

The Museum of Annotation

- *best* practice in empirically-based dialogue research in ancient times
- major theoretical and technical breakthroughs in the past

Phase 1: Annotation with pencil and paper

- ca. 1995-1996
- anaphora resolution: Text by German writer Heiner Müller
- discourse structure: Text by German writer Uwe Johnson

MUELLER.TXT

(5) Er ging weiter auf und ab, so, daß er sie sehen konnte, sah aber nur auf, wenn er sicher war, sie konnte ihn nicht sehen.

Cb: Noddin

Cf: [Has, Noddin
Rabin

4.F		X.F	→	1.F	2.F	3.F	4.F	1.F	2.F	3.F
4.F		1.F	→	2.F	3.F	4.F	1.F	2.F	3.F	4.F
	→	2.F		3.F	4.F	1.F	2.F	3.F	4.F	1.F
		3.F		4.F	1.F	2.F	3.F	4.F	1.F	2.F
4.F		X.F	→	1.F	2.F	3.F	4.F	1.F	2.F	3.F
4.F		1.F	→	2.F	3.F	4.F	1.F	2.F	3.F	4.F
	→	2.F		3.F	4.F	1.F	2.F	3.F	4.F	1.F
		3.F		4.F	1.F	2.F	3.F	4.F	1.F	2.F

(6) Vor ihren Augen bemühte er sich, ein imponierendes Gesicht zu machen.

Cb: Has

Cf: [Has, Noddin, Rabin, Gesicht]

Smith - Hut

X.F		1.F
4.F		2.F
		3.F

(7) Zugleich versuchte er gelangweilt auszusehen, damit die Umstehenden nichts merkten.

Cb: Has

Cf: [Has, Umstehende]

Ca bin

X.F		1.F
4.F		2.F
		3.F

(8) Das war anstengend.

X.F		1.F
4.F		2.F
		3.F

MUELLER2.TXT

(36) Was sie denn so mache abends.

Ch: Nöckle

Ch: [Nöckle]

Ca Wum

1x.T	4.F	1.F	2.F	3.F
4.F		2.F	3.F	

(37) Sie beugte sich vor und er sah den Ansatz ihrer Brust, weiß unter dem gebräunten Hals.

Ch: Nöckle

Ch: [Nöckle, Hals, Brust, Ansatz, Hals]

Ca Wum

1x.T	4.F	1.F	2.F	3.F
1x.T	4.F	1.F	2.F	3.F

(38) Sie hat kleine Brüste, dachte er und fragte, ob sie "Die Junge Garde" kenne, das Buch von Fadejew.

Ch: Nöckle

Ch: [Nöckle, Hals, Brust, ...]

Ca Wum

1x.T	4.F	1.F	1.T
1x.T	4.F	2.F	2.F
1.T	4.F	3.F	3.T
2.F		1.F	—
3.T		2.F	
3.T		3.F	

(39) Es paßte ihm nicht, als sie nun Fragen stellte, und er antwortete einsilbig.

Ch: Nöckle

Ch: [Nöckle, Hals, Brust, Frage]

1x.T	4.F	1.F	1.F
1x.T	4.F	2.F	1.F

Fehler in (1), (2), (3)

falsche Auflösung	64 0	/	25 24 21 Fehler in (1), (2), (3)
Relevanz auf Proposition	2		
plurale Anapher	7 8	/	
Relevanz auf Element einer Menge	6 7		
Relevanz auf ? (-es)	1		

in (1) 1 zusätzliche falsche Auflösung = ~~21~~⁶ Fehler

7 ~~8~~ falsche Auflösungen fallen mit Semantik weg

in (2) ~~7~~³ zusätzliche falsche Auflösungen = ~~21~~²⁴ Fehler

8 ~~7~~ falsche Auflösungen fallen mit Semantik weg

12 ~~11~~ ¹⁷ Fehler

Heinrich Müller - Lichen geschichte

Teil 1 + 2

153 ~~152~~ Anaphora
867 Wörter

22
122
101 ~~101~~ selektivpropaguli Anaphora
29 ~~29~~ ~~29~~ selektiv propaguli Anaphora

A) ohne Samenli

	(1)	(2)	(3)	(4)
T	27 27 27 27 27 83,7%	91 91 91 91 91 64,7%	125 125 125 125 125 11,7	105 105 105 105 105 7,2
F	26 26 26 26 26 16,3%	54 54 54 54 54 25,3%	28 28 28 28 28 11,1%	34 34 34 34 34 24,2%

B) mit Samenli

	(1)	(2)	(3)	(4)
		114		

(21) Jeden dritten Tag war er mit Wasser und Poliermitteln zugange, rieb die Ledersitze ein, wischte sie blank, putzte den Lack.

X
0+0

Ch: -

Cf: E.B., Spucke, Ledersitze, Lack, Tap, Elwan, Poliermittel

-

(22) Noch an pulvertrockenen Tagen hatte das hochbeinige sprungs"ichtige Ding am Rinnstein ein Aussehen, als st"unde es im Schaufenster.

X
+

Ch: Spucke

Cf: E Spucke, Tap, Rinnstein, Aussichts, Klappen

Stoß-Platz

(23) Er bedankte sich f"ur die Bewunderung der umstehenden Jungen und fuhr sie hin und her "uber den Marktplatz.

X
+

Ch: -

Cf: E.B., Bewunderung, Tap, Marktplatz

-

(25a) B. glaubte auch zu bemerken, da"s er es mit diesem Wagen leichter hatte bei den M"adchen,

X+X

CG: -

CF: [B, Sportwagen, Modell]

-

(25b) und einmal h"atte er sich fast geradeheraus erkundigt.

CG: B

CF: [B]

Continu

(26) Aber er verga"s es "uber dem vollkr"aftigen Toben der Maschine auf der dreibahnig ausgebauten Umgehungsstra"se, die er allabendlich aufsuchte, dem Motor zuliebe.

</par>

O
+
+

CG: B

CF: [B, Sportwagen, Motor, Toben, Umgehungsstra"se]

Continu

<par>

(27a) Er f"uhrte den Wagen mit sich wie andere Leute die Uhr,

X

CG: B

CF: [B, Wagen, Uhr]

Johnson

1	-	B	1	
2	C	B		
3	C	B		
4	C	B		
5	C	B		
<hr/>				
6	R	Values	2	
7	SI	Values		
8	SI	B		
9	C	B		
10	C	B	3	
<hr/>				
11	-	Spillage		4
12	R	B		
13	SI	Bank		5
14	R	R		
15	SI	B		
<hr/>				
16	R	Spillage	6	
17	C	B		
18	C	B		
19	R	Spillage		
20	C	B		
21	C	B		
<hr/>				
22	SI	Spillage	6	
23	-	B		
24	C	B	6	
25	-	Bank		

Johns

falsch paratext bei Anspiel

(240) "die Urein", Anbruch in (22) nicht erreichbar,
 stattdessen Anspiel auf $(p(\text{L.A.}, \text{Kosmos}, \text{a.d.}), = (16))$
 "Spurwege"

(311) "die Wege", Anbruch in (28) o. (29) nicht
 erreichbar, stattdessen Anspiel auf $(p(\text{L.A.}, \text{Kosmos}, \text{a.d.}), = (16))$
 "Spurwege"

falsch paratext bei behandelte Elipen

(16) "die Konstruktion", Anbruch in (14) nicht erreichbar,
 stattdessen Anbruch in $(p(\text{L.A.}, \text{Kosmos}, \text{a.d.}), = (11))$
 "flüchtig"

(26) "die Maschine", Anbruch in (21) nicht erreichbar,
 stattdessen Anbruch in (16) = $(p(\text{L.A.}, \text{Kosmos}, \text{a.d.}), = (11))$

Folium

Anticidat amur Amplicon

U:		10
Cp(C ₁ , U ₁₋₁)		36
Cp(C ₁ -A, U ₁ C ₁ -A ad3)		9
Cp(C ₁ -A, U ₁ C ₁ -A ad3)		1
Cp(C ₁ -2, U ₁ C ₁ -2 ad3)		

Other

U ₁ Amplicon		2
-------------------------	--	---

 2 false positives, U₁ Amplicon

58

Anticidat amur Kalkulle Cligpion

Cp(C ₁ , U ₁ -1)		8
Cp(C ₁ -A, U ₁ C ₁ -A ad3)		3
Cp(C ₁ -A, U ₁ C ₁ -A ad3)		1
Cp(C ₁ -2, U ₁ C ₁ -2 ad3)		

Johnson

Antrodiaet emu Amylopsis

U ₁		10
U ₁₋₂		32
U ₁₋₂		10
U ₁₋₂		2
U ₁₋₂		1
U ₁₋₅		
U ₁₋₆		1
		<hr/> 56
P ₁₋₆		2
		<hr/> 58

Antrodiaet emu Hebeella Obipic

U ₁₋₁		8
U ₁₋₂		4
U ₁₋₃		

Summary: Annotation with pencil and paper

- Advantages:
 - easy to produce
 - allows to get good overview
- Disadvantages:
 - analysis/report manually
 - impossible to reproduce
 - impossible to exchange or reuse

Phase 2: Annotation machine-readable, reporting semi-automatically

- ca. 1997-1998
- anaphora resolution, text taken from NYT
- pronoun resolution in spoken dialogue, Switchboard

```

emacs@chimera.myth.com
Buffers Files Tools Edit Search Mule Help

\item
\orig{When \ent{Kenneth L. Curtis}{1} was wheeled into \ent{court}{2}
nine \ent{years}{3} ago, mute, dull-eyed and crippled,}

\cb{--}{--}

\begin{bfp}
\item[Curtis] Curtis
\item[Court] court
\item[Years] years
\end{bfp}

\item
\orig{it seemed clear to nearly \ent{everyone}{4} involved}

\cb{--}{--}

\begin{bfp}
\item[Everyone] everyone
\end{bfp}

\item[a.]
\orig{that it would be pointless to put \pana{him}{1} on
\ent{trial}{5} for the \ent{murder}{6} of \poss{his}{1} former
\ent{girlfriend}{7}, \ent{Donna Kalson}{8}, and the \ent{wounding}{9}
of \poss{her}{7} \ent{companion}{10}.}

\cb{--}{--}

--:-- curtis.tex Mon Oct 16 1:29PM (LaTeX Fill)--L45-- 2%-----

```


intra - Fric / Ham

Shot in Head, Suspect Goes Free, Then to College

Jane Fritsch

In: The New York Times. Vol. CXLVII, No. 50, 999, Dec. 7, 1997, p.A45+A48

Paragraph 1

Segment 1

- (1) When Kenneth L. Curtis₁ was wheeled into court₂ nine years₃ ago, mute, dull-eyed and crippled,

Cf: [CURTIS: Curtis
COURT: court
YEARS: years]

- (2) it seemed clear to nearly everyone₄ involved

Cf: [EVERYONE: everyone]

- a. that it would be pointless to put *him*₁ on trial₅ for the murder₆ of his₁ former girlfriend₇, Donna Kalson₈, and the wounding₉ of her₇ companion₁₀.

Cf: [EVERYONE: him, his wrong (inter)
GIRLFRIEND: girlfriend, her wrong (inter)
MURDER: murder correct
COMPANION: companion
WOUNDING: wounding
TRIAL: trial]

Paragraph 2

- (3) It had been a year₁₁

Cf: [YEAR: year]

- (4) since *Mr. Curtis* had slammed his pickup truck into *them*, break-

(6) and, finally, ϵ_1 fired a bullet₁₅ into his₁ own brain₁₆.

Cf: [CURTIS:	- , his	<i>correct</i>
BRAIN:	brain	
BULLET:	bullet]	

(7) *Mr. Curtis*₁ lingered in coma₁₇ for months₁₈,

Cf: [CURTIS:	Mr. Curtis
COMA:	coma
MONTHS:	months]

(8) then ϵ_1 awoke to a world₁₉ of paralysis₂₀, pain₂₁ and mental confusion₂₂

Cf: [CURTIS:	-
WORLD:	world
{PARALYSIS, PAIN, CONFUSION}:	paralysis, pain, confusion]

(9) from which psychiatric experts₂₃ said

Cf: [WORLD:	which
EXPERTS:	experts]

a. *he*₁ would never emerge.

Cf: [CURTIS:	he]	<i>correct</i>
---------------	------	----------------

Paragraph 3

Segment 2

(10) One expert₂₄ calculated his₁ I.Q.₂₅ at 62₂₆.

Cf: [EXPERT:	expert, his	<i>wrong</i> (<i>seg-</i>
I.Q.:	I.Q.	<i>ment</i>)

Statistics:

Words	1425
Utterances	90
Anaphora	169
Pronominal Anaphora	66
Common	64
Local	0
Plural with Splitted Antecedents	2
Nominal Anaphora	28
Common	24
Demonstrative	3
Plural with Splitted Antecedents	1
Possessive Pronouns	32
Zero Pronouns	0
Ellipses	12
Discourse Deictics	0
Physical Deictics	0
Anaphoric Proper Names	43
Textual Ellipsis	0

Pronoun Resolution:

Utterances	90
Pron. Anaphora & Poss. Pron.	86
Correct	63
Wrong	23
Wrong (inter)	16
Wrong (split)	2
Wrong (event)	0
Wrong (chain)	3
Wrong (other)	0
Wrong (strategy)	1
Wrong (ambiguity)	0
Wrong (segment)	1
Wrong (inferred)	0

```

emacs@chimera.myth.com
Buffers Files Tools Edit Search Mule Help
BACKGROUND_A: 2
BACKGROUND_B: 2
REMARKS:      None.

\end{verbatim}

\begin{longtable}{lllp{4in}l}

\sy{x-} & \da{I} & A.1: & \na{Okay, well}, having, \na{uh},
personally, \na{uh}, helped, \na{uh}, \ent{two children} attend,
\na{or}, or to select \ent{colleges}, \splitnana{one of \ipana{them}} not
very successfully \na{[laughter]}, \panai{I}'m not sure if \panai{I}'m
any authority on \demmp{this topic} at all. & \\\

\sy{-x} & \da{R} & B.2: & Well, & \\\

\sy{x-} & \da{I} & & which \ent{colleges} did \ipana{they} go to? & \\\

\sy{-x} & \da{R} & A.3: & Well, & \\\

\sy{x-x} & \da{I} & & \ent{a daughter}, \na{uh}, went to \ent{the
University of Virginia}, & \\\

\sy{x-} & \da{I} & & \idem{that} was a good
selection. \na{\ipana{She}}, & \\\

\sy{-xx-} & \da{R/I} & B.4: & \idem{That}'s a beautiful school. & \\\

\sy{-x} & \da{R} & A.5: & Yeah, \elli really is. & \\\

--:-- sw3117.tex      Mon Oct 16 1:20PM 0.28 (LaTeX Fill)--L25-- 2%-----

```


Switchboard 3117.txt

FILENAME: 3117_1292_1312
 TOPIC#: 324
 DATE: 910627
 TRANSCRIBER: glp
 DIFFICULTY: 2
 TOPICALITY: 2
 NATURALNESS: 2
 ECHO_FROM_B: 2
 ECHO_FROM_A: 2
 STATIC_ON_A: 1
 STATIC_ON_B: 1
 BACKGROUND_A: 2
 BACKGROUND_B: 2
 REMARKS: None.

1- I A.1: ~~Okay~~, well, having, uh, personally, uh, helped, uh, two children attend, or, or to select colleges, *one of* **them** not very successfully [laughter], I'm not sure if I'm any authority on this topic at all.

-1 R B.2: Well,

2- I which colleges did **THEY** go to?

-2 R A.3: Well,

3- I *a daughter*, uh, went to the University of Virginia, **THAT** was a good selection. She,

+ -3/4- R/I B.4: **THAT**'s a beautiful school.

-4 R A.5: Yeah,

5- I e really is.

-5 R B.6: Beautiful school.

+ 6-6 I A.7: And she was, **SHE** was very, [singing] very happy *there* and v-, and very successful in what **SHE** picked to do,

+ 7-7 I so, uh, **SHE**'s just graduated from *there*,

+ 8-8 I so **THAT** was a good selection.

9-9 I Uh, ~~our son~~ has this kind of, you know, **HE**'s, well, **HE** started out going Stephen F Austin

21-21 I so. Uh, but, I, I think, you know, just in terms of, of what happened, I think that first of all I, I would know the reputation of the school and, uh, for more than just academics [barking], but, you know, if [IT] is a party school or not.

+ 22- I Although, I think every school [noise] is a party school to some degree personally [laughter].

-22 R B.12: Well,

23-23 I that's true.

+ 24- I Of course, uh, you know, for me, my son, [HE] is only six -

-24 R A.13: Right.

25-25 I B.14: - and, but I had the experience with my sister going to school and [SHE] was, I treated my sister almost like my daughter -

-26 R A.15: Uh-huh.

+ 27- I B.16: - because [SHE] was nine, ten years younger than I was.

-27 R A.17: Right.

+ 28- I B.18: And [SHE] ended up going to the University of Oklahoma.

-28 R A.19: Uh-huh.

+ 29-29 I B.20: I can say that because [IT] was a big well known school,

+ 30-30 I [IT] had a well known education

+ 31- I and [IT] was a wonderful education -

-31 R A.21: Yeah.

32- I B.22: - but, yet [IT] was a small school.

-32 R A.23: Right.

33- I B.24: Whereas I went to the University of Texas -

-33 R A.25: Yeah.

34- I B.26: - which was, you know, a million people go there -

-34 R A.27: Yeah.

35-35 I B.28: - I think that for me it's going to be one of those things, yes,

- 36-36 I I want to make sure [IT] has a good education,

+ 37- I but I also want to keep [IT] small enough that you can enjoy [IT] because Texas to me was almost too big.

-37 R A.29: Yeah,

38- I I've heard that from people who have gone to Texas.

+ -38/39- R/I B.30: [IT] is way too big.

-39 R A.31: Yeah.

40- I I went to, uh, Iowa State, which is also a big eight school, -

-40 R B.32: Yes.

41- I A.33: - like Oklahoma,

-41 R B.34: Yes,

42- I I'm, I'm quite aware.

-42 R A.35: Yeah.

+ 43- I And, uh, when I was there [IT] was maybe eleven or twelve thousand people which

-52 R B.46: So,
 53- I uh, course, now ~~we're~~, we're already telling our son, you know, Tex-, uh Okla-
 homa's good and Arkansas is good.

-53 R A.47: Uh-huh.
 54- I B.48: Now, of course, *Texas* is good only because, you know -
 -54/55- R/I A.49: *you* went to school *there*, yeah.
 -55/56- R/I B.50: - ~~I~~-, because *I* went *there*, of course.
 -56 R A.51: Right, uh-huh.
 57- I B.52: But, uh, ~~we're going~~, my parents let *me* choose on my own -
 -57 R A.53: Uh-huh.
 + 58-58 I B.54: - and *I* want to pretty much allow HIM to choose on his own,
 + 59- I but *I* kind of keep an ear open as to what HE's doing.
 -59 R A.55: Sure, oh, yeah, yeah.
 + 60-60 I Well, our daughter chose, uh, *Virginia* because SHE didn't want to go to school
 in *Texas*.
 + 61- I I mean SHE just sort of didn't want to go to school *here*.
 -61/62- I B.56: Well, ~~I was~~ when *I* was in Louisiana in Shreveport -
 -62 R A.57: Uh-huh.
 63- I B.58: - *I* would go to any school except L S U.
 -63 R A.59: Yeah.
 64- I B.60: *I* did not want to school -
 - R A.61: Same deal.
 - I B.62: - in-the, in the same state.
 -64 R A.63: Yeah.
 + 65-65 I And, uh, in fact, SHE was accepted at two schools in *Texas*, uh, *Rice* and *Trinity*,
 + 66-66 I but SHE decided SHE wanted to go out of state
 + 67-67 I and, uh, *Virginia* was, uh, you know, in the east.
 + 68-68 I IT wasn't a, you know, IT wasn't an Ivy League school,
 + 69-69 I but IT was, well, IT's quite a bit further
 + 70-70 I and IT was, uh, of the state schools that SHE could have gone to IT was one of
 the cheapest, yet one of the best, +, you know, with the academics and the tradition
 and everything that is *there*.
 + 71-71 I Uh, SHE made a really good selection, we thought
 + 72-72 I and, so, uh, SHE pretty much made THAT her-, herself,
 + 73-73 I but SHE did have some others in the wings.
 + 74-74 I You know, there's another school that SHE, SHE would have, have gone to if SHE
 hadn't been accepted *there*.
 75-75 I So, uh, [lipsmack] *I* don't know,
 76- I I think it really depends a lot on the child, because our daughter is, was just a
 lot more levelheaded about her-proc, the process.
 -76/77- R/I B.64: Luckily *I* still have twelve more years to worry about it.

Statistics:

Words	0
Utterances	0
Anaphora	295

Anaphoric Form

Pronominal Anaphora (3rd)	95
Pronominal Anaphora (1st/2nd)	125
Demonstratives	28
NP	34
Proper Names	0
Demonstrative NP	1
Possessive Pronouns	23
Zero Pronouns	0
Ellipses	2
Total	295

Types of Reference

Individual		262
Plural with Split Antecedents	1	
Split Sing. Anaphors with Plural Antecedents	7	
First/Second Person Sing.	125	
Physical Deictics	0	
Abstract		37
Discourse Deictics	20	
Vague Reference	8	
IEP	9	
Total		295

Pronoun Resolution:

Utterances	0
Pron. Anaphora & Poss. Pron.	0
Correct	0
Wrong	0
Wrong (inter)	0
Wrong (split)	0
Wrong (event)	0
Wrong (chain)	0
Wrong (other)	0
Wrong (strategy)	0

Sw 3/17 pro

me.tex	ms.tex					
1	1	A.1	19	19	ipana	ipana
2	2	B.2	5	5	ipana	ipana
3	3	A.3	16	16	ipana	ipana
4	4	A.7	2	2	ipana	ipana
5	5	A.7	4	4	ipana	ipana
6	6	A.7	18	18	ipana	ipana
7	7	A.7	24	24	ipana	ipana
8	8	A.7	45	45	ipana	ipana
9	9	A.7	48	48	ipana	ipana
10	10	A.7	75	75	ipana	ipana
11	11	A.7	84	84	ipana	ipana
12	12	A.7	95	95	ipana	ipana
13	13	A.7	109	109	ipana	ipana
14	14	A.7	115	115	ipana	ipana
15	15	A.9	3	3	ipana	ipana
16		B.10	20	20	ipana	splitantep
17	16	B.10	22	22	ipana	ipana
18	17	A.11	7	7	ipana	ipana
19	18	A.11	9	9	ipana	ipana
20	19	A.11	25	25	ipana	ipana
21	20	A.11	29	29	ipana	ipana
22	21	B.12	14	14	ipana	ipana
23	22	B.14	15	15	ipana	ipana
24	23	B.16	3	3	ipana	ipana
25	24	B.18	2	2	ipana	ipana
26	25	B.20	6	6	ipana	ipana
27	26	B.20	13	13	ipana	ipana
28	27	B.20	20	20	ipana	ipana
29	28	B.22	4	4	ipana	ipana
30	29	B.28	22	22	ipana	ipana
31	30	B.28	33	33	ipana	ipana
32	31	B.28	40	40	ipana	ipana
33	32	B.30	1	1	ipana	ipana
34		A.35	8	8	vagpro	ipana
35	33	A.35	24	24	vagpro	vagpro
36	34	A.35	26	26	vagpro	vagpro
37	35	B.44	11	11	ipana	ipana
38	36	B.44	12	12	ipana	ipana
39		A.45	6	6	vagpro	ddpro
40	37	B.54	9	9	ipana	ipana
41	38	B.54	26	26	ipana	ipana
42	39	A.55	12	12	ipana	ipana
43	40	A.55	24	24	ipana	ipana
44	41	A.63	6	6	ipana	ipana

65	62	B.78	13	13	ieppro	ieppro	
66		B.78	16	16	vagpro	ipana	
67	63	A.93	12	12	ipana	ipana	
68	64	A.93	52	52	ddpro	ddpro	+
69	65	A.95	6	6	ddpro	ddpro	+
70		B.96	11	11	ipana	vagpro	
71	66	A.97	19	19	vagpro	vagpro	
72	67	A.97	35	35	ipana	ipana	
73	68	A.101	1	1	ipana	ipana	
74	69	B.102	3	3	ipana	ipana	
75	70	B.102	7	7	ipana	ipana	
76	71	A.105	19	19	ipana	ipana	
77	72	A.105	27	27	ipana	ipana	
78	73	B.106	6	6	ipana	ipana	
79	74	B.106	12	12	ipana	ipana	
80	75	B.106	18	18	ipana	ipana	
81	76	B.108	4	4	ipana	ipana	
82	77	A.111	10	10	ieppro	ieppro	
83	78	B.118	1	1	ipana	ipana	
84		B.120	1	1	ipana	vagpro	
85	79	A.123	2	2	ieppro	ieppro	
86		A.123	6	6	ipana	vagpro	
87	80	A.123	18	18	ieppro	ieppro	
88	81	A.123	20	20	ieppro	ieppro	
89	82	A.123	44	44	ieppro	ieppro	
90	83	A.127	16	16	ieppro	ieppro	
91	84	A.127	26	26	ieppro	ieppro	
92	85	A.127	28	28	ipana	ipana	
93	86	A.127	67	67	ipana	ipana	
94	87	A.127	70	70	ipana	ipana	
95	88	B.140	2	2	vagpro	vagpro	
96	89	B.140	3	3	vagpro	vagpro	
97	90	B.140	10	10	ipana	ipana	

All Turns: 701 177 1225 524

Pronouns:

IPANA: 148

DDPRO: 5

VAGPRO: 20

IEPPRO: 20

N: 97

Z: 90

T: 194

PA: 0.927835051546392

PE: 0.603916462960995

KAPPA: 0.817803716374857

Promoucs

	SW 2403	SW 3117	SW 3241	Σ
IPANA	120	148	5	273
DDPRO	33	5	9	47
VAG PRO	31	20	26	77
LEPRO	24	20	86	132
N	104	107 97	63	264
Z	53	90	58	281
T	208	194	126	528
PA	180 0.7980	213 0.9271	92 0.9206	275 0.8777
PE	119 0.3935	80 0.6019	52 0.5151	159 0.5777
KAPPA	167 0.6670	122 0.8178	74 0.8367	205 0.105

$$\chi^2 = (273/528)^2 + (47/528)^2 + (77/528)^2 + (132/528)^2 -$$

Summary: Machine-readable annotation, reporting semi-automatically

- Advantages
 - reproducible
 - can be corrected after the fact
 - reporting semi-automatically including statistics
 - allows to get good overview
- Disadvantages:
 - hard to produce because no graphical user interface
 - reporting only semi-automatically
 - almost impossible to reuse data

Phase 3: Tool-based annotation, reporting automatically

- ca. 1999-2000
- pronoun resolution, dialogue act tagging in spoken language, Switchboard
- anaphora resolution in written text, Brown

Referee: sw_0380_4572.du
□ X

File Tools Settings Mode
Help

really the government that MUMBLEx .

So , in **that sense** , **it** , at least , philosophically , **it** 's really borderline legal .

SpeakerA19 .

Uh-huh .

Well I , I think **the** , I think that , you know , a person's **competence** should be more determined by , you know , **their** and **their behavior**

and if **they're** , you know , if **they're** on **some kind of** , you know , **controlled substance** then , you know , **it** kinds of away from **the job of** , you know , **the supervisor or manager person** , you know , **who is able to evaluate them** .

I think **they** should be evaluated **that way** rather than with **know** , **a chemical** , you know ,

SpeakerB20 .

Well , what about **this** , uh , **AIDS epidemic now** ? Where you , uh , uh , you know , **certain people are actually are** , are **in their disregard for it-** , **others** .

SpeakerA21 .

Add Coreferring Expr	Ex_1 (3.25) "you" Ex_2 (3.53) "privacy" Ex_3 (7.11) "I"
Coreferring exprs:	Ex_117 "they" Ex_127 "them" Ex_131 "they"
Current expression:	Ex_131 (138.8) "they"

Set attributes for expression
□ X

they

IT Class none Expl IPro DDPro VagPro

Semantic Role none VOC DIR LOC MNR PRP TMP

Dem Class none IDem DDDem VagDem

THEY Class none IPro IEPro

NP Form none PRP PRP\$ defNP indefNP DTpro NNP

Case none NOM OBJ OBL

NP Depth 0 1 2 3 4 5 6 7

Text Cat none HLN LST TTL

S Depth 0 1 2 3 4 5 6 7

Grammatical Role none CLF NOM ADV LGS PRD SBJ TPC CLR

Notes:

Mark Un-Indir Un-Indir Un-Indir

Un-Indir Un-Indir Un-Indir

Coref Coref Coref

debian 2.2
1
2
3
4
chi...
Ref...
Set...
Ref...
Ref...
Set...
14:20

Annotation based on Penn Treebank

```
( (CODE (SYM SpeakerA3) (. .) ))
( (S
  (INTJ (UH Oh) )
  ( , , )
  (NP-SBJ (PRP I) )
  (VP (VBP do) (RB n't)
    (VP (VB know) ))
  (. .) (-DFL- E_S) ))
( (S
  (NP-SBJ-1 (PRP I) )
  (VP (VBD had)
    (NP
      (NP
        (ADJP
          (NP-ADV (DT a) (JJ little) (NN bit) )
          (JJR more) )
          (NN time) )
        (SBAR
          (WHADVP-2 (-NONE- 0) )
          (S
            (NP-SBJ (-NONE- *-1) )
            (VP (TO to)
              (VP (VB think)
                (PP (IN about)
                  (NP (PRP it) ))
                  (ADVP-TMP (-NONE- *T*-2) ))))))))
  (. .) (-DFL- E_S) ))
```

Referee: Segment Marking

File Tools Settings Mode

SpeakerA1 .
So , uh , what things do you consider an inva
SpeakerB2 .
Ooh , well I don't know .
What do you think ?
SpeakerA3 .
Oh , I don't know .
I had a little bit more time to think about i
I was thinking of , like , uh ,
I don't know ,
I was ju- , started to think about all the bi
bases they have with all the information about
credit reports and all those , you know , dem
they do that , um , you know , have , who kno
, stuff about all the purchases that you've m
tract somewhere
SpeakerB4 .

Mark Segment Vi

Referee: View Segments

File Tools Settings Mode Help

<BEGIN Seg 33>
So what that means is , to me , that really it 's the government that
's requiring this ,
<END Seg 33>

<BEGIN Seg 34>
although they don't actually e- , execute it themselves , it 's
really the government that MUMBLEx .
<END Seg 34>

<BEGIN Seg 35>
So , in that sense , it , at least , philosophically , it 's really
borderline legal .
<END Seg 35>

SpeakerA19 .
<BEGIN Seg 36>
Uh-huh .
<END Seg 36>

<BEGIN Seg 37>
Well I , I think the , I think that , you know , a person's
competence should be more determined by , you know , their actions
and their behavior
<END Seg 37>

Close Refresh

Set attributes for segment

No segment selected

Type NO R I R/I

Apply

debian 2.2

1 2 3 4 chi... Ref... Set... Ref... Ref... Set... 14:28

File structure in Referee

```
/home/strube/exx/dial/annot/katy/second/4572
```

```
(0) 130> ls -al
```

```
total 120
```

```
drwxr-xr-x  2 strube  eml           4096 Feb 28  2000 .
drwxr-xr-x  7 strube  eml           4096 Mar 23  2000 ..
-rw-r--r--  1 strube  eml        23803 Mar  2  2000 .sw_0380_4572.du.attr
-rw-r--r--  1 strube  eml        5173 Mar  2  2000 .sw_0380_4572.du.info
-rw-r--r--  1 strube  eml           0 Mar  2  2000 .sw_0380_4572.du.link
-rw-r--r--  1 strube  eml         334 Mar  2  2000 .sw_0380_4572.du.note
-rw-r--r--  1 strube  eml        1595 Mar  2  2000 .sw_0380_4572.du.seg
-rw-r--r--  1 strube  eml        1526 Mar  2  2000 .sw_0380_4572.du.segat
-rw-r--r--  1 strube  eml           0 Mar  2  2000 .sw_0380_4572.du.time
-rw-r--r--  1 strube  eml        5157 Feb 17  2000 sw_0380_4572.du
-rw-r--r--  1 strube  eml       39428 Feb 17  2000 sw_0380_4572.mrg
-rw-r--r--  1 strube  eml       19835 Feb 17  2000 sw_0380_4572.new1
```

```
(0) 131>
```


Coreference

27	22	64	26	15	0	0
28	22	21	26	15	5	0
29	21	35	26	15	0	0
30	36	4	36	5	0	0
31	38	0	38	1	0	0
32	38	22	38	31	0	0
33	38	48	38	52	0	0
34	41	4	41	5	0	0
35	41	11	41	15	0	0
36	41	20	41	24	5	28
37	41	27	41	28	0	0
38	41	34	41	38	0	0
39	41	44	41	49	5	28
40	52	13	52	23	0	0
41	51	55	52	23	0	0
42	51	23	52	23	6	0
43	58	0	58	4	6	42
44	58	11	58	15	6	43
45	58	18	58	22	6	44
46	58	41	58	42	0	0

Attributes on markables (referring expressions)

(1)(S Depth)(0)(Semantic Role)(none)(NP Form)(PRP)(Grammatical Role)(SBJ)
(2)(S Depth)(0)(Semantic Role)(none)(NP Form)(none)(Grammatical Role)(non
(3)(S Depth)(0)(Semantic Role)(none)(NP Form)(PRP)(Grammatical Role)(SBJ)
(4)(S Depth)(0)(Semantic Role)(none)(NP Form)(PRP)(Grammatical Role)(SBJ)
(5)(S Depth)(0)(Semantic Role)(none)(NP Form)(PRP)(Grammatical Role)(SBJ)
(6)(S Depth)(0)(Semantic Role)(none)(NP Form)(PRP)(Grammatical Role)(SBJ)
(7)(S Depth)(0)(Semantic Role)(none)(NP Form)(indefNP)(Grammatical Role)(
(8)(S Depth)(1)(THEY Class)(none)(Case)(OBL)(NEUTER Class)(Anaph)(NP Form
(9)(S Depth)(0)(Semantic Role)(none)(NP Form)(none)(Grammatical Role)(non
(10)(S Depth)(0)(Semantic Role)(none)(NP Form)(PRP)(Grammatical Role)(SBJ
(11)(S Depth)(0)(Semantic Role)(none)(NP Form)(PRP)(Grammatical Role)(SBJ
(12)(S Depth)(0)(Semantic Role)(none)(NP Form)(PRP)(Grammatical Role)(SBJ
(13)(S Depth)(0)(Semantic Role)(none)(NP Form)(PRP)(Grammatical Role)(SBJ
(14)(S Depth)(1)(THEY Class)(IEPro)(Case)(NOM)(NEUTER Class)(none)(NP For
(15)(S Depth)(1)(Semantic Role)(none)(NP Form)(PRP)(Grammatical Role)(non
(16)(S Depth)(1)(THEY Class)(Anapha)(Case)(OBL)(NEUTER Class)(none)(NP Fo
(17)(S Depth)(1)(Semantic Role)(none)(NP Form)(PRP)(Grammatical Role)(SBJ
.sw_0380_4572.du.attr line 17/245 8%

Summary: Tool-based annotation, reporting automatically

- Advantages:
 - reproducible
 - easy to go back and to correct mistakes
 - saves time and unnecessary work by preprocessing software
 - reporting automatically – allows detailed error analysis
- Disadvantages:
 - still a lot of work (until the annotator's wrist hurts)
 - difficult to get overview because view restricted to window on the screen (however, statistical analysis and error analysis may help)
 - because of non-standard data format difficult to access, convert, reuse,
...

Phase 4: XML-based annotation, standardized

- ca. 2001-2002
- anaphora resolution in written text, HTC

MMAX V0.6 C:\htc\002_htc.anno

File Tools Settings Help **Delete Member attributes**

Das Stadttheater .

Im Gegensatz zu anderen Städten steht **das Heidelberger Stadttheater** nicht an herausgehobener Stelle , sondern es fügt sich in die Straßenflucht ohne Vorplatz ein . Der Haupteingang zeigt **flachen Segmentbögen** . Erst 1874 über **ein** von einer privaten Initiative getragene Theater . Es wurde in der Folge stark verändert . Nach dem Innen-Umbau von 1880 durch Hermann Behagel gestaltete Fritz Haller 1924 **das Haus** erneut um . Trotz Aufstockung und Verbreiterung im Stil des Neoklassizismus wurden aber Teile der Straßenseite erhalten . 1990 wurde seitlich ein gläsernes Foyer von Rudolf Biste und Kurt Gerlin

es

type none anaphoric bridging

ante_sub_anaphoric none direct pronominal IS-A other

np_form none NE defNP indefNP PPER PPOS PDS

grammatical_role none SBJ OBJ other

agreement none 3M 3F 3N 3P 1S 2S 1P 2P

Apply Undo changes

MMAX V0.6 E:\htc\114_htc.anno

File Tools Settings Help Delete Member attributes

Überlegungen zur teilweisen Rekonstruktion des Heidelberger Schlossgarten in den siebziger Jahren .

In den Jahren 1976/77 wurden Überlegungen darüber angestellt , wie man den historischen Garten für den Besucher durch weitere Restaurierungs- und Rekonstruktionsarbeiten attraktiver gestalten könnte . Deshalb wurden verschiedene Untersuchungen durch den Schloßgartenbeirat vorgenommen . Das Ergebnis war die Erstellung von Leitlinien für eine gestalterische Behandlung des Hortus Palatinus unter besonderer Berücksichtigung kunstgeschichtlicher und denkmalpflegerischer Gesichtspunkte . Danach plante man , die Zwischenterrasse den historischen Unterlagen entsprechend wiederherzustellen . Ebenso dachte man an die Neugestaltung der kleinen östlichen Terrasse wo einst die Kabinette oder Gemächer

1 Das Ergebnis

type none anaphoric bridging

ante_sub_bridging none part-whole cause-effect entity-attribute other

np_form none ME defNP indefNP PPER PPOS PDS

grammatical_role none SBJ OBJ other

agreement none 3M 3F 3N 3P 1S 2S 1P 2P

Apply Undo changes

MMAX file structure

```
(0) 23> ls -al 002* coref_scheme.xml *.dtd *.xsl
-rwxr-xr-x    1 strube   strube           139 Mar 20 17:59 002_htc_abn.anno
-rwxr-xr-x    1 strube   strube          5888 Mar 20 18:01 002_htc_abn_marka
-rwxr-xr-x    1 strube   strube           564 Jun 23  2002 002_htc_text.xml
-rwxr-xr-x    1 strube   strube          3850 Jun 23  2002 002_htc_words.xml
-rw-rw-r--    1 strube   strube          3452 Mar 20 18:05 coref_scheme.xml
-rwxr-xr-x    1 strube   strube           242 Jun 23  2002 markables.dtd
-rwxr-xr-x    1 strube   strube           208 Jun 23  2002 text.dtd
-rwxr-xr-x    1 strube   strube          1314 Jun 23  2002 text.xsl
-rwxr-xr-x    1 strube   strube           166 Jun 23  2002 words.dtd
(0) 24>
```

Summary: XML-based annotation, standardized

- Advantages:
 - reproducible
 - easy to go back and to correct mistakes
 - saves time and unnecessary work by preprocessing software
 - reporting automatically – allows detailed error analysis
 - standoff annotation
 - allows use of suite of XML tools for processing
- Disadvantages:
 - still a lot of work (until the annotator's wrist hurts)
 - difficult to get overview because view restricted to window on the screen (however, statistical analysis and error analysis may help)
 - usually only one kind of annotation at one time (i.e. either coreference or dialogue acts, but not both together)

XML-based annotation, multi-level

- ca. 2003-
- anaphora resolution in spoken dialogue, Switchboard

Summary: XML-based annotation, multi-level

- Advantages:
 - arbitrary many levels of annotation on top of base-level annotations
 - maximizes use and possible reuse of annotations
 - allows to study interaction between many phenomena
- Disadvantages:
 - requires some planning
 - correcting base-level data may be difficult