EAGLES

Provisional Recommendations and Guidelines for the Syntactic Annotation of Corpora

DRAFT FINAL REPORT

EAGLES DOCUMENT EAG—TCWG—SASG/1.8

APPENDIX A

Contents

A Appendix: sample phrase structure annotations

The samples in this appendix illustrate the application of the annotation scheme proposed in this document to a number of European languages. The samples include the subcategorisations discussed in section 3.3.1 and also syntactic functions. For the sake of clarity, single node constituents have not been labelled (see section 4.4 for a discussion of the advantages and disadvantages of labelling single word constituents.)
It is hoped that these samples will have illustrative value, but they are not intended to be model implementations of the foregoing guidelines. They show rather minimal annotations, including the recommended categories of section 3.2, but few optional elements. For many purposes, a richer annotation scheme will be required.

A.1 Dutch

Source: *NRC Handelsblad*, Weekeditie, 23 May 1995

[S [NP Begin juni NP] [Aux worden Aux] [VP [PP in [NP het Scheveningse Kurhaus NP] PP] [NP de
start june are in the Scheveningen 'spa' the

Verenigde Naties NP-Subj] [AdvP weer AdvP] nagespeeld VP] . S]
United Nations again enacted

[S [AdvP Net als afgelopen jaren AdvP] [Aux zullen Aux] [NP-Subj studenten [PP uit [NP de hele
just as previous years will students from the whole

wereld NP] PP] NP-Subj] [VP deelnemen aan [NP-Obj zittingen [PP van [NP [NP de Algemene Vergadering
world participate in sessions of the General Assembly

NP] en [NP het Internationale Gerechtshof NP] NP] PP] NP-Obj] VP] . S]
and the International Court of Justice

[S [NP-Subj Ze NP-Subj] [Aux zullen Aux] [AdvP er AdvP] [VP [NP-Obj de VN-agenda [PP voor
They will there the UN agenda for

[NP het komende jaar NP] PP] NP-Obj] behandelen VP] . S]
the next year discuss

[S [NP-Subj The European International Model United Nations (Teimun) NP-Subj] , [CL zoals [NPhet
as the

evenementNP] heet CL] , [VP duurt [NP een kleine week NP] VP] . S]
event is called lasts a small week

[S [NP Het eindverslag [PP van Teimun NP] PP] [Aux wordt Aux] [PP door [NP de secretaris-generaal
The final report of Teimun is by the secretary general

Boutros-Ghali NP] PP] [VP [PP in [NP New York NP] PP] [PP aan [NP de echte Verenigde Naties NP] PP]
in to the real United Nations

gepresenteerd VP]. S]
presented

[S [NP-Subj Teimun NP-Subj] [VP is [NP een initiatief [PP van [NP de Studentenvereniging [PP voor [NP
is an initiative of the student association for

Internationale Betrekkingen NP] PP] (SIB) NP] PP], [NP een [PP van [NP de drie [AdjP grote internationaal
International Relations one of the three big internationally

georiënteerde AdjP] studentenverenigingen NP] PP] [PP in Nederland PP] NP] VP]. S]
orientated student associations in Holland

[S [NP-Subj De andere twee NP-Subj] [VP zijn VP] [NP [NP de European Law Students Association
the other two are the

(ELSA) NP] en [NP de Association [PP des [NP États Généraux [PP des [NP Étudiants [PP de l'Europe
and the Association of the States General of the Students of Europe

PP] PP] NP] PP] (AEGEE) NP] NP] . S]

[S [PP Behalve [NP de VN-in-het-klein [NP PP] [VP organiseert VP] [NP-Subj de SIB [PP [NP met
Besides the UN-in-the-small organizes the SIB with

bijna tweeduizend , [ADJP uitsluitend Nederlandse ADJP] leden , [CL-Vless onder wie veel
almost two thousand exclusively Dutch members among whom many

politicologie- en talenstudenten CL-Vless NP] PP] NP-Subj], [AdvP jaarlijks AdVP] [NP-Obj
political science and language students yearly

tientallen congressen en lezingen NP-Obj] . S]
tens of congresses and lectures

A.2 Danish

Source: *Berlingske Tidende*, 6 August 1995

[S [NP-Subj Trudjman NP-Subj] [VP spiller [NP-Obj hœjt spil NP-Obj] VP] . S]
plays high play

[S [NP-Subj Kroatiens præsident NP-Subj] [VP er [PP under [NP internt pres NP] PP] VP] [CL [PP [VP
Croatia's president is under internal pressure

for at generobre [NP-Obj Krajina NP-Obj] VP] CL] . S]
(for) to recapture Krajina

[S [AdvP Men AdvP] [NP-Subj offensiven NP-Subj] [VP kan give [NP-Obj bagslag NP-Obj] VP] .
But offensives can give repercussion

S]

[S [NP-Subj Der NP-Subj] [VP står [ADVP meget på spil ADVP] [PP for [NP præsident Franco Trud-
There stands much on play for President

jman NP] PP] [PP1 med [NP1 iværksættelsen [PP2 af [NP2 det omfattende indgreb [PP3 mod [NP3
with initiation of the extensive intervention against

de oprœrske serbere NP3] PP3] NP2] PP2] NP1] PP1] VP] . S]
the rebellious Serbs

[S [NP-Subj Han NP-Subj] [VP [Aux kan Aux] [VP risikere [NP-Obj en intervention [PP fra [NP det
He can risk an intervention from the

serbisk ledede Jugoslavien NP] PP] NP-Obj] VP] . S]
Serbian led Yugoslavia

[S [NP-Subj Flere faktorer NP-Subj] [VP lader til [CL at [VP [AUX have AUX] fået [NP-Obj Trudjman
Several factors seem to have made

NP-Obj] VP] til at slå til [ADVP nu ADVP] CL] VP] VP] . S]
(to) to beat to now

[S [NP-Subj Trudjman NP-Subj] [VP nærer [ADVP også ADVP] [NP-Obj stœrre tillid NP-Obj] [PP
places also greater confidence

til [NP det kroatiske militærs overlegenhed NP] VP] . S]
to the Croatioan army's superiority

[S [NP-Subj Hans hær NP-Subj] [VP [AUX har Aux] nået op [PP på [NP mindst 100.000 mand NP] PP]
His army has come up to at least men

VP] . S]

A.3 English

From: LOB Corpus, Fiction K

[S [NP-Subj He NP-Subj] [VP was [NP a tiny man NP] , [PP of [NP fanatical neatness NP] PP] VP], [CL-Vless [NP-Subj his remaining hair NP-Subj] [ADJP [ADJP snowy], and [ADJP cropped [PP like [NP a soldier's NP] PP] ADJP] ADJP] CL-Vless] . S]

[S [NP-Subj His white shirt cuffs NP-Subj] [VP were [ADVP actually ADVP] starched VP]: [NP-Subj he NP-Subj] [VP [VP protruded [PP from them PP] [NP-Obj his [ADJP surprisingly thick and hairy ADJP] wrists NP-Obj] VP] and [VP began [VP to wash VP] VP] VP] . S]

"[S [PP As a matter of fact PP] [NP-Subj a good hard frost NP-Subj] [VP seems [VP to clear [NP the old tubes NP] VP] VP] . S] [S[ADJP Much better [PP for me PP] [PP than [NP the rain NP] PP] ADJP] . S]"

[S "[ADJP Good ADJP]" [VP said VP] [NP-Subj Colmore NP-Subj] .S] [S "[ADJP Excellent ADJP]". S]

[S [NP-Subj Fairbanks NP-Subj] [VP hummed [NP-Obj a few bars NP-Obj] VP] [PP in [NP a voice [VP made resonant [PP by [NP the very weakness [PP of [NP his chest NP] PP] NP] PP] VP NP] PP] . S]

[S [S [NP-Subj Colmore NP-Subj] [VP was [ADJP ready [VP to leave VP] ADJP] VP] S], [S but [S [VP delayed [NP-Obj his departure NP-Obj] VP] CL], [ADVP as [NP one [CL-Rel [NP-Subj who NP-Subj] [VP dare not [VP go to bed early VP] VP] [PP for [NP fear [PP of [VP missing [NP some [ADJP [ADJP wholly unanticipated ADJP] but [ADJP remotely possible ADJP] ADJP] event [PP of [NP absorbing interest NP] PP] NP] VP] PP] NP] PP] VP] CL-Rel] NP] PP] S] . S]

[S [NP-Subj He NP-Subj] [VP took up [NP-Obj a clothes brush NP-Obj] VP] . S]

A.4 German

Source: *Der Spiegel*, 12 June 1995, p. 202.

[S [NP-Subj Sonnen NP-Subj] [Aux werden Aux] [VP [PP im Dunkeln PP] geboren VP] . S]
suns are in the dark born

[S [CL-Adv Abgeschirmt [PPvon [NP dichten Staubwolken NP] PP] CL-Adv], [VP reifen [NP-Subj die
screened off from dense clouds of dust mature the

Sternenembryos NP-Subj] [AdvP viele Jahrmillionen lang AdvP] heran VP] . S]
embryonic stars many millions of years long V-Prep

[S [NP-Subj Kein Lichtstrahl [PP von [NP ihrem Werden NP] PP] NP-Subj] [VP dringt [ADVP
no ray of light of their development emits

gewönlich ADVP] [PP nach außen PP] VP] . S]
normally to outside

[S [AdvP Doch AdvP] [NP-Obj dem erdumkreisenden Hubble-Weltraumteleskop NP-Obj] [Aux sind Aux]
yet the earth circling Hubble telescope are

[AdvP jetzt AdvP] [NP-Subj spektakuläre Schnappschüsse [VP [PP aus [NP dem Brutkasten [NP einer
now spectacular photographs from the ‘womb’ of a

entstehenden Sonne NP] NP] PP] NP-Subj] gelungen VP] . S]
developing sun succeeded

[S [NP-Subj Die [CL [VP [NP lezte Woche NP] veröffentlichten VP] CL] Aufnahmen NP-Subj] [VP zeigen
the last week released pictures show

[NP-Obj gewaltige , [CL [VP [NP Milliarden [PP von [NP Kilometern NP] PP] NP] [PP durchs All PP]
enormous billions of kilometers through galaxy

geschleuderten VP] CL] Gasfontänen , [CL-Rel [NP-Subj die NP-Subj] [VP aussehen [PP wie [NP der
hurled gas fountains which look like the

heiße Atem [NP eines Feuerschluckers NP] NP] PP] VP] CL-Rel] NP-Obj] VP] . S]
hot breath of a fire eater

[S [CL-Adv Wo [NP-Subj sie NP-Subj] [VP [PP mit [NP der umliegenden Materie NP] PP]
where they with the surrounding matter

zusammenprallen VP] CL-Adv], [VP scheint [NP-Subj der Weltraum NP-Subj] [VP zu brennen VP] VP]
collide seems the galaxy to burn

. S]

[S [VP Ausgestoßen [Aux hat Aux] [NP-Obj die Flammenzungen NP-Obj] VP] [NP-Subj eine [CL [VP [NP
expelled has the flame tongues a

rund 1500 Lichtjahre NP] [PP von [NP der Erde NP] PP] entfernte VP] CL] Nachwuchssonne, [CL-Rel
circa 1500 light years from the earth distanced new generation sun

[NP-Subj die NP-Subj] [VP vermutlich unmittelbar [PP vor [NP ihrer Zündung NP] PP] steht VP]
which presumably immediately before her ignition stands

CL-Rel] NP-Subj] . S]

[S [S [NP-Subj Sie selber NP-Subj] [Aux ist Aux] [VP [PP auf [NP den Hubble-Fotos NP] PP] nicht zu
she herself is on the Hubble pictures not to

sehen VP] S] , [S [NP-Subj sie NP-Subj] [VP verbirgt [NP-Obj sich NP-Obj] noch [PP hinter [NP ihrer
see she hides herself still behind her

Dunkelwolke NP] PP] VP] S] . S]
dark clouds

A.5 Italian

Source: PLNLP Italian corpus

[S [NP-Subj Le sollecitazioni [ADJP dovute [PP a [NP corti circuiti NP] PP] ADJP] NP-Subj]
the stresses due to short circuits

[VP costituiscono VP] [ADVP infatti ADVP] [NP-Obj un punto [ADJP fondamentale ADJP] [PP di [NP
represent in-fact a point fundamental of

indagine NP] PP] NP-Obj] .S]
investigation

[S [VP Viene approvato VP] [NP-Subj il verbale [PP della [NP seduta [ADJP precedente ADJP] NP] PP]
is approved the minutes of-the meeting previous

NP-Subj] . S]

[S [NP-Subj Tale progetto NP-Subj] [VP consentirá VP] [NP-Obj la interconnessione [PP di [NP centri
this project will-enable the interconnection of centres

[PP di [NP calcolo NP] PP] [ADJP (dotati [PP di [NP calcolatori [PP di [NP case costruttrici
of computing supplied with computers of manufacturers

diverse NP] PP] NP] PP] ADJP]) [CL-Rel [NP-Subj che NP-Subj] [VP operano VP] [PP in [NP ambito
various which operate in framework

[ADJP scientifico italiano ADJP] NP] PP] CL-Rel] NP] PP] NP-Obj] .S]
scientific Italian

[S [NP-Subj Il CNUCE NP-Subj] [VP lavora VP] [PP nel [NP campo [PP delle [NP reti [PP di [NP
the CNUCE works in-the field of-the networks of

elaboratori NP] PP] NP] PP] NP] PP] [PP fino dal [NP 1972 NP] PP] .S]
computers since

[S [PP Sulla [NP base [PP della [NP [ADJP nostra precedente ADJP] sollecitazione NP] PP] NP] PP] [VP
on-the basis of-the our previous solicitation

abbiamo ricevuto VP] [NP-Obj una ventina [PP di [NP contributi NP] PP] NP-Obj] .S]
we-have received a twenty of contributions

[S [NP-Subj La XXX NP-Subj] [VP ha VP] [ADVP recentemente ADVP] [VP annunciato VP] [CL-That
the has recently announced

che [VP adotterá VP] [NP-Obj la tecnologia Z-Ring NP-Obj] [PP per [NP i [ADJP suoi futuri ADJP]
that will-adopt the technology for the their future

prodotti [ADJP commerciali ADJP] NP] PP] CL-That] .S]
products commercial

A.6 Spanish

From: ITU Corpus

[S [NP-Subj Las caracteristicas [PP de [NP estos canales [PP para [NP gran velocidad [PP de [NP
The characteristics of these channels for high speed of

modulacion NP] PP] NP] PP] NP] PP] NP-Subj] [VP se definen [PP en [NP las recomendaciones [NP
modulation are defined in the recommendations

R.37 ([NP canales [PP de [NP 100 baudios [PP con [NP 240 Hz [PP de [NP separacion NP] PP] NP] PP] NP]
channels of baud with of separation

PP] NP]) NP] y [NP R.38 ([NP canales [PP de [NP 200 baudios NP] y [NP 360 Hz [PP de [NP separacion
and channels of baud and of separation

NP] PP] NP] PP] NP]) NP] NP] NP] VP] . S]

[S [NP-Subj Los canales [PP de [NP 200 baudios/360 Hz NP] PP] NP-Subj] [VP se instalaran [ADVP
The channels of are installed

segun [NP el orden [PP de [NP frecuencia siguiente NP] PP] NP] ADVP] VP] : [NP 5,4,6,3,2,1 NP], [PP
according to the order of frequency following

en [NP sustitucion [PP de [NP los canales correspondientes [PP de [NP 50 baudios NP] PP] NP] PP] NP] PP]
in substitution of the channels corresponding of baud

. S]

[S [NP-Subj Los numeros [PP de [NP canales NP] PP] NP-Subj] [VP se ajustan [PP a [NP la
The numbers of channels are adjusted to the

Recomendacion R.38 NP] PP] VP] . S]
recommendation

[S [PP En [NP el caso [PP de [NP un sistema mixto [CL-Rel [PP en [NP el [CL que [VP se utilicen [NP
In the case of a system mixed in the which are used

canales [PP con [NP tres velocidades [PP de [NP modulacion NP] PP] diferentes NP] PP] NP] VP] CL] NP]
channels with three speeds of modulation different

PP] CL-Rel] NP] PP] NP] PP], [NP-Subj el orden [ADJP indicado [PP en [NP el sector 3.2 NP] PP] ADJP]
the order indicated in the sector

NP-Subj] [VP tiene [NP prioridad [PP con [NP respecto [PP al [NP [ADJP indicado [PP en [NP el sector
has priority with respect to the indicated in the sector

2.1 NP] PP] ADJP] NP] PP] NP] PP] NP] VP] . S]

[S [PP Para [NP la combinacion [PP de [NP canales [PP de [NP 240 Hz [PP de [NP separacion [PP con
For the combination of channels of of separation with

[NP canales [PP de [NP 480 Hz [PP de [NP separacion NP] PP] NP] PP] NP] PP] NP] PP] NP] PP] NP] PP]
channels of of separation

NP] PP] , [NP-Subj los canales [PP de [NP 480 Hz [PP de [NP separacion NP] PP] NP] PP] NP-Subj] [VP
the channels of of separation

se instalaran [ADVP segun [NP el orden [PP de [NP preferencia siguiente NP] PP] NP] ADVP] VP] :
are installed according to the order of preference following

[NP 4,3,5,2,6 NP] . S]

A.7 Portuguese

Source: CRPC-Corpus

[S [NP-Subj O clima mundial NP-Subj] [VP tem vindo a evoluir [ADVP lentamente ADVP] VP] [PP
the climate global has been evolving slowly

no [NP decurso [PP de [NP milénios NP] PP] , [PP de [NP séculos NP] PP] e [PP de [NP decénios NP]
in the course of millenia of centuries and of decades

PP] NP] PP] e [CL [VP continuará a modificar-se VP] [PP no [NP futuro NP] PP] . S]
and will continue to be modified in the future

[S [ADVP Ora ADVP] , [CL-Cond se [NP-Subj o homem NP-Subj] [VP pode tirar proveito [PP das
now if the man may take advantage of the

[NP condições climáticas favoráveis NP] PP] VP] CL-Cond] , [VP é [ADVP também ADVP] vulnerável [PP
conditions climatic favourable is also vulnarable

às [NP mudanas NP] e [PP às [NP variações NP] PP] [PP do [NP clima NP] PP] PP] VP] . S]
to the alterations and to the variations of the climate

[S [NP-Subj [NP A alimentação NP] , [NP a aágua NP] , [NP a energia NP] , [NP o habitat NP]
the food the water the energy the environment

e [NP a saúde NP] NP-Subj] [VP são [NP outros tantos aspectos [PP da [NP vida humana NP] PP]
and the wealth are other many aspects of the life human

[CL-Rel que [VP dependem [ADVP fortemente ADVP] [PP do [NP clima NP] PP] VP] CL-Rel] VP] .
which depend strongly on the climate

S]

[S [ADVP Ora ADVP] , [ADVP sem querer ADVP] , [NP-Subj o homem moderno NP-Subj] [VP modifica
now without willing the man modern modifies

[NP o clima NP] VP] [PP [PP à [NP escala local NP] PP] e , [ADVP em menor medida ADVP] , à
the climate at the scale local and to smaller extent at the

[NP escala regional NP] PP] . S]
scale regional

[S [NP-Subj A investigação NP-Subj] [VP permitiu determinar [NP um grande número [PP das [NP
the research allowed to determine a large number of the

características fundamentais [PP das [NP mudanças climáticas [ADJP observadas [PP no [NP passado
characteristics fundamental of the alterations climatic observed in the past

NP] PP] ADJP] NP] PP] NP] PP] NP] VP] , e [VP permite [ADVP tambén ADVP] fazer [NP projecções [PP
and allows also to make predictions

para [NP o clima futuro NP] PP] NP] VP] . S]
for the climate future