub>”
3. “There<sub>F</sub> are<sub>F</sub> earrings<sub>A</sub> on<sub>S</sub> [the<sub>E</sub> table<sub>C</sub>]<sub>A</sub>”
4. “Es<sub>F</sub> gibt<sub>F</sub> Ringe<sub>A</sub> auf<sub>S</sub> [dem<sub>E</sub> Tisch<sub>C</sub>]<sub>A</sub>”<sup>de</sup>

NSS: what about presentational sentences: “here is my card”, “this is a dax”?

**Question Words.** Question words should be annotated with the same category as their respective component in a given answer.

1. “How<sub>D</sub> did you fix your car?”
2. “Who<sub>A</sub> shot the sheriff?”
3. “[Which<sub>E</sub> car<sub>C</sub>]<sub>A</sub> did you buy?”
4. “Why<sub>H</sub> haven’t you called me?”
5. “When<sub>T</sub> will they arrive?”

The same applies to indirect questions:

1. “Tell<sub>P</sub> me<sub>A</sub> [what<sub>A</sub> happened<sub>P</sub>]<sub>A</sub>”
2. “I<sub>A</sub> wonder<sub>P</sub> [where<sub>A</sub> he is going]<sub>A</sub>”

Some of these words can also be used as Relative pronouns. In such cases they are not interrogative but merely relate the E Scene with the elaborated entity, so they should be marked as Rs.

1. “the<sub>E</sub> man<sub>C</sub> [who<sub>R</sub> was<sub>F</sub> n’t<sub>D</sub> there<sub>S</sub> (*man*)<sub>A</sub>]<sub>E</sub>”
2. “the<sub>E</sub> tiger<sub>C</sub> [which<sub>R</sub> lost<sub>P</sub> [his<sub>S</sub> hair<sub>A</sub>]<sub>A</sub> (*tiger*)<sub>A</sub>]<sub>E</sub>”
3. “the<sub>E</sub> city<sub>C</sub> [ [in which]<sub>R</sub> John<sub>A</sub> lives<sub>P</sub> (*city*)<sub>A</sub>]<sub>E</sub>”

**Non-contiguous Linkers.** Sometimes a linkage relation is expressed by several words, which are not contiguous in the text, but evoke a single relation. We mark them by convention as two separate linkers and not as a non-contiguous unit.

1. “[Either]<sub>L</sub> you buy it [or]<sub>L</sub> you don’t”

**Dates and Names.** Dates and names are treated as unanalyzable. Therefore, no sub-units should be annotated:

1. “I live [in<sub>R</sub> [New York]<sub>C</sub>]<sub>A</sub>”
2. “The event took place [on<sub>R</sub> [May 17th, 1832]<sub>C</sub>]<sub>T</sub>”

NSS: What about a date with internal grammar: “the 17th of May”?

**Reflexives.** Reflexives are the words that (in their primary sense) state that two participants of an event are one and the same (“himself”, “themselves”, “to one another” etc.). In UCCA, we mark them as part of the P/S, which is considered unanalyzable. Note, however, that in some cases reflexives are not used in their primary sense. In these cases, they should be analyzed according to their meaning in the context.

1. “John<sub>A</sub> [washed himself]<sub>P</sub>”
2. “Mary<sub>A</sub> [talked herself]<sub>P</sub> [into coming]<sub>A</sub>”
3. “[He<sub>C</sub> himself<sub>F</sub>]<sub>A</sub> spoke<sub>P</sub> [to the manager]<sub>A</sub>.” (“himself” here does not introduce a participant, but rather emphasizes that it was “he” and not someone else)
4. “He did it [all<sub>E</sub> [by<sub>R</sub> himself<sub>C</sub>]<sub>C</sub>]<sub>D</sub>” (it’s a D since the expression basically means that he did it alone)
5. “John<sub>A</sub> [relieved himself]<sub>P</sub> [in<sub>R</sub> the<sub>E</sub> backyard<sub>C</sub>]<sub>A</sub>”
6. “John<sub>A</sub> [established himself]<sub>P</sub> [as<sub>R</sub> a<sub>E</sub> lecturer<sub>C</sub>]<sub>A</sub>”
7. “John hat<sub>F</sub> [sich gewaschen]<sub>P</sub>”
8. “[Studieren]<sub>P</sub>]<sub>A</sub> [lohnt sich]<sub>P</sub>” *de*

**Complex Prepositions.** Some prepositions are multi-worded. They should be annotated as complex units (or as unanalyzable if they have no parts with significant semantic input). *In German this could be “auf Grund”, “an der Seite von”, “des Weiteren” etc.*

1. “[According to]<sub>S</sub> John<sub>A</sub>, [ [the<sub>E</sub> soup<sub>C</sub>]<sub>A</sub> is<sub>F</sub> salty]<sub>S</sub>]<sub>A</sub>”
2. “Mary<sub>A</sub> is<sub>F</sub> [on top of]<sub>S</sub> [this<sub>E</sub> task<sub>C</sub>]<sub>A</sub>”
3. “[ [later in]<sub>R</sub> 1988<sub>C</sub>]<sub>T</sub>, John<sub>A</sub> bought<sub>P</sub> [a<sub>E</sub> car<sub>C</sub>]<sub>A</sub>”

**Frame of reference.** Some Scenes are accompanied by a background statement which explains in what way it should be construed. If the background does not refer to the same event as the Scene itself, it should be treated as a separate Scene.

1. “[Under European law]<sub>H</sub>, [this is an infringement]<sub>H</sub>”
2. “Historically<sub>H</sub>, [governments favored city dwellers]<sub>H</sub>”
3. “[According to]<sub>L</sub> [the<sub>E</sub> figures<sub>C</sub>]<sub>H</sub>, [you lost a lot of money]<sub>H</sub>”

**Several Coordinated Predicates.** When encountering several coordinated predicates, each predicate should be annotated as an independent scene.

1. “[John is [a businessman]<sub>S</sub>]<sub>H</sub>, [politician<sub>S</sub> (John)<sub>A</sub>]<sub>H</sub> and<sub>L</sub> [Author<sub>S</sub> (John)<sub>A</sub>]<sub>H</sub>”
2. “[John<sub>A</sub> wrote<sub>P</sub> (song)<sub>A</sub>]<sub>H</sub>, [recorded (John)<sub>A</sub> (song)<sub>A</sub>]<sub>H</sub> and<sub>L</sub> [performed<sub>P</sub> [the<sub>E</sub> song<sub>C</sub>]<sub>A</sub> (John)<sub>A</sub>]<sub>H</sub>”

## 7.7 Morphology.

**Inflectional and Derivational Morphology.** UCCA does not annotate them in the current layer. Therefore the word “dogs” has no sub-units and neither does the word “talked”. This will be added in future layers.

**Coersed Word/Phrase.** Several words that were coersed into one and obtained their own idiosyncratic meaning. In this layer of UCCA they should be analyzed as a single unit, without sub-units.

1. “There are pickpockets<sub>A</sub> in this side of town”
2. “he<sub>A</sub>’s<sub>F</sub> [a<sub>E</sub> have-been<sub>C</sub>]s”
3. “Let’s<sub>D</sub> go<sub>P</sub> [to<sub>R</sub> the<sub>E</sub> merry-go-round<sub>C</sub>]A”

## 8 Chapter E: Criteria for compound splitting in German

Some of the examples are adapted from Schulte im Walde et al., 2016. This section is co-authored with Jakob Prange and Nathan Schneider.

### Criterion 1: Is the compound semantically transparent or opaque?

1. Split transparent compounds.
  - The meaning of *Ahornblatt* (maple leaf) can be derived from the meanings of *Ahorn* (maple) and *Blatt* (leaf).
2. Don’t split opaque compounds.
  - The meaning of *Maulwurf* (mole) cannot be derived from the meanings of *Maul* (mouth of an animal) and *Wurf* (throw).
3. Don’t split partially/asymmetrically transparent compounds.
  - The meaning of *Zeitungssente* (newspaper hoax) cannot be derived from the meaning of *Ente* (duck), but it can be derived from the meaning of *Zeitung* (newspaper).
  - *Murmeltier* (marmot) is a *Tier* (animal) but it does not involve either the noun *Murmel* (marble) or the verb *murmeln* (murmur).
  - *Sonnenkönig* (“Sun King”, aka King Ludwig XIV) is a *König* (king), but it doesn’t involve a *Sonne* (sun). It’s more of a name, and hence should not be split.
  - *Geduldsfaden* (thread of patience) refers to the extent of one’s patience, but doesn’t involve a thread. Note that this is different from the metaphorical use of *Faden* (thread) as part of a conversation. Also, you cannot paraphrase it with *Faden der Geduld*, cf. Criterion 2.
  - *Schriftzug* (logo) refers to something written (*Schrift* = writing), but it doesn’t have to be an actual hand movement *Zug* (stroke) anymore, although it is derived from that originally.

### Criterion 2: Can the compound be paraphrased as a noun phrase with the same noun head?

If it can be paraphrased, it should be split.

1. *Kaufleute* (salesmen) → *Leute, die kaufen und verkaufen* (people that buy and sell).
2. *Kinderbuch* (children’s book) → *ein Buch für Kinder* (a book for children)
3. *spindelförmig* (spindle-shaped) → *hat die Form einer Spindel* (has the shape of a spindle)

**Note:** Even if the head of the compound is a metaphor, if the same metaphor can be used in a paraphrase, the compound is considered compositional and should be split: *Bergkette* → *eine Kette von Bergen* (a chain of mountains), even though it’s not an actual chain, but rather a chain-like arrangement of mountains.

**Criterion 3: Is the pattern of the compound productive? That is, can one or both of the words of the compound be altered, while retaining a similar meaning?**

1. If it is, it should be split.

- *Fruchtsaft*, *Apfelsaft*, *Orangensaft* (types of juice)
- *Schiffsherr* (ship owner), *Haus herr* (house owner)
- *Braunbär*, *Schwarzbär*, *Grizzlybär* (different species of bears); BUT: *Waschbär* (raccoon), *Armeisenbär* (anteater) should not be split.
- *Gebirgszug* (mountain range), *Siegeszug* (triumphal march), *Vogelzug* (bird migration) are all related, BUT: *Schriftzug* (logo) doesn't have much to do with the above compounds and should not be split.

2. Where one of the words of the compound cannot be used as a free word, or has a very different meaning when used that way, it should not be split.

- *Uhrwerk*, *Fachwerk*, *Triebwerk*, *Schuhwerk*, *Blattwerk* are all related, BUT *Werk* is an opus, a piece of art or a factory and therefore should not be split (borderline).

## 9 Chapter F: Possible Post-processing Notes

- In E Scenes, put the Cs elaborated on in the Elaborator Scene.
- In Ground, extract the G from the Scene they are positioned in, and add a root node whose children are the G and the Scene.
- Flag: turn all the Rs into Fs, especially if a PSS layer is included.
- Include the determiners within the main relation if they are in an A-Scene noun phrase.
- Possessive pronouns should be S+A
- Negative polarity relators (without, neither) should be annotated both as negation and as L/R.

## 10 Chapter G: Plain Text Notation

In order to make UCCA's annotation legible and standardized, we give here guidelines for UCCA's notation in plain text. We note that the hierarchical structure formed by UCCA can be annotated by standard bracketing. The abbreviation of the category should be either adjacent to the left or to the right side of the category. For example, annotating the word "apple" with the category X should look like "[X apple]" or "[apple X]". We use the following abbreviations for the categories:

T – time  
Q – quantifier  
H – parallel Scene  
A – participant  
C – center  
L – linker  
D – adverbial  
E – elaborator

G – ground  
S – state  
N – connector  
P – process  
R – relator  
F – function

**Non-contiguity:** We use a dash to indicate a continuation of a unit. For example, if "word1 ... word2" is a non-contiguous unit then we mark it "[X- word1] [Y] [Z] [W] [-X word2]".

"[John A] [P- took] [Mary A] [up on -P] [ [her A] [promise P ] A]"

In case there are two non-contiguous units nested within one another, and of the same category, we may use indices to disambiguate. For example, in the sequence "w1 w2 w3 w4 w5", if "w1 ... w4" is a non-contiguous unit of category X and "w2 ... w5" is also a non-contiguous unit of category X, we mark it "[X1- w1] [X2- w2] w3 [-X1 w4] [-X2 w5]".

**Remote Units:** We place the Remote unit inside its Parent unit at the end of the phrase in round brackets and assign it with the relevant category:

- "John got home and [took a shower (John A)]"

**Implicit Units:** Implicit units are marked much like remote units, the only difference is that we add a fixed expression "IMPLICIT" inside the round brackets.

- "[Not going there any more (IMP A)]" (Who is not going there is implicit)

## 11 Caveats and Future Work

- Annotation of redundant elements
- Suggestion engine, including pre-marking of Hs by heuristics.