Defining UMR

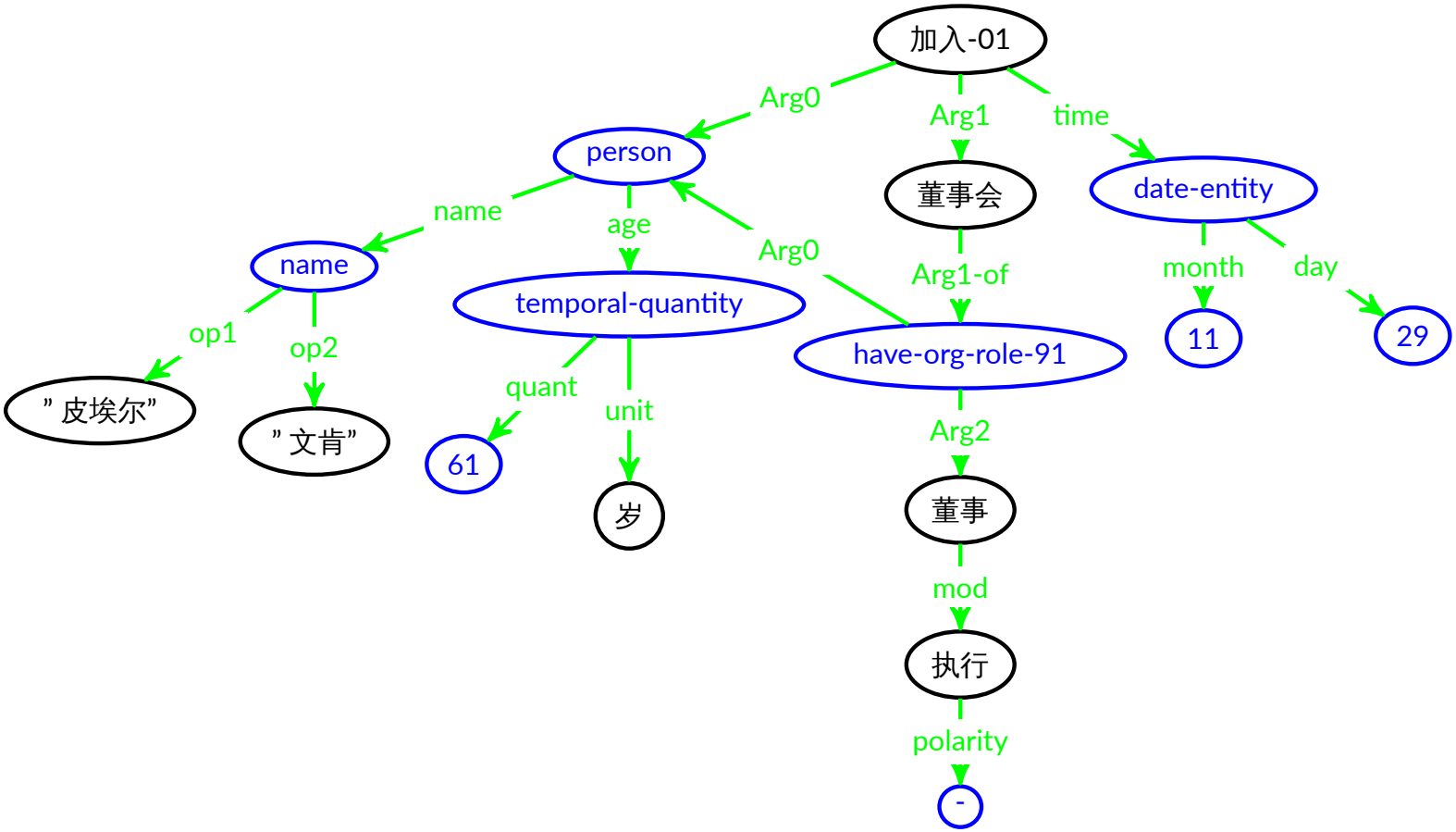
(Van Gysel et al., 2021)

- At the sentence level, UMR adds:
  - An *Aspect* attribute to eventive concepts
  - *Person* and *Number* attributes for pronouns and other nominal expressions
  - A principled set of discourse relations
  - Quantification scope between quantified expressions
- At the document level UMR adds:
  - Temporal dependencies in lieu of tense
  - Modal dependencies in lieu of modality
  - Coreference relations beyond sentence boundaries

UMR is a cross-lingual meaning representation

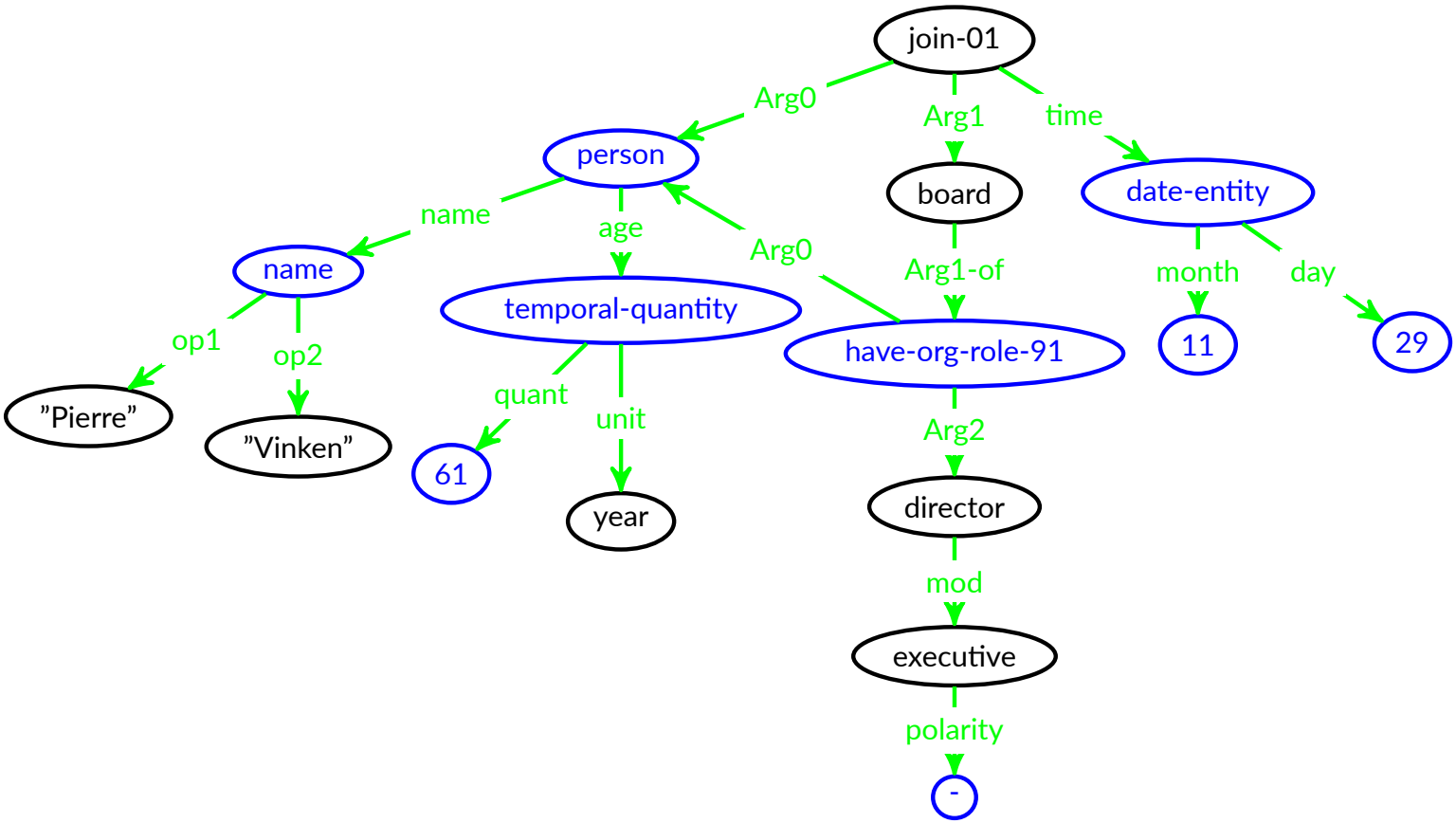
- Abstract concepts (e.g., *person*, *thing*, *have-org-role-91*) are uniform across languages
  - Concepts that do not always have explicit lexical support but can be inferred from context
- UMR defines a set of general participant roles (e.g., *agent*, *theme*, *causer*) and non-participant relations that are uniform across languages
- But UMR is still not an Interlingua:
  - Lexical concepts include sense-disambiguated lemmas or simple lemmas and are language-specific (e.g., Mandarin 加入*.01* vs. English *join-01* vs. Sanapaná *empahlkay'a*)
  - Languages can define their own lexicalized participant roles (e.g., *:ARG0* of 加入.01)
- In general, grammatical meaning is language-independent while lexical meaning is language-specific

Language-independent vs language-specific aspects



“61 岁的 Pierre Vinken 将于 11 月 29 日加入董事会，担任非执行董事。”

Language-independent vs language-specific aspects



"Pierre Vinken , 61 years old , will join the board as a nonexecutive director Nov. 29 ."

Abstract concepts in UMR

- Abstract concepts inherited from AMR:
  - Standardization of quantities, dates etc.: *have-name-91*, *have-frequency-91*, *have-quant-91*, *temporal-quantity*, *date-entity*...
- New concepts for abstract events: "non-verbal" predication.
- New concepts for abstract entities: entity types are annotated for named entities and implicit arguments.
- Scope: *scope* concept to disambiguate scope ambiguity to facilitate translation of UMR to logical expressions (see sentence-level structure).
- Discourse relations: concepts to capture sentence-internal discourse relations (see sentence-level structure).

Where do we find abstract eventive concepts?

Semantic type and information packaging (Croft 2001):

	Reference	Modification	Predication
Entities	UNMARKED NOUNS	relative clauses, PPs on nouns	**predicate nominals, complements**
States	deadjectival nouns	UNMARKED ADJECTIVES	**predicate adjectives, complements**
Processes	**event nominals, complements, infinitives, gerunds**	**participles, relative clauses**	**UNMARKED VERBS**

Where do we find abstract eventive concepts?

- Sentence-level information packaging is not always predicational:
  - *I have a book* - “thetic”, “presentational”, all new information
  - *The book belongs to me* - “predicative”, possessee is known information
- AMR does not distinguish these meanings, UMR does only in typically “non-verbal” contexts:
  - Possession
  - Location
  - Object/Property predication

Where do we find abstract eventive concepts?

- Languages use different strategies to express these meanings:
  - Overt predication: English *I have a book*

Where do we find abstract eventive concepts?

- Languages use different strategies to express these meanings:
  - Juxtaposition: Tiwi

```
ngawa    mantani teraka
our      friend  wallaby
`Our friend has a wallaby,
lit. [as for] our friend, wallaby.'
```

Where do we find abstract eventive concepts?

- Languages use different strategies to express these meanings:
  - Predicativized possessum: Yukaghir

```
pulun-die   jowje-n'-i
old.man-DIM net-PROP-3SG.INTR
`The old man has a net,
lit. The old man net-has.'
```

  - UMR trains annotators to recognize the semantics of these constructions and select the appropriate abstract predicate and its participant roles
  - UMR does not require alignment between concepts and words

Sample abstract events

Clause type	Predicate	ARG0	ARG1	ARG2
thetic poss.	have-91	possessor	possessum	
pred. poss.	belong-91	possessum	possessor	
thetic loc.	exist-91	location	theme	
pred. loc.	have-location-91	theme	location	
property pred.	have-mod-91		theme	property
object pred.	have-role-91	theme	ref. point	object category
equational	identity-91	theme	equated referent	

Example abstract events

```
an-yetn-eye'          ko'o    vakka-hak       ah-angkok
2/3F-exist-V1.NFUT    1SG:PRO cow-old/broken 1SG-POS
'I have a book.' lit. 'My book exists.'
(h / have-91
   :ARG0 (p / person
            :ref-person 1st
            :ref-number Singular)
    :ARG1 (v/ vakkahak 'book')
    :aspect State)
```

Named entities in UMR are hierarchically organized

Class	Type	Subtype
animal, plant, language, nationality, person, cultural-activity, award, food-dish, computer-program, variable		
social-group	family, ethnic-group, regional-group, religious-group, clan	
	organization	international-organization, business, company, government-organization, political-organization, criminal-organization, armed-organization, academic-organization, association, sports-organization, religious-organization

Named entities in UMR are hierarchically organized

Class	Type	Subtype
geographic-entity	ocean, sea, lake, river, gulf, bay, strait, canal, peninsula, mountain, volcano, valley, canyon, island, desert, forest	
celestial-body	moon, planet, star, constellation	
region	local-region, country-region, world-region	
geo-political-entity	city, city-district, county, state, province, territory, country	
facility	airport, station, port, tunnel, bridge, road, railway-line, canal, building, theater, museum, palace, hotel, worship-place, market, sports-facility, park, zoo, amusement-park	
vehicle	ship, aircraft, aircraft-type, spaceship, car-make	

Named entities in UMR are hierarchically organized

Class	Type	Subtype
cultural-artifact	work-of-art, picture, music, literature, dance, show, broadcast-program	
	publication	book, newspaper, magazine, journal
event	incident, natural-disaster, earthquake, war, conference, game, festival, ceremony	
notational-system	writing-script, music-key, musical-note	
biomedical-entity	molecular-physical-entity, small-molecule, protein, protein-family, protein-segment, amino-acid, macro-molecular-complex, enzyme, nucleic-acid, pathway, gene, dna-sequence, cell, cell-line, species, taxon, disease, medical-condition	

Road Map

- Lexical resources and grammatical analysis is not available for many languages
- UMR aims to be available for semantic annotation of languages from the very beginning of analysis. It is therefore structured as a “Road Map”
  - Early stages of Road Map must not rely on availability of resources or analysis
  - Annotations at earlier stages must still be compatible with more fine-grained annotations at later stages

Road Map

- Participant Roles:
  - Stage 0: General participant roles
  - Stage 1: Language-specific frame files
  - UMR-Writer allows for the creation of lexicon with argument structure information during annotation
- Morphosemantic Tests:
  - Stage 0: Identify one concept per word
  - Stage 1: Apply more fine-grained tests to identify concepts
- Annotation Categories with Lattices:
  - Stage 0: Use grammatically encoded categories (more general if necessary)
  - Stage 1: Use (overtly expressed) fine-grained categories
- Modal Dependencies:
  - Stage 0: Use simplified modal annotation
  - Stage 1: Fill in lexically based modal strength values

Language-independent vs -specific participant roles

- Core participant roles are defined in a set of frame files (valency lexicon, see Palmer et al. 2005). The semantic roles for each sense of a predicate are defined:
  - E.g. boil-01: *apply heat to water*
    ARG0-PAG: *applier of heat*
    ARG1-PPT: *water*
- Most languages do not have frame files
  - But see e.g. Hindi (Bhat et al. 2014), Chinese (Xue 2006)
- UMR defines language-independent participant roles
  - Based on ValPaL data on co-expression patterns of different micro-roles (Hartmann et al., 2013)

Language-independent roles: An incomplete list

UMR Annotation	Definition
Actor	Animate entity that initiates the action
Undergoer	Entity (animate or inanimate) affected by the action
Theme	Entity (animate or inanimate) moving from one entity to another, spatially or metaphorically
Recipient	Animate entity that gains possession (or at least temporary control) of another entity
Force	Inanimate entity that initiates the action
Causer	Animate entity that acts on another animate entity to initiate the action
Experiencer	Animate entity that cognitively or sensorily experiences a stimulus
Stimulus	Entity (animate or inanimate) that is experienced by an experiencer

Language-independent vs language-specific participant roles

- Language-independent participant roles are assigned independent of grammatical relations in a language
  - [She]$_{actor}$ ate.
  - [She]$_{undergoer}$ fell.
- As lexical resources are developed for a language, annotations with language-independent participant roles can be linked to predicate-specific numbered argument roles:
  - E.g. Sanapaná Stage 0 – entoma: *eat*
    Actor: *eater*
    Undergoer: *food*
  - Sanapaná Stage 1 – entoma-01: eat
    ARG0: Actor
    ARG1: Undergoer

Argument structure alternations

- Information-packaging alternations (e.g., passives) are not distinguished by the participant role annotation
- Semantic alternations (e.g., causatives, reflexives) are reflected in the participant role annotation

Berber (Guerssel 1986:52)

```
T-ttw-asy            tfirast.
3F.SG-DETR-pick.up   pear
'The pear was picked up.'

(t/ tttwasy-00 'pick up'
     :undergoer (t2/ tafirast 'pear'
          :ref-number Singular)
     :aspect Performance
     :modstr FullAff)
```

Kukama

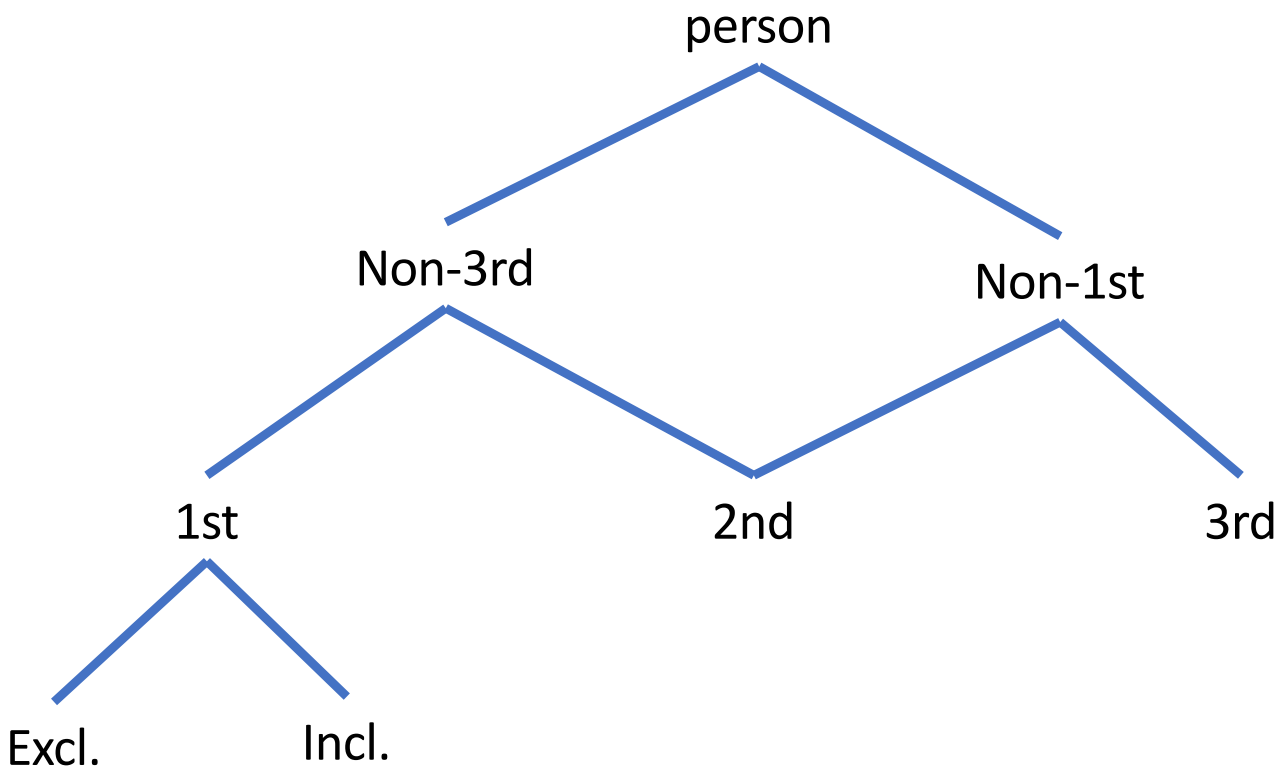
```
nai          kurata-ta    churan=ui    uni=pu
grandmother  drink-CAUS   kid=PST      water=INS
'Grandmother made the kid drink the water.'

(k/ kuratata 'make drink'
     :causer (p/ person
          :ARG0-of (h/ have-rel-role-91
               :ARG2 (n/ nai 'grandmother'))
          :ref-number Singular)
     :actor (c/ churan 'kid')
          :ref-number Singular
     :undergoer (u/ uni 'water')
     :aspect Performance
     :modstr FullAff)
```

How UMR accommodates cross-linguistic variability

- Not all languages grammaticalize/overtly express the same meaning contrasts:
  - English: *I* (1SG) vs. *you* (2SG) vs. *she/he* (3SG)
  - Sanapaná: *as-* (1SG) vs. *an-/ap-* (2/3SG)
- However, there are typological patterns in how semantic domains get subdivided:
  - A 1/3SG person category would be much more surprising than a 2/3SG one
- UMR uses lattices for abstract concepts, attribute values, and relations to accommodate variability across languages.
  - Languages with overt grammatical distinctions can choose to use more fine-grained categories

Category Lattices

- Semantic categories are organized in “lattices” to achieve cross-lingual compatibility while accommodating variability
  - Lattices for Aspect, Modal Strength, Person, Number, Discourse Relations, Modification Relations



Wordhood vs concepthood across languages

- The mapping between words and concepts in languages is not one-to-one: UMR designs specifications for complicated mappings between words and concepts.
  - Multiple words can map to one concept (e.g., multi-word expressions)
  - One word can map to multiple concepts (morphological complexity)

Multiple words map to one concepts

Consistent standards for annotating MWEs cross-linguistically

```
(i / intrigue-01
      :Aspect State
      :ARG0 (a / aspect
            :ARG1-of (m / moral-02)
            :poss (m2 / movement-07))
      :ARG1 (p/ person
            :ref-person 3rd
            :ref-number Singular)
      :mod (a2 / as-well)
      :modstr FullAff)
```

The moral aspects of the movement intrigued him as well

```
(x0/敲竹杠-01
   :Aspect Activity
   :arg0 (x1/他)
   :arg1 (x2/人
            :mod (x3/老)
            :mod (x4/可怜)
            :mod (x5/那)
            :unit (x6/个)))
```

对那个可怜的老人，他还 敲竹杠。

Concepts can map to words that are discontinuous

```
(x0/帮忙-01
   :aspect Performance
   :arg0 (x1/地理学)
   :affectee (x2/我)
   :degree (x3/大))
```

地理学帮了我很大的忙。

```
(w / want-01
      :Aspect State
      :ARG0 (p / person)
            :ref-person 3rd
            :ref-number Singular
      :ARG1 (g / give-up-07
            :ARG0 h
            :ARG1 (t / that)
            :aspect Performance
            :modpred w)
      :ARG1-of (c / cause-01
            :ARG0 (a / umr-unknown))
      :aspect State)
```

"Why would he want to give that up?"

One word maps to multiple UMR concepts

- One word containing predicate and arguments

Sanapaná:

```
yavhan  anmen   m-e-l-yen-ek
honey   alcohol NEG-2/3M-DSTR-drink-POT
"They did not drink alcohol from honey."
```

```
(e / elyama
     :actor (p / person
        :ref-person 3rd
        :ref-number Plural)
     :undergoer (a / anmen
        :material (y/ yavhan))
     :modstr FullNeg
     :aspect Habitual)
```

- Argument Indexation: Identify both predicate concept and argument concept, don't morphologically decompose word

One word maps to multiple UMR concepts

- One word containing predicate and arguments

Arapaho:

```
he'ih'iixooxookbixoh'oekoohuutoono'
he'ih'ii-xoo-xook-bixoh'oekoohuutoo-no'
NARR.PST.IPFV-REDUP-through-make.hand.appear.quickly-PL
``They were sticking their hands right through them
[the ghosts] to the other side.''

(b/ bixoh'oekoohuutoo `stick hands through'
    :actor (p/ person :ref-person 3rd :ref-number Plural)
    :theme (h/ hands)
    :undergoer (g/ [ghosts])
    :aspect Endeavor
    :modstr FullAff)
```

- Noun Incorporation (less grammaticalized): identify predicate and argument concept

One word maps to multiple UMR concepts

- One word containing predicate and arguments

Arapaho:

```
hoono' nuhu' tihciinii'eihiinit, he'ih'etoocein nuhu' hitiine' nuhu' hoote
hoono'  nuhu'   tih-cii-nii'eihiini-t,
not.yet this    when.PST-NEG-be.eagle-3S
he'ih-'etoocein                    nuhu'   hi-tiin-e'      nuhu'   hoote
NARR.PST-pull.rope-like.thing.out  this    3S-mouth-LOC    this    sinew
``At the [time] when he wasn't yet an eagle,
he took [it] out of his mouth, the sinew.''
```

```
(e/ 'etoocein `pull rope-like thing out'
    :actor (p/ person :ref-person 3rd :ref-number Singular)
    :theme (h/ hoote `sinew')
    :source (h2/ hitiine' `mouth'
        :Part-of p)
    :aspect Performance
    :modstr FullAff)
```

- Noun Incorporation (more grammaticalized): identify predicate, ID argument only if independently expressed

One word maps to multiple UMR concepts

- Derivational valency-changing morphology vs. auxiliaries

Kukama:

```
nai          kurata-ta   churan=ui   uni-pu
grandmother  drink-CAUS  kid-PST     water-INST
``Grandmother made the kid drink the water.''
```

```
(k/ kuratata `make drink'
    :causer (n/ nai `grandmother')
    :actor (c/ churan `kid')
    :undergoer (u/ uni `water')
    :aspect Performance
    :modstr FullAff)
```

```
(d/ drink
    :cause (m/ make
        :actor (g/ grandmother)
        :aspect Performance
        :modstr FullAff)
    :actor (k/ kid)
    :undergoer (w/ water)
    :aspect Performance
    :modstr FullAff)
```

- Independent negation test to judge if there are two events

One word maps to multiple UMR concepts

- Derivational TAM morphology vs. auxiliaries

Arapaho:

```
ceesisnoo'oebiicitiit
ceesis-noo'oe-biicitii-t
IC.begin-around-bead.st-3S
``She is starting to bead around it.''
```

```
(b/ biicitii `bead st.'
    :actor (p/ person
        :ref-person 3rd
        :ref-number Singular)
    :undergoer (t/ thing)
    :aspect Activity
    :modstr FullAff)
```

```
(b/ bead st.
    :actor (p/ person
      :ref-person 3rd
      :ref-number Singular)
    :undergoer (t/ thing)
    :aspect Activity
    :modstr FullAff)
```

- Aspect: not identified as separate event regardless of morphosyntactic expression

One word maps to multiple UMR concepts

- Derivational TAM morphology vs. auxiliaries

Arapaho:

```
xonouu niibeetwon3eiinein
xonouu      nii-beet-won-3eiin-ein
immediately IPFV-want-ALL-put.inside.a.place-3S/2S
``Right away he wants to go and put you in jail.''

(b/ beetwon3eiin `want to go and put s.t. inside a place'
    :actor (p/ person
        :ref-person 3rd
        :ref-number Singular)
    :theme (p/ person
        :ref-person 2nd
        :ref-number Singular)
    :aspect Habitual
    :modstr NeutAff)
```

```
(w/ want
    :experiencer (p/ person
      :ref-person 3rd
      :ref-person Singular)
    :stimulus (p2/ put
      :actor p
      :theme (p3/ person
            :ref-person 2nd
            :ref-number Singular)
        :goal (j/ jail)
        :aspect Habitual)
        :modpred w
    :aspect State
    :modstr FullAff)
```

- Semi-modals: Independent modalization test to judge if there are two events

One word maps to multiple UMR concepts

Sanapaná:

```
apk-el-vet-angv-ay-akm-e'        hlema   nenhlet
2/3M-DSTR-see-arrive-TI-TRM-V1   one     person
``They arrived and saw a person.''

(v/ engvetangvayam `arrive and see'  (a/ and
    :experiencer (p/ person              :op1 (a2/ arrive
        :ref-person 3rd                      :actor (p/ person
        :ref-number Plural)                    :ref-person 3rd
    :stimulus (p2/ person)                       :ref-number Plural)
    :aspect State                            :aspect Performance
    :modstr FullAff)                             :modstr FullAff)
                                         :op2 (s/ see
                                             :experiencer p
                                             :stimulus (p2/ person)
                                             :aspect State
                                             :modstr FullAff))
```

- Associated Motion: Independent argument structure test to judge if there are two events

UMR sentence-level additions

- An *Aspect* attribute to event concepts
  - *Aspect* refers to the internal constituency of events - their temporal and qualitative boundedness
- *Person* and *number* attributes for pronouns and other nominal expressions
- A set of concepts and relations for discourse relations between clauses
- Quantification scope between quantified expressions to facilitate translation of UMR to logical expressions