ANAPHORA AND AMBIGUITY

ANNOTATION OF ANAPHORIC INFORMATION

CONTENT

- My Project
- Relevant Notions
- Corpora
- Procedure
- Results
- Further Research

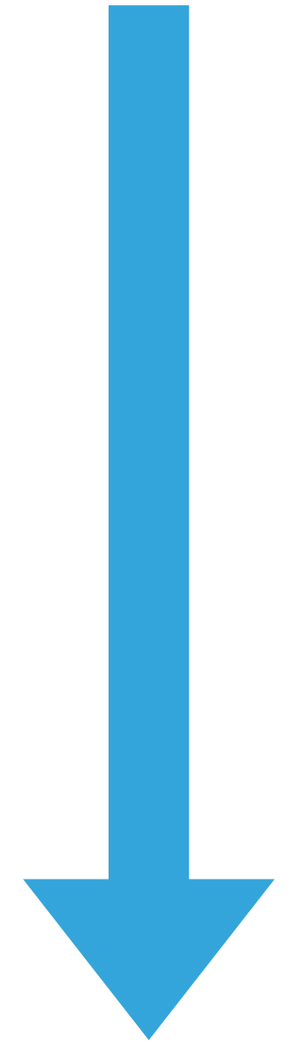
THE PROJECT

GENERAL OBJECTIVE

- analysis of the disagreement and ambiguities in anaphoric judgements between annotators recorded in two corpora
  - Phrase Detectives (Yu et al, 2023; Poesio et al, 2019)
  - ARRAU (Uryupina et al, 2020)

GENERAL OBJECTIVE

- examination of the corpora and their annotation schemes ✔
- evaluation of annotations marked as ambiguous ✔
- Phrase Detectives (PD): scheme for annotating disagreements
  - evaluation of PD 3 Gold according to disagreement annotation scheme
- compute inter-annotator agreement
- analysis of cases of disagreement
  - new cases of ambiguity?

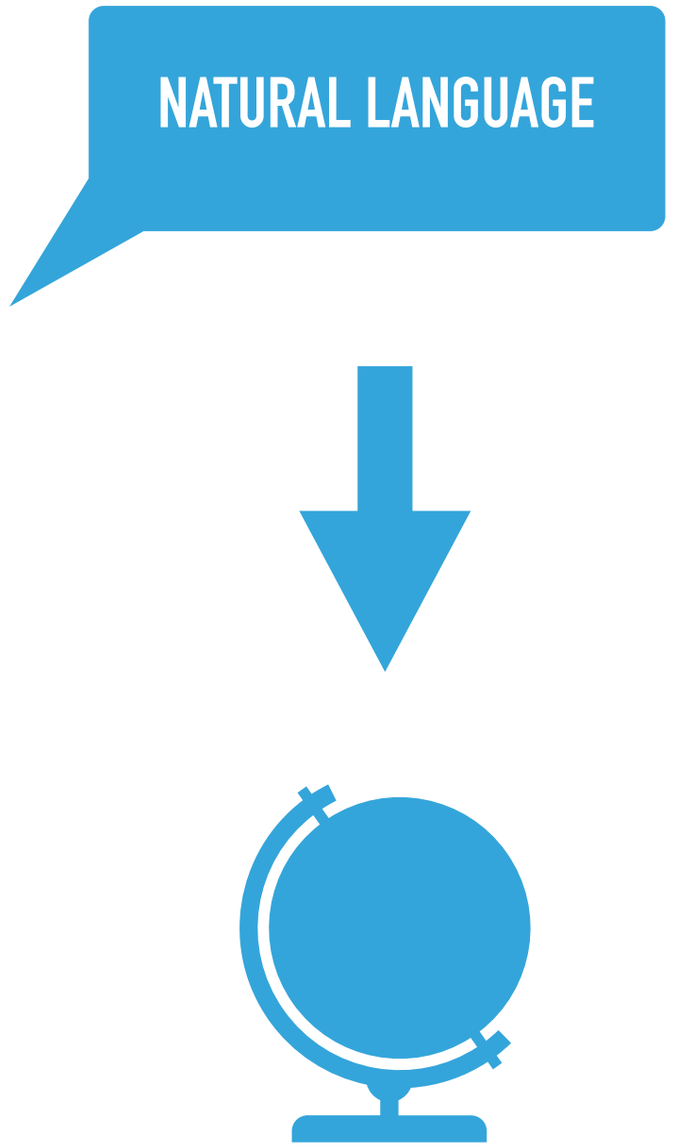


Master Thesis

SOME RELEVANT NOTIONS

SOME TERMINOLOGY

- (co-)reference



SOME TERMINOLOGY

- (co-)reference
- anaphora
- entity context dependence
- interpretation depends on relation to (discourse) entities
  - e.g. pronouns
    - [A bird] sits in the tree. [It] is green.

SOME TERMINOLOGY

- (co-)reference
- anaphora
- bridging

We can’t take [the car]. [The wheel] is broken.

SOME TERMINOLOGY

- (co-)reference
- anaphora
- bridging
- (discourse) deixis

- use of a linguistic expression to point to some entity or property
- entity is contextually available
  - in context and also extralinguistically
- visual context: "Can you open the door?"
- spatial deixis: "John lives [there]."
- discourse deixis:
  after listening to a story
  "[That] is a funny story."

SOME TERMINOLOGY

- (co-)reference
- anaphora
- bridging
- (discourse) deixis
- markable

S : okay

M : uh / and then

and then uh

send [E2]$_{old}$

send [/ E2]$_{old}$'s already going [there]$_{old}$

and as soon as [it]$_{old}$ gets [there]$_{old}$

couple [the cars]$_{old}$ and take [it]$_{old\ phrase}$ to uh

S : well [we]$_{old}$ n / [we]$_{old}$ need [a boxcar]$_{new}$

to take [the oranges]$_{old}$

SOME TERMINOLOGY

- (co-)reference
- anaphora
- bridging
- (discourse) deixis
- markable
- generic
- expressing some kind of world knowledge
  - "[Ravens] are black."
- expressing a generalization
- reference to a kind

THE CORPORA

ARRAU

- multi-genre corpus for English
  - news texts (RTS, 72013 markables)
  - medical and art history genre (GNOME, 6562 markables)
  - task-oriented dialogues (TRAINS, 16999 markables)
  - spoken narratives (PEAR, 4008 markables)
- 350000 tokens

ARRAU – GOAL OF ANNOTATION

- collection of data about how
  - context is constructed
  - context is used in natural language
- two phenomena in particular
  - anaphoric reference
  - deictic reference

ARRAU – WHAT IS ANNOTATED

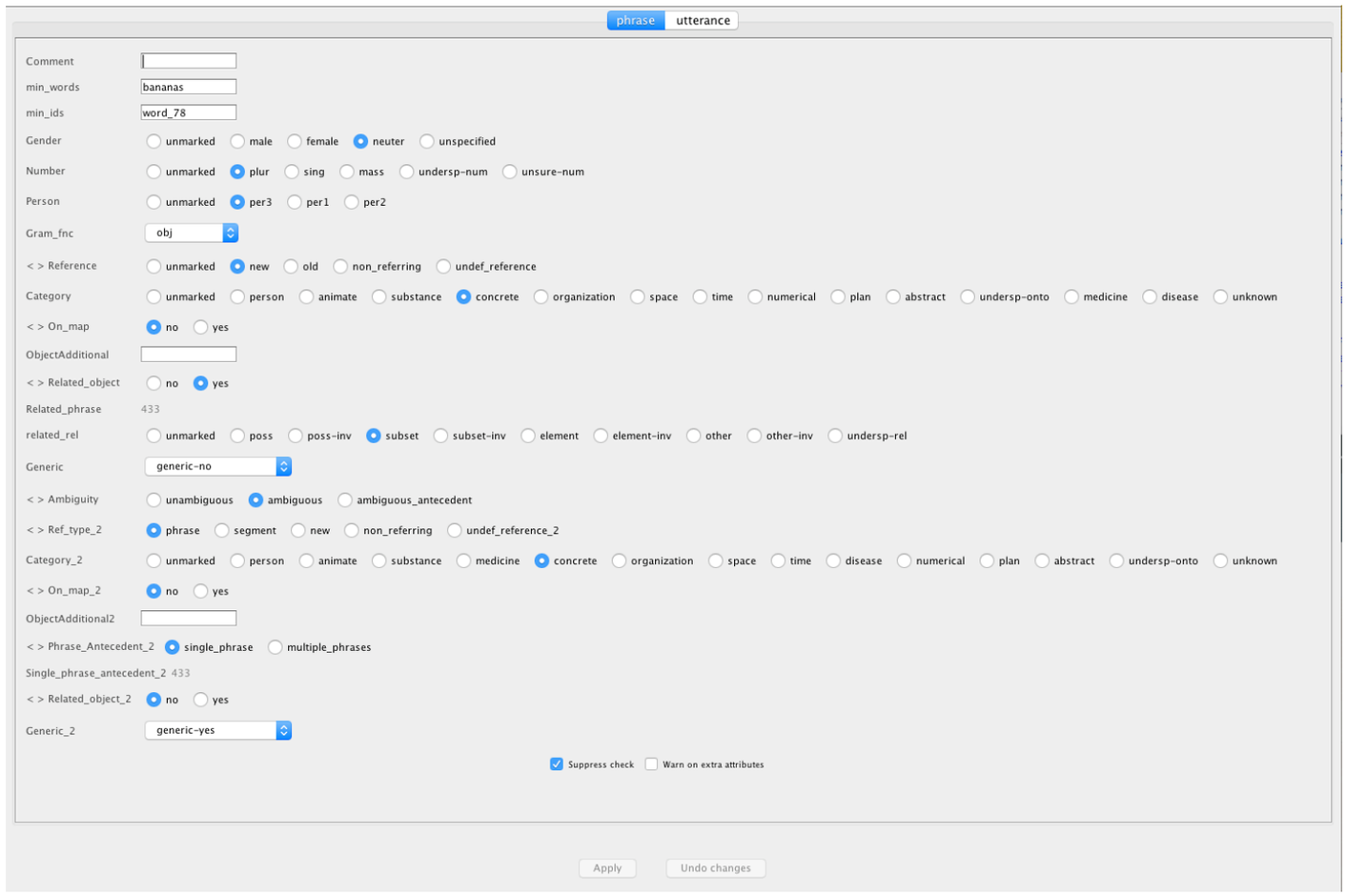
- complex and linguistically motivated annotation scheme
  - referring expressions
  - non-referring expressions
    - bridging references
    - discourse deixis
- annotation of a number of semantic properties
  - most importantly genericity
- annotation of anaphoric ambiguity (explicit)
- ALL NPs are considered markables

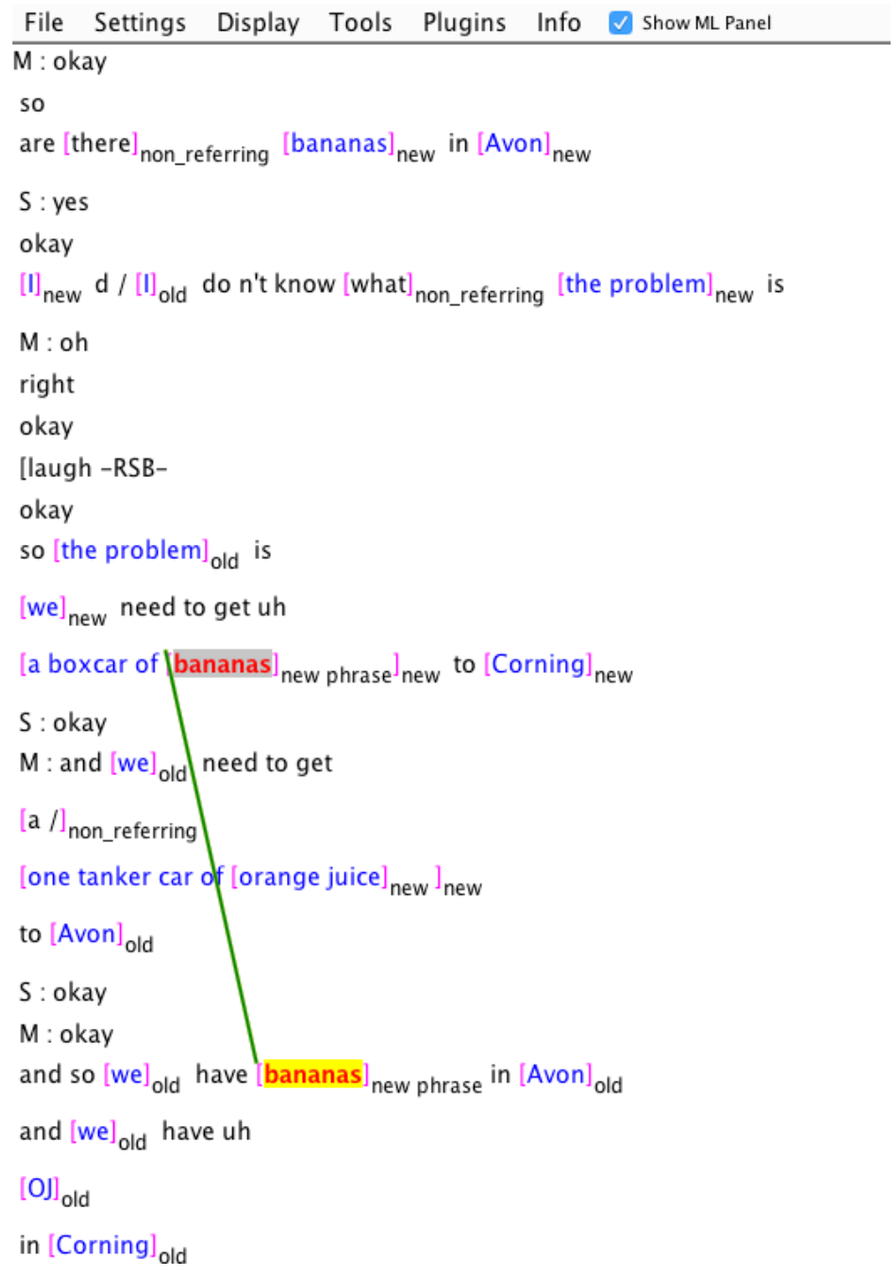
ARRAU – WHAT FOR

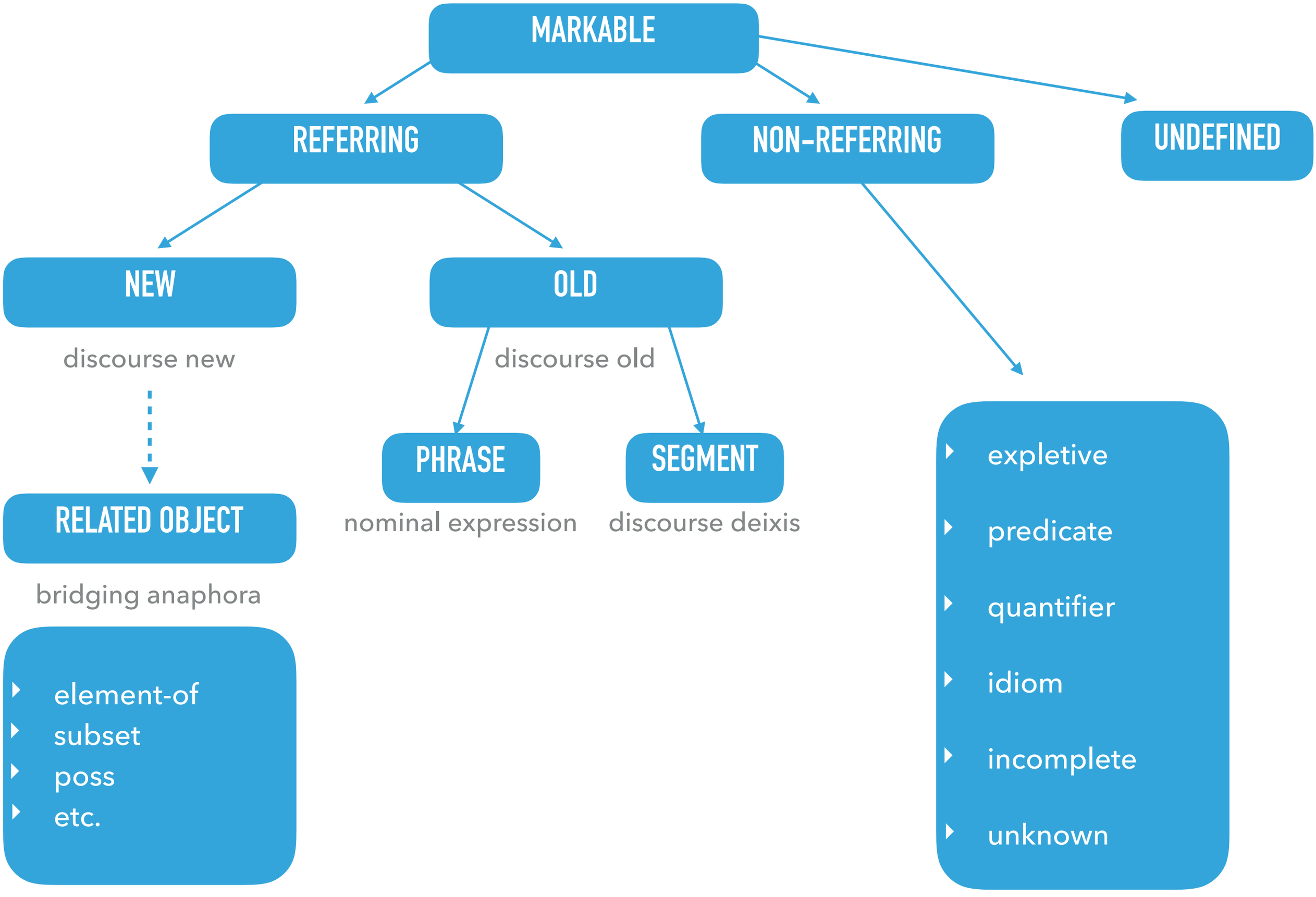
- resource for next generation of anaphora or coreference resolvers
- allows for joint modeling and deeper linguistic analysis of different anaphora-related phenomena

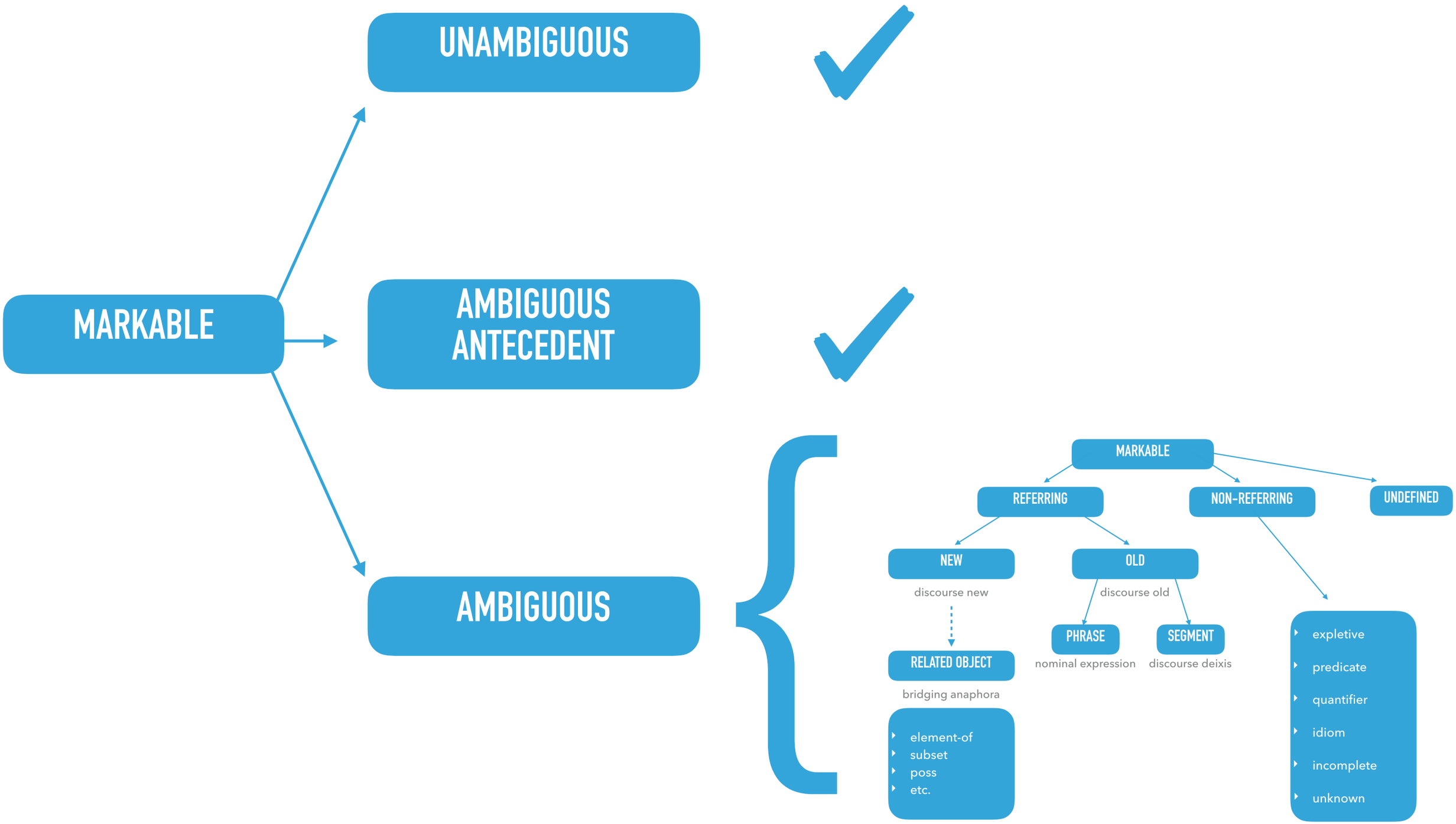
PROCEDURE

ANNOTATION TOOL

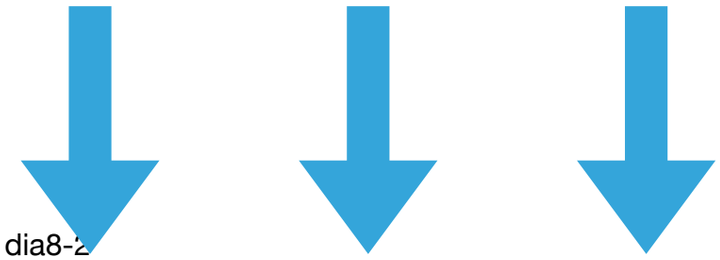





MARKABLE
REFERRING
NON-REFERRING
UNDEFINED
NEW
OLD
discourse new
discourse old
RELATED OBJECT
PHRASE
SEGMENT
nominal expression
discourse deixis
bridging anaphora
element-of
subset
poss
etc.
expletive
predicate
quantifier
idiom
incomplete
unknown


MARKABLE
UNAMBIGUOUS
AMBIGUOUS ANTECEDENT
AMBIGUOUS
MARKABLE
REFERRING
NON-REFERRING
UNDEFINED
NEW
discourse new
OLD
discourse old
RELATED OBJECT
bridging anaphora
element-of
subset
poss
etc.
PHRASE
nominal expression
SEGMENT
discourse deixis
expletive
predicate
quantifier
idiom
incomplete
unknown

SUBJECTIVE EVALUATION OF ANNOTATIONS



min_ids	min_words	NP	comment (annotator)	1st reading	2nd reading	Reason for Amb	Notes	keywords
word_14	I		-	DN	-		is this a fragment? the utterance is but not the NP. if yes, then non_referring in oder to annotated as incomplete	
word_413	it		nb you can't load an engine	DO; operator-instruction; antecedent 79 (E2 on-map)	DO; generic-no; antecedent is empty	referring to the engine while actually meaning the boxcar or mereological entity of the train	2nd has no antecedent	
word_1006	boxcar	a boxcar	semantic vs pragmatic coreference	DN; on_map; operator-question	DO; not on_map; generic-no	?	DN in 1st because of operator?	
colours	mereology	uncertainty	something is unmarked or missing a value	not annotated as ambiguous				

SUBJECTIVE EVALUATION OF ANNOTATIONS

wsjarrau_0037

min_ids	min_words	NP	comment (annotator)		1st reading						2nd reading						Reason for Amb	Notes	keywords
				Mereology	Generic	Reference (Info_Status)				Related (Bridging)	Generic	Reference (Info_Status)				Related (Bridging)			
						non-ref	new	old				non-ref	new	old					
								phrase (antecedent)	segment (deixis)					phrase (antecedent)	segment (deixis)				
word_585	players	[players]	mp: changed to operator-modal	no	modal	no	no	single	no	no	modal	no	yes	no	no	yes	old vs. new+bridging (subset-inv)		
word_663	it	[it]'s not worth ..	-	no	-	yes (explet.)	no	no	no	no	-	-	-	-	-	-		ambiguous: 2nd= old (word_645 or part of it)	no_amb
word_667	it	not worth [it]	-	no	-	yes (explet.)	no	no	no	no	-	-	-	-	-	-		ambiguous: 2nd=old (word_647 = money); a third reading of this and the one above could also be idiomatic	no_amb
word_1661	part	[[Hungary's]part of	changed to ambiguous between old and new	no	g-no	no	yes	no	no	yes	g-no	no	no	single	no	no	new+bridging (undersp-rel) +abstract vs. old+concrete	project vs. dam itself	
word_1684	dam	[the dam]	-	no	g-no	no	no	single	no	no	g-no	no	no	single	no	no	identifying right antecedent	the Hungarian damn seems to be the obvious choice; though meriological issues complicate the matter	mereology
Total of Markables Annotated as Amb in File: 3																			
colours	operator-scope lost	not annot. as amb.	highlight																

EXAMPLES

M : all right

so

put [the oranges]$_{old}$ in [the boxcar]$_{old}$

[that one [you]$_{old}$ picked up from [Dansville]$_{old}$]$_{old}$

S : right

M : hook on [the tanker car]$_{old}$

bring [it]$_{old\ phrase}$ to [Elmira]$_{old}$

examples

min_ids	min_words	NP	comment (annotator)		1st reading						2nd reading						Reason for Amb	Notes
				Mereology	Generic	Reference (Info_Status)				Related (Bridging)	Generic	Reference (Info_Status)				Related (Bridging)		
						non-ref	new	old				non-ref	new	old				
								phrase (antecedent)	segment (deixis)					phrase (antecedent)	segment (deixis)			
word_1529	it	[it]	mereology	yes	g-no	no	no	multiple	no	no	no	no	no	single	no	no	identifying the right antecedent	
word_498.. word_499	bowery mission	[the Bowery Mission]	mp: marked as ambiguous referent/ bridging	no	g-no	no	no	single	no	no	g-no	no	yes	no	no	yes	old+concrete vs. new+ bidging (undersp-rel)) +space	
Total of Markables Annotated as Amb in File:																		
colours		not annot. as amb.	highlight															

PROCEDURE

EXAMPLES

At [night]$_{non\_referring}$ [he]$_{old}$ returns to [the condemned building [he]$_{old}$ calls [home]$_{non\_referring}$]$_{new}$.

[[His]$_{old}$ life , including [[his]$_{old}$ skirmishes with [a competing sketch artist]$_{new}$]$_{new}$]$_{new}$, seems carefree .

[He]$_{old}$ is [[his own]$_{old}$ man]$_{non\_referring}$.

Then , just as [the Tramp]$_{old}$ is given [a blind girl to cure]$_{new}$ in [`` City Lights , ``]$_{new}$ [the Artist]$_{old}$ is put in [charge of returning [[a two-year-old waif]$_{non\_referring}$ Nicole Alysia , [[whose]$_{old}$ father]$_{new}$ has been murdered by [thugs]$_{new}$, to [[her]$_{old}$ mother]$_{new}$]$_{new}$]$_{non\_referring}$.

[This cute child]$_{old}$ turns out to be [[a blessing]$_{non\_referring}$ and [a curse]$_{non\_referring}$]$_{non\_referring}$.

[She]$_{old}$ gives [the Artist]$_{old}$ [a sense of [purpose]$_{new}$]$_{new}$, but also alerts [him]$_{old}$ to [the serious inadequacy of [[his]$_{old}$ vagrant life]$_{old}$]$_{new}$.

[The beds at [the Bowery Mission]$_{old\ new}$]$_{new}$ seem far drearier when [he]$_{old}$ has to tuck [a little girl]$_{old}$ into [one of [them]$_{old}$]$_{new}$ at [night]$_{old}$.

min_ids	min_words	NP	comment (annotator)		1st reading						2nd reading						Reason for Amb	Notes
				Mereology	Generic	Reference (Info_Status)				Related (Bridging)	Generic	Reference (Info_Status)				Related (Bridging)		
						non-ref	new	old				non-ref	new	old				
								phrase (antecedent)	segment (deixis)					phrase (antecedent)	segment (deixis)			
word_1529	it	[it]	mereology	yes	g-no	no	no	multiple	no	no	no	no	no	single	no	no	identifying the right antecedent	
word_498.. word_499	bowery mission	[the Bowery Mission]	mp: marked as ambiguous referent/ bridging	no	g-no	no	no	single	no	no	g-no	no	yes	no	no	yes	old+concrete vs. new+ bidging (undersp-rel)) +space	
Total of Markables Annotated as Amb in File:																		
colours		not annot. as amb.	highlight															

RESULTS

PRELIMINARY

REGARDING AMBIGUITIES

1st int	2nd int	RST	TRAINS	GNOME	PEAR	TOTAL
DO	DO	31	112	4	28	175
	DN	37	4	2	1	44
	DD	8	1	0	2	11
	NR	0	4	0	0	4
DN	DO	0	0	0	0	0
	DN	0	0	0	0	0
	DD	0	0	0	0	0
	NR	0	0	0	0	0
NR	DO	0	0	0	0	0
	DN	0	0	0	0	0
	DD	0	0	0	0	0
	NR	0	0	0	0	0
Total		76	121	6	31	234

Table 6: Distribution of ambiguity in the subdomains of ARRAU.

Uryupina et al. (2020:20)

	RST	TRAINS	GNOME	PEAR	OVERALL
all	72013	16999	6562	4008	99582
generic-yes	1438	728	12	74	2252 (2%)
operator-conditional	90	231	201	2	524
operator-instruction	15	163	211	-	389
operator-iquant	7	6	-	-	13
operator-modal	443	1080	147	16	1686
operator-question	54	432	39	10	535
operator-tquant	16	4	-	-	20
Total operator bound					3167 (3%)
underspecified-disease	-	-	84	-	84
underspecified-replicable	37	1	2	21	61
underspecified-substance	692	431	160	-	1283
underspecified-generic	1	3	-	-	4
Total underspecified					1432 (1.4%)

Table 3: Distribution of generic mentions in ARRAU.

Uryupina et al. (2020:14)

REGARDING AMBIGUITIES

- categorisation of ambiguities
  - identification of certain patterns in the annotation of attribute values
  - generalisation of reoccurring cases/patterns and assign a label to them, i.e. reason for ambiguity (e.g. "DO vs. DN", "DN vs. DN+bridging", "identifying the right antecedent")

REGARDING THE SCHEME AND MANUAL

- the attribute ambiguous_antecedent is not used
- complexity of generic attribute
- quantifiers cannot serve as antecedent
  - [Some [cats]] came in. [They] started to purr.

	RST	TRAINS	GNOME	PEAR
all	72013	16999	6562	4008
singletons	39431	5308	2887	1320
generic	2772	3071	819	122
generic	1434	728	12	74
episodic-no	-	4	-	-
no-generic	385	1285	470	-
operator-conditional	89	231	184	2
operator-instruction	15	163	205	-
operator-iquant	7	6	-	-
operator-modal	437	1078	142	16
operator-question	54	429	39	10
operator-tquant	16	4	-	-
underspecified-decease	-	-	75	-
underspecified-generic	1	3	-	-
underspecified-replicable	35	1	2	21
underspecified-substance	684	431	160	-
non-referential	9477	2351	1033	607
coordination	2408	231	326	37
expletive	444	851	75	122
idiom	638	148	29	42
incomplete	-	149	1	36
predicate	4252	145	352	79
quantifier	1724	817	250	132
unknown	9	6	-	159

Table 3: Distribution of markables in the second releases of ARRAU

Uryupina et al. (2016:61)

REGARDING THE SCHEME AND MANUAL

- the attribute ambiguous_antecedent is not used
- complexity of generic attribute
- quantifiers cannot serve as antecedent
  - [Some [cats]] came in. [They] started to purr.

	RST	TRAINS	GNOME	PEAR
all	72013	16999	6562	4008
singletons	39431	5308	2887	1320
generic	2772	3071	819	122
generic	1434	728	12	74
episodic-no	-	4	-	-
no-generic	385	1285	470	-
operator-conditional	89	231	184	2
operator-instruction	15	163	205	-
operator-iquant	7	6	-	-
operator-modal	437	1078	142	16
operator-question	54	429	39	10
operator-tquant	16	4	-	-
underspecified-decease	-	-	75	-
underspecified-generic	1	3	-	-
underspecified-replicable	35	1	2	21
underspecified-substance	684	431	160	-
non-referential	9477	2351	1033	607
coordination	2408	231	326	37
expletive	444	851	75	122
idiom	638	148	29	42
incomplete	-	149	1	36
predicate	4252	145	352	79
quantifier	1724	817	250	132
unknown	9	6	-	159

Table 3: Distribution of markables in the second releases of ARRAU

Uryupina et al. (2016:61)

FURTHER RESEARCH

TO BE CONTINUED . . .

HOW TO CONTINUE

- design of a scheme:
  - automatise the process to identify certain feature patterns
- identification of outliers of this scheme
  - are there regularities in those outliers?
- new insights about anaphoric ambiguity

HOW TO CONTINUE AND OPEN QUESTIONS

- generic attribute
  - should it be reevaluated?
- ambiguity attribute: make real use of the value ambiguous_antecedent
  - run a test study?
- quantified NPs
  - how can relationships of quantified NPs to other markables be integrated into the scheme?

REFERENCES

- Cohen, Ariel. 2016. Genericity. In Maria Aloni & Paul Dekker (eds.), The Cambridge Handbook of Formal Semantics, 285–310. 1st edn. Cambridge University Press
- Poesio, Massimo, Maris Camilleri, Paloma Carretero-Garcia & Ron Artstein. 2022. ARRAU 3 Annotation Manual Version 1.1.
- Uryupina, Olga, Ron Artstein, Antonella Bristot, Federica Cavicchio, Francesca Delogu, Kepa J. Rodriguez & Massimo Poesio. 2020. Annotating a broad range of anaphoric phenomena, in a variety of genres: the ARRAU Corpus. Natural Language Engineering 26(1). 95–128.
- Uryupina, Olga, Ron Artstein, Antonella Bristot, Federica Cavicchio, Kepa J Rodriguez & Massimo Poesio. 2016. ARRAU: Linguistically-Motivated Annotation of Anaphoric Descriptions. In Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16).