Propositional Framing coding schema

Yukinori Kimoto (University of Hyogo)
Norikazu Kogura (Tokyo University of Foreign Studies)
Nicholas Evans (The Australian National University)
Danielle Barth (The Australian National University)

Please cite as:
Kimoto, Yukinori, Norikazu Kogura, Nicholas Evans, and Danielle Barth (2022). Propositional framing coding schema. http://scopicproject.run-the-task.wordpress.com/2022/01/coding-schema-propositional-framing.pdf. Accessed on [insert date here].

This coding schema is designed to investigate the typological variations for complement clause and its alternative strategies across languages. The approach taken here is *onomasiological* or function-based: we first identify certain functional domains, and annotate how the domain is encoded in discourse. Our focus is thus not on structural variations in complement clause per se. We are concerned with preferences for complement clause in individual languages, by examining the relative frequency between complement clause and alternative strategies such as parataxis, and non-complement subordination.

1. Preparation in ELAN

Parent Tier: create "Comp@speaker"
(Linguistic Tier type: create "Comp")

Note tier: create "Comp_Notes@speaker"
(Linguistic Tier type: use the existent "Variable Notes")

2.1 What are coded and not coded

The expressions that fit either (i) or (ii) should be coded.

(i) a. the expression fits one of the functional types (e.g. UTT, THINK, IMM, see Table 1)
b. then the content is a proposition, and realized as such in the surface structure.

Included: I told her that she should start the project. (finite)
Included: *I saw **him swimming*** (nonfinite)
Included: *I saw **his swimming**.* (nonfinite, nominalization)
Included: *She said "Oh no!"* (the speech complement can be fragmentary, but this must be indicated in the notes)

Excluded: *He told **it** to me.* (pronoun)
Excluded: He was talking loudly. (no content proposition)
Excluded: *I saw **the guy**.* (perceiving a thing)
Excluded: *Who's this? **Maybe his mother**?* (not a complete proposition, so not tagged)

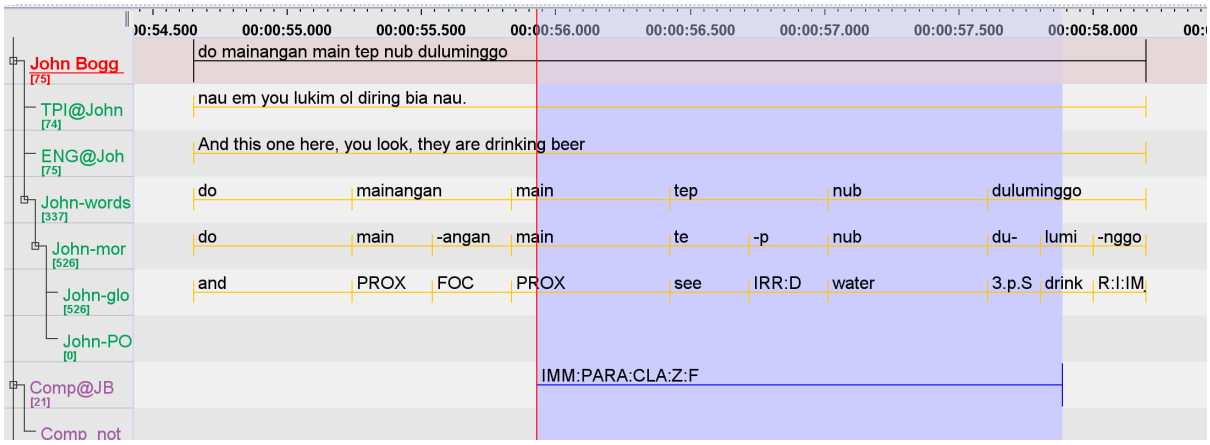
(ii) The content proposition is standing alone without framing elements.
Included: *He came in. "**Hey, I'm home!**"*. (no framing elements)

Table 1 Propositional Framing Coding schema (each linked with ":")

functional type (§3)	overall construction (§4)	framing element (§5)	connective (§6)	content proposition (§7)
UTT (utterance) PROB (probability judgement) THINK (thinking) PRET (pretence) COM (commentative) KNO (knowledge) FEAR (fear) DES (desiderative) IMM (immediate perception)	INDP (independent) COMP (complement) FUSE (fused) SUB (other subordinate propositions) ADV (adverbial frame) PARA (paratactic) PRTH (parenthetical) COORD (coordination)	CL (clause) CLA (clause with endophoric) NP (nominal) FRAG (fragment) INFL (inflection) PART (particle, adverbial) Z (zero)	C (connective) Z (zero)	For speech and thought D (direct) IF (indirect finite) INF (indirect nonfinite) N (neutral: indirect or direct) XE (compressed) For other content NF (nonfinite) F (finite)

2.2 How to Code

Use tags without terms. Each slot of the tag must be filled with an element. Elements are separated by colons. There are 5 slots for different aspects of the proposition, its frame and associated elements as seen in Table 1 above.



3. Functional type

This parameter specifies the function type of framing elements. Many categories follow Noonan (1985), but with additions being probability judgment (PROB) -similar to Noonan's ATT, and thought (THINK)- a subtype of Noonan's ATT.

For some categories, one might encounter causative or passive versions (e.g. 'remind' to 'remember', or 'be seen' to 'see'); these should be housed under the corresponding basic type (e.g. 'remind' would go under the KNO set. The same goes for lexicalised causatives like 'show' (i.e. cause to see) which go under the IMM set.

It can be difficult to distinguish functions such as THINK and PROB. In some languages these are clearly distinct (e.g., verb of thought v particle of probability) but often the same form serves both functions. Use what you can to decide, e.g., maybe the person and the TAM values give a clue. Also consider: could this be treated as D/I etc? If not, maybe it should be a PROB or a KNO or similar.

UTT (Utterance) expresses utterance and communication with (primarily) verbs or related nouns e.g. say, tell, report, ask, promise, news, story

–John said that Norm left. **UTT**:COMP:CL:C:IF
–Nell asked if Norm left.
–I heard [the news/story that Norm fell down into the river]. **UTT**:COMP:NP:C:F

PROB (Probability assessment) expresses speaker assessment of the likelihood of a state of affairs, e.g. reckon, it may be that

– I reckon he'll be home by now **PROB**:COMP:CL:Z:F
– It may be that he's unhappy with their decision **PROB**:COMP:CL:C:F
– Looks like it's snowing out there.

Other devices - including reductions to adverbs like *maybe* – should also be included here (so there'll be a lot of overlap with the Stance domain). Likewise if a language has an inflectional category expressing a probability judgement (e.g. a hypothetical mood). Note that in the case of languages that have something like a so-called 'irrealis' spanning a wide range of meanings (e.g. polysemous between probability desiderative and future meanings) only the using expressing probability evaluations should be coded for under this rubric.

Seeming and simulative constructions can be treated as PROB
e.g., 'it seems that'; 'it's like X'.

THINK (Thought) Used for thinking in the sense of thinking something, having something in mind. The canonical use is representation of thoughts, e.g.
e.g. 'He says in his mind, I should really hold off on the drinking.'
Dream predicates also fall under the THINK functional type.

PRET ('Pretence predicates' like *pretend*, *make believe*)

COM ('Commentative predicates (factive)) e.g. *regret, be sorry, it's good that, it's bad that, be angry that, be upset that*

KNO ('Predicates of knowledge and acquisition, loss or recall of knowledge') like *know, discover, realize, find out, understand*, *forget, remember* or causative versions like *remind.* These will usually be factives. To *mean* will also be coded as KNO in the sense of X means that Y.
–Can't you **see/find** that she is angry? **KNO**:COMP:CL:C:F
–ipalpalagip=na nga adiay nangiyari diay asawa=na 'He is reminding them of what happened to his wife.'
-His sitting like this means that he is sad **KNO**:COMP:CL:C:F

FEAR (fearing) e.g. be afraid, fear, worry, be anxious
Sometimes a FEAR may look a bit like a COMM
e.g. in Dalabon 'he feels bad about…' = 'he worries about…'
One test for this is, is it factive? If yes, it should be COMM, if not, it should be FEAR.

DES (desiderative) e.g. *want, be eager to, desire, -tai* 'want' (Japanese), *kayat* 'want' (Ilokano), and also like, enjoy

IMM ('Immediate perception') e.g. see, hear or causative or passive versions of perception like 'show' or 'be seen' or equivalents like 'demonstrate'

This is primarily for situations where the perception framing and the proposition are simultaneous or overlapping events.
e.g. “I saw them selling pumpkins at the market”
not
“I saw that they had gone to the market”

–He **saw** the man punching her. **IMM**:COMP:CL:Z:NF

– He showed/demonstrated that this cooking takes a long time. **IMM**:COMP:CL:C:F
– I demonstrated how to cook dumplings. **IMM**:COMP:CL:C:NF

What about perception verbs adjacent to independent sentences?
E.g. ‘Look! They’re in the garden.’
This is not annotated unless there is a strong indication that the perception verb frames the proposition in this instance.
When there seems to be a clear relation between the perception verb and the proposition then they should be coded; if you do not think this then do not code.
put two contrastive examples
“Hey, did you see they’re going to the market?”
→ would usually be annotated

e.g. “Oh, look! I think that they’re going to the market.”
→ would NOT usually be annotated

NOTE: There are some situations where a perception verb like ‘see’ might be more appropriately coded as PROB, THINK, KNOW. This might include, for example, inferential situations where one is talking about observing visual evidence of a past event.

4. Overall construction
This parameter concerns the type of construction used for framing propositions.

INDP (independent clause without a framing element)
In this category, the content clause may stand alone or take connective.
Independent constructions do not have a framing element.
– He came in. “**I am here now!**”.
UTT:**INDP**:**Z**:Z:D

– “***Apuroochi sita-de***” *to.* (アプローチしたで、と。) (insubordination)
approach do-decl comp
‘... “He approached her.” ...’
UTT:**INDP**:Z:**C**:D

FUSE (fused)
Fused construction refers to the expression that framing or content phrases are fused into one clause. Two patterns are included in this class:

1. Framing element constitutes part of content clause.

Adverbial modifier in content clause
– *He is **supposedly** fine.* (adverbial modifier)

PROB:**FUSE**:PART:Z:F

Verb inflection in Matukar Panau
- nubnga-lumi-**mbawi**
beer1sg-drink-**IRR:I:DESID**
'I want to drink beer'
DES:**FUSE**:INFL:Z:NF

Another example: Clause with a quotative particle where there is no matrix verb. The particle is coded as the framing element in 5.

2. Content expression is fused into NP, occupying the monoclausal object slot

– She reported **his attack on her**. cf. She reported [that he had attacked her].
UTT:**FUSE**:CL:Z:XE

COMP (complement clause structure)
This is defined as the clause occupying one of core argument slots in the matrix clause. The content clause should retain clausal structure, which would be completely lost in fused NP above.

–She reported that he had attacked her.
UTT:**COMP**:CL:C:IF

SUB (other constructions where the proposition is in a subordinate clause to the framing clause)
This is a broad category of non-complement clause structure, but still a subordinate clause, including chained clauses, converbs, participial constructions, headless relative clause, etc.

-Clause chain in Matukar Panau
nub di-pan-au-dop nga-lumi-ndop nga-mai-ye
beer 3pl-give-1sg-IRR:D 1sg-drink-IRR:D1sg-don't.want-R:I:PFV
'They gave me beer and I didn't want to drink it'
FEAR:**SUB**:CL:Z:NF

ADV (adverbial, where the framing element is in a subordinate clause to the proposition)

G|ui: The framing element is an adverbial clause; the content is a main clause.
cia chu ʔama muu ka, ʔabi (chu) ha ci |ii=sa jia.
me(SUB) PAST him see when he (PAST) ASP song=ACC dance
'When I saw him, he was dancing (yesterday).' IMM:**ADV**:CL:C:F

A note about determining whether you have an ADV or SUB overall structure: Look at the relation between the framing element and the proposition. Or ask yourself, which clause is subordinate/dependent. If the proposition is subordinate to the framing element, use SUB. If the frame is subordinate to the proposition, use ADV.

PARA (parataxis)
Parataxis is applied to cases where the content clause does not occupy the core argument position of the matrix clause, but the framing element is expressed before/after the content clause. Clauses should be sequential/adjacent and uttered by a single speaker.

–He spoke to me. "I'm coming tomorrow."

UTT:**PARA**:CL:Z:D

Note: if either of the two clauses show some dependency on another, it should be coded as SUB/ADV.

For borderline cases between COMP and PARA, make sure your explanatory document outlines your usual criteria for decisions, and include a note in the notes tier in cases where you diverge from these.

PRTH (parenthetical) The main event is presented as a main clause, and the semantically framing quotative, perceptual or cognitive predicate is placed inside it, usually set off by prosodic cues. (In some languages, such as Kayardild, the parenthetical clause gets special morphosyntactic treatment but that is unusual).

'Scandinavians, they say, have an inborn dislike of corruption.' UTT:**PRTH**:CL:Z:F
'You, I see, are a jazz buff.' IMM:**PRTH**:CL:Z:F
'John, I believe, has stopped coming to class.' PROB:**PRTH**:CL:Z:F
'Kazu, I have heard, has a new phone' IMM:**PRTH**:CL:Z:F

COORD (coordination) the framing element and proposition are in a coordinate relationship to each other, normally with a connective. This also includes cases in verb chaining languages where the framing and proposition clauses are part of a chain whose finite form comes later.

G|ui:
cire chu ʔama muu ʔabi ja ha ci |ii=sa jia.
I PAST(yesterday) him see he CONJ ASP(progressive) song=ACC dance
'I saw him dancing (yesterday)' (lit. 'I saw him and he was dancing'.) IMM:**COORD**:CL:C: F

Matukar Panau:
nug ta-hun-i-ndop, **ha-d aipainim di-te-ndop**, kagin manag main-ami di-nage-mba.
RFLX 1pl.incl-hit-3s-IRR:D, **CL-1pl.icl children 3spl-see-IRR:D**, behavior like.that this-only 3pl-do-IRR:FUT
we fight each other, our children see it, they will also do this kind
of behaviour. IMM:**COORD**:CL:Z:NF

5. Framing elements (CL, CLA, NP, PART, INFL, Z)
This parameter concerns the formal category of semantically framing elements.
If you encounter examples with multiple framing elements framing the ***same*** propositional content, you can list multiple framing elements. If so, list the framing elements in ***alphabetical order,*** not the order in which they occur in your data, ***separate the elements with only a comma***, no spaces!), and if you have multiple frames of the same type for a single proposition, you can list doublets. **(PART,PART)**
yes: CL,PART
yes: CL,PART,PART
no: CL, PART
no: PART,CL
no: PART, CL, PART

CL (clause)
CL can be a verb plus arguments or an adjective

–We sighed “I'm never gonna get a perfect coding schema."
UTT:PARA**:CL**:Z:D

Note on ‘bracketed’ structures
This refers to examples where there is more than one clause that ‘goes with’ a framed proposition
e.g., ‘He said ‘No way’ he said.’
This can be coded with two elements, e.g. CL,CL

If the linkage between the two CL is important then can use ^ to indicate bracketing construction CL^CL

e.g., UTT:PARA:CL,CL:Z:D
If the proposition is clearly a complement of one verb and not the other, then it is likely more sensible to just code the COMP
e.g., ‘He said that he wouldn’t drink, he said.’

CLA (clause with anaphoric/cataphoric/endophoric element that refers to the proposition)
-He said this. “I came here”
UTT:PARA**:CLA**:Z:D

other possibilities with CLA:
we think this might be rare/unlikely but possible
-He said this that he is going to go
UTT:COMP:**CLA**:C:IF

We think UTT:**PARA:CLA:C**:D would be impossible, but let us know if you encounter it.
However, we do think it would be possible to have CLA as an option for both PARA or COMP (or perhaps others).

INFL (inflection) this will most likely occur with FUSE constructions

- nubnga-lumi-**mbawi** (Matukar Panau)
beer1sg-drink-IRR:I:DESID
‘I want to drink beer’
DES:FUSE:**INFL**:Z:NF

‘Inflection’ is intended broadly so this category also includes clitics and other bound elements that it is not appropriate to term “particles” (PART).

FRAG (fragment)
This includes cases where elements of a clause or other larger structure appear to have been elided and a ‘fragment’ is doing the framing work. (This can include framing elements that are formally NPs; but not all NP frames will be FRAG, see the NP category below.)

Examples include situations such as constructions where the ‘author’ of the reported speech is referred to, but without a verb:

e.g. ‘<u>he-ERG</u> “blah blah”’

‘He “I’m going now”’

“No idea how this should continue”

‘He to me “They are fighting”’

‘Like “They are fighting”’

Put in your notes what is happening here and what you think it being elided.

NP (nominal)
– I heard **the story** that the man fell down into the river.
UTT:COMP:**NP**:C:F

Notes: NP can take a complement clause (*noun complementation, appositive clause*), which is structurally different from relative clause in that the complement clause does not have null argument. (cf. relative clause: ***the story*** *that Ø surprised everyone*)

PART (particle, adverbial)

– *He is **supposedly** fine.* (adverbial modifier)
PROB:FUSE:**PART**:Z:F

Z (zero)

– He came in. “**I am here now!**”.
UTT:INDP:**Z**:Z:D

6. Connective
This parameter specifies whether content propositions and the relevant framing element are connected with an explicit formal device (C) or not (Z). Connectives may be morphologically complex, like *as to whether* in English, *mitaina/to iu (hanasi)* (みたいな/という(話)) ‘a story like/saying’ in Japanese or maybe something like case marking. Prototypically is something like a complementizer (*that*) or a conjunction (*and*).
Infinitive forms do not have a connective.
If you are coding a PARA construction, think about whether you have CLA:Z or CL:C if there is some element between the frame and proposition.

7. Content Proposition
The content may be represented as reported speech/thought or not. This is independent of whether the function is THINK or UTT or another kind, as many languages may use speech content as the complement in DESID, IMM or other functions. Look at the proposition and decide whether it seems like it is representing an actual or purported quote or enactment. If it does, use the first series: D/I/N/XE. If it does not, use the second series: F/NF.

For speech and thought : This series follows Reported Speech coding schema.

D: direct speech and thought and action may be indicated by a shift in tone of voice or prosody. Ideophones and depictions are also a kind of 'D', but make a note if this is occurring in your data.

I: indirect speech and thought indicated by: (a) shift in person; (b) shift in TAM; (c) deixis from the purported time of speech/thought/action and (d) other language specific indicators.

FOR ALL speech coded as I, the status as finite or non-finite must also be indicated
i.e. the choice is between IF or INF

e.g. He asked her to go
UTT:COMP:CL:Z:INF

Note: if the speech is mixed/biperspectival, follow the conventions of the reported speech and thought schema, and add an underline for the device which indicates the speech's mixed status, i.e. I_log or I_tense. This should be also showing up in the reported speech and thought tiers. Please also let Nick & Danielle know if you observe this in your data.

N: not determinable as to whether it is direct or indirect

XE: compressed expressions representing speech or thought, more compact than direct or indirect speech, but still maintaining content, e.g. She reported his attack on her to the police, dia melaporkan pemukulan oleh polisi 'he reported his beating by the police'. Compared to Direct or Indirect Speech/Thought/Action there is substantial reworking of the construed words of the original report (above and beyond deictic shifts); compared to CE there is actual content to the material. Typically this will be an argument of a speech or thought predicate (where clauses are regarded as a type of argument).

For non-speech/thought categories
F: finite this is the default category if the language has no TAM or person marking to help determine finiteness. This means the embedded proposition is the same (no changes) regardless of whether it's a stand-alone utterance or a framed/embedded one. For example, Auslan:

BUT PT:PRO1 SEE SAME PT:DET BOY GO HOLIDAY OR GO SCHOOL OR WHATEVER
But I see (it's) like that boy go holiday or go school or whatever…
UTT t
The embedded clauses 'go holiday' and 'go school' have no TAM or person marking that is different to how they could be signed stand alone.

NF: nonfinite
If some TAM markings are helpful for distinguishing finite from nonfinite, use this as the criterion to divide F vs. NF. If it is not available or does not work well in the language, use other criteria (e.g. subject/object agreement). Make a note of these criteria. Nominalizations, gerunds, and some kinds of subordinate clauses will probably be nonfinite

If there is a logophoric element, this can be indicates as I/D/F/NF_LOG (etc.)

Are you having trouble deciding when to code for I/D and when F?
→ note that I/D is for propositions that represent speech or things going on in someone's head
I/Ds could also (in principle) include desire, remembrance, etc.

8. Notes
The notes tier can include more than one note (e.g., separated by comma).

GCAC - **grammaticalisation of complements and complementizers/connectives -** appropriate for borderline cases of PARA:CL:C/ COMP:CL:C that hinge on some connective being considered a complementiser or not. See 9.1 for more details.

ES - **Embedded in Speech** In each Comp_Notes@speaker tier, make a note if the frame-proposition unit is occurring inside reported speech and thought with **ES**. This is similar to what we are doing for Stance. However, the reported speech will also be coded within this schema, so **put the annotations in sequence**.

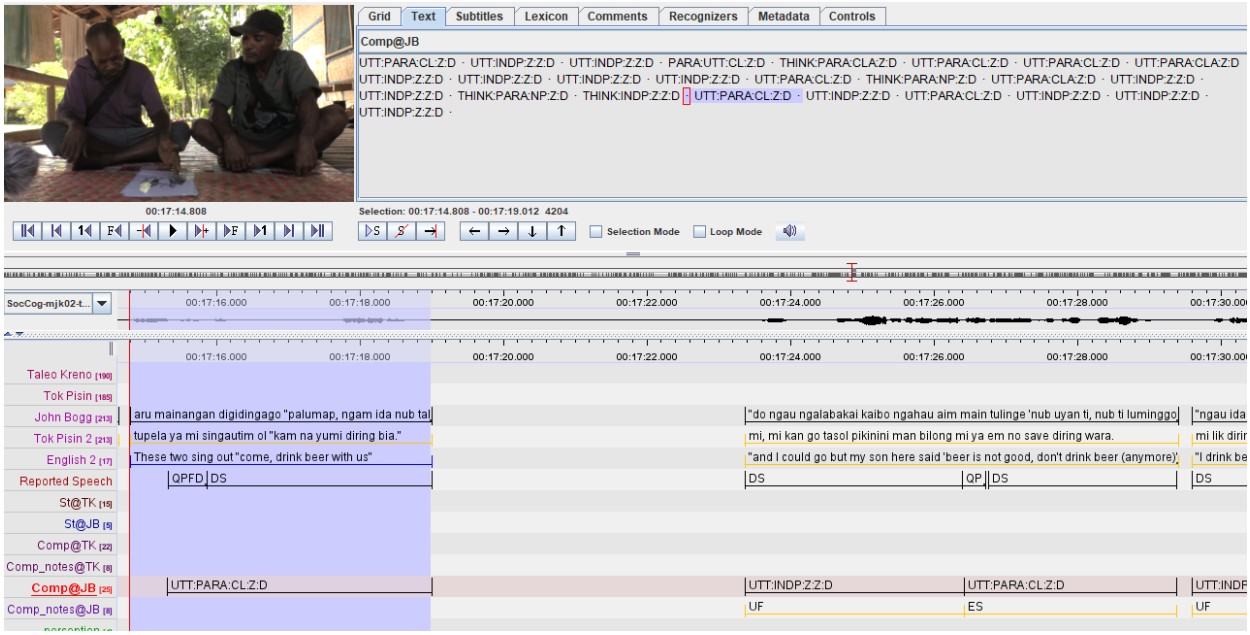
E - Embedded In each Comp_Notes@speaker tier, make a note if the frame-proposition unit is embedded inside another frame-proposition unit that is not speech (D/I/N/XE). Put **E** in the notes to indicate this.

Where there are multiple layers of embedding:
For those beyond the first layer, include numbers to show the level of embedding, e.g., ES2, E3, ES4, etc.
The count for the multiple layers can encompass both E and ES. So you may have something like ES, ES2, E3.

UF - Umbrella Framed In cases where there is a string of sentences/utterances that all have a common frame either preceding or following the string. Put this note under the utterances that don't have the adjacent frame.

The Figure below shows notes with both UF and ES. The first utterance has a frame (PARA). The next utterance has no frame (INDP). However, it is in a way under the umbrella of the first frame. This annotation receives the note UF. This second utterance also contains an utterance inside the direct speech. That annotation receives the note ES. The third annotation, partially visible in the figure, is also an utterance, under the overall umbrella frame from the first annotation, it receives then a UF in its note. There are 3 annotations in the utterance (ts) level tier, 4 annotations at the propositional framing (comp) tier.
Note: UF is not applied to turns from another speaker.
E.g.
Speaker A: "And then he says, 'I'm not going'" Speaker B: "No way, I'm staying."
In this case, Speaker B's turn is *not* annotated as UF.

Note, UF is typically for reported speech/thought, and is NOT used for situations such as:
'This card shows that he is standing outside. He is in the sunshine, and he's happy.'
This is treated like a frame with a single proposition. The sentence 'He is in the sunshine...' is NOT coded as an independent UF proposition.



NEG - Negated

PD – Primary Discourse Make a note if the proposition-framing unit or the framing is at the level of primary discourse. Many of the PROB categories will have the frame at the PD level and will need a PD in the notes.
Note: Narrative discourse is assumed to be the normal level for this category.

If the proposition is a finite clause that has some element elided then this can be treated as 'F' but the notes can indicate the presence of ellipsis

e.g. He's coming back from school. Or **maybe** from work.

'or maybe from work' is not a complete clause but can be coded as F with a note about the ellipsis in the notes line.

Some coding examples

Ku Waru:
Here is an example with RS annotation meaning 'He said this "You are at fault, you should not do it"'.

en	lawa	ting	naa	tiai	i	nirim
you:PL	wrong; fault;	do	not	do	here; this	said
pro	n	v	av	v	DEM	v
DS					PRU	QPFM_nyi.

We coded this as **UTT:COMP:CL:C:D**

9. Common coding problems and quandaries

9.1.1 Grammaticalisation from paratactic to complement constructions - CL OR CLA?

A common problem arises in coding quoted speech constructions on the cusp of grammaticalising from direct quoted speech, paratactic, introduced by a demonstrative like this or that, to complement constructions where the demonstrative word has grammaticalised to a complementiser (like English that). The latter coded in our schema as one type of combination of COMP plus CL plus C (connective), where the framed clause is now clearly a complement clause rather than in a paratactic construction. The former is coded as PARA:CLA:Z, with the demonstrative word contributing to the 'A' in the coding decision (He said this...).

In other words, a very common historical development is from a structure like

CLA | PARA

Pred<speech> DEM (this, that, like.this) "Direct speech"

to

Pred<speech? Complementizer | COMP
Complement clause

It can be hard to decide which way to code these, in given cases. Factors to take into account are:

(a) prosody, i.e. are the two clauses prosodically independent (the paratactic case) or more tightly linked (the complement case)

(b) semantic and phonological properties of the demonstrative/connective – e.g. in English *that* is obligatorily /ðæt/ when used as a demonstrative, but more commonly /ðət/ when used as a complementiser, and very specific semantic connections are needed to use that in a paratactic construction, where this is much more likely

(c) there may be further evidence from deictic values, where deictic shifts characteristic of indirect speech are evidence in favour of the complement clause analysis. Cf the difference in interpreting who is the referent of 'I', Mengistu or Lila, in a situation of Lila saying:

'Mengistu said that I am a hero.' (Lila is the hero) COMP:CL:C:IF
'Mengistu said this: "I am a hero"' (Mengistu is the hero) PARA:CLA:Z:D
As this example makes clear, there are various dependencies in coding here – e.g. between indirect speech (I) for the content proposition, COMP for the overall construction, and coding of the this/that word as a C (connective) or the A part of the CLA.

9.1.2 Grammaticalisation from paratactic to complement constructions - GCAC

Because we are interested in this grammaticalisation path, and the vital examples are often hard to rule on, we encourage annotators to identify such examples overtly by
(a) in the relevant notes tier, insert the abbreviation **GCAC** (mnemonically: grammaticalisation of complements and complementizers/connectives) to make this findable by R
(b) in the same tier and annotation give your prose rationale for your decision.

E.g. for the Dalabon example

kahdjakurnhwonawonang kahyininj, nunh, 'Ngayeburlhmiyan [etc.]
‘He was thinking like ‘When I get out (of gaol), ...’

there are two plausible options. (Note that kahdjakurnhwonawonang means ‘he was thinking and thinking’, kahyininj is ‘he said/did/thought’, and nunh is ‘(like) this’).

(a) treat this as a paratactic construction comparable to English

‘He was thinking, he said (mentally) like: “When I get out...’
which would be annotated in the following way. Treat kahdjakurnhwonawonang as not being part of any framing construction, but merely setting up the fact he was thinking. Then kahyininj nunh ‘he said like this’ is in a paratactic relationship to the following direct speech. The whole sequence kahyininj nunh ngayeburlhmiyan would then be annotated as UTT:PARA:CL:Z:D

(b) treat this as a complement construction in which kahyininj nunh ‘said that’ has grammaticalised to a complementiser. Now kahdjakurnhwonawonang IS treated as a complement-taking predicate, with the kahyininj nunh being a complementiser. In that case the annotation would be either

(i) THINK:PARA:CL:C:D with kahyininj nunh being the C (but the overall construction still being paratactic rather than a proper complement construction, or

(ii) THINK:COMP:CL:C:D in which we treat it as a full-fledged complement construction

The problem with (b.i) and (b.ii), as analyses, is that there is nothing reduced or conventionalised about kahyininj here. It is a fully inflected finite verb capable of standing by itself with the meaning 'he said' (and nunh is a demonstrative introducing upcoming material).

So the only reason for including this here as a complement-taking predicate would be the close prosodic integration between the two, suggestive of a potentially emerging complementizer deriving from a speech act verb, which is why we opt for coding choice (a).

However, it is cases like this are likely to be grammaticalisation sources for how complement clauses emerge. To save them from being invisible in our corpus, we add the **GCAC** tag in the COMP_notes tier .

9.2 Grammaticalisation of complement constructions into stance markers

Another widespread phenomenon is the grammaticalisation of erstwhile framers like

I think (that)
I reckon (that)
You know (that)
Do you see (that)

to more tag-like stance markers.

These should be coded as PROB functional type and borderline cases should be noted in the notes tier with a prose comment