

Annotationsrichtlinien/Entitäten v1.1

Seitenübersicht

1. Grundsätzliches
 1. Maximale Nominalphrasen
 2. Teil- vs. Vollannotation
 3. Verschachtelungen
 4. Status der Entitäten
 5. Gruppenbildungen
 6. Zeitliche Veränderungen
 7. Kompositionen
 8. Adjektivierungen
2. Klassifizierung
 1. Personen/Figuren (PER)
 2. Orte (LOC)
 3. Organisationen (ORG)
 4. Ereignisse (EVT)
 5. Werke (WRK)
 6. Abstrakte Konzepte (CNC)
3. Einzelfälle
 1. Umgang mit ‘Gott’
 4. Workflow beim Annotieren allgemein
 5. Offene Fragen / Baustellen / schwierige Fälle

Link zur Annotationsumbegung CRETAAnno: <https://clarin09.ims.uni-stuttgart.de/cretanno/>

Version 1.1. Dies ist ein erster überarbeiteter Entwurf für gemeinsame Annotationsrichtlinien für CRETA-Entitäten. Einfügungen gegenüber der Version 1 sind rosa gefärbt.

Beispielsätze sind durchnummeriert, die zu annotierenden Textstellen sind mit eckigen Klammern [] markiert.

NR: Anmerkungen, Kommentare und Rückfragen zu den Richtlinien bitte entweder an der entsprechenden Stelle oder am Ende unter *Offene Punkte / Baustellen* einfügen. In jedem Fall bitte als AutorIn zu erkennen geben (Initialen sind CRETA-eindeutig).

Unter Entitäten verstehen wir spezifische, durch Benennung unterscheidbare Objekte in einer echten oder fiktiven Welt. Die Objekte können jeweils einer Kategorie aus einer groben Kategorieneinteilung zugeordnet werden. In Texten annotieren wir Referenzen auf solche Entitäten, also Bezeichnungen, die jeweils auf eine einzelne Entität oder eine Menge von Entitäten hinweisen. Einige der Kriterien und Beschreibungen in den Richtlinien beziehen sich auf “echtweltliche” Eigenschaften, andere auf die sprachliche, linguistische Form der Referenz. Im Kontext sollte jeweils klar sein, was gemeint ist.

Im Kontinuum zwischen Eigennamen (Named Entites à la Computerlinguistik) und

Koreferenzresolution bewegen wir uns eher auf der Seite der Koreferenzresolution, wobei wir uns erstmal auf die Referenzausdrücke konzentrieren. D.h. in diesem Schritt werden Textstellen annotiert, die auf eine Entität referieren. Welche das ist, oder ob zwei Stellen auf die gleiche Entität referieren, blenden wir erstmal aus.

1 Grundsätzliches

Wir verwenden die Bezeichnung “Referenzausdruck” für alle Textstellen, die wie oben beschrieben referieren.

Referenzausdrücke können sein:

Eigennamen (“Named Entities”, “Gahmuret von Anshevîn”, “Warschauer Pakt”, “Der Teppich” [als Werktitel]) referieren typischerweise auf eine (im Textzusammenhang) eindeutig identifizierbare Entität.

Gattungsnamen (Appellativa, “Haus”, “Tier”, “Bauer”) bezeichnen Gattungen. Eine Gattung ist ein Menge von Entitäten mit gleichen Eigenschaften (im Ggs. zu einer einzelnen Entität dieser Menge), sind also für unsere Annotationsaufgabe in dieser Verwendung irrelevant. Gattungsnamen werden aber auch verwendet, um jeweils auf eine bestimmte Instanz einer solchen Gattung (also eine Entität) zu referieren (“Das ist aber ein niedlicher Hund!”). Oft werden Gattungsnamen auch als Gruppenreferenzen verwendet (“drei Bauern”) oder in Kombination mit Eigennamen (“der Engel Lotte”). Wir annotieren vor allem Gattungsnamen, wenn sie sich im konkreten Textzusammenhang auf eine oder mehrere Instanz(en) einer Gattung beziehen. Die CNC-Kategorie beinhaltet *auch* Gattungsnamen, die auf Klassen/Gattungen referieren.

- (1) Besonders langsam war wie immer [der Esel Benjamin]. Er ließ [das Tier] auf dem Feld frei. (Entitäten)
- (2) Der Löwe lebt im Unterschied zu anderen Katzen in Rudeln. (keine Entität)

Pronomen (er, sie, es) können auf viele verschiedenen Dinge referieren. Da es sich um eine (ab-)geschlossene Wortklasse handelt, die wir gut automatisch im Text finden können, lassen wir sie bei der manuellen Annotation weg. Pronomen werden nicht annotiert.

Es sind also nur die Eigen- und Gattungsnamen zu annotieren, die auf eine Entität referieren.

1.1 Maximale Nominalphrasen

Referenzausdrücke sind (in diesen Richtlinien) maximale Nominalphrasen (NPs), also Substantive mit sich davor/-nach anschließenden, das Substantiv weiter spezifizierenden Textteilen.

Nominalphrasen beinhalten z.B.

bestimmte und unbestimmte Artikel

- (3) Er war [ein Tor].
- (4) Er war [der König].

Komplemente und Adjunkte

- (5) [Die Burg oberhalb [des Flusses]] sah majestätisch aus.
- (6) [Das Kind [des Bauern]] wurde auf den Namen Peter getauft.
- (7) [Die Ermordung [Cäsars]]_{EVT} war von langer Hand geplant.
- (8) [[Obelix]' Hund]_{PER} ist klein und weiß.

Attribute

- (9) Mit [diesem Schwert hier] kann der Drache erschlagen werden.
- (10) [Die magische Wand] widerstand den Versuchen, sie einzubrechen.

Relativsätze

Relativsätze werden, sofern sie Teil der Nominalphrase sind, mit annotiert. Dadurch werden die Annotationen mitunter sehr lang. Abgetrennte Relativsätze wie in (12) werden nicht mit annotiert¹.

- (11) [Die Magd, die die meiste Verantwortung innehatte], war Anna.
- (12) Sie hatte [ein Haus] gebaut, das grüne Fensterläden hatte.

Appositionen

- (13) [Meine Schwester, Clara Reuss], liebt [einen Mann].
- (14) [Das Werk "Schlafes Bruder"] beschreibt ...
- (15) [Mein Nachbar, ein Arzt namens Doc Brown], empfahl ...

Nicht Teil der Nominalphrase sind vorangestellte Präpositionen (16), außer es handelt sich um zusammengesetzte Formen (17). Bei **zusammengesetzten Formen** (am, im etc.) wird die Form mit in den Referenzausdruck aufgenommen.

- (16) Auf [der Bank] sitzt ein Vogel.
- (17) [Am Hamburger Rathaus] leben viele Tauben.

Substantivierte Adjektive ... werden als Referenzausdruck (zusammen mit dem Rest ihrer NP) annotiert. Die Kategorie hängt davon ab, auf wen/was die NP referiert.

- (18) [Die Schöne]_{PER} warf ihr Haar herab.

¹Über diese Frage haben wir lange diskutiert. Für eine Aufnahme von Relativsätzen spricht a) Kompatibilität mit existierenden Koreferenz-Annotationen (TüBa-D/Z, NoSta-D), b) mehr semantischer Gehalt der annotierten Textstellen und c) dass es sich um eine syntaktische Konstituente handelt. Dagegen spricht, dass a) es die Annotationen aufwändiger und langsamer macht, b) es die Frage nach nicht-kontinuierlichen Annotationen aufwirft und c) wir ja auch noch gar nicht wissen, wie wir den Relativsatz dann benutzen. Angesichts der Überlegung, dass lange Relativsätze dann vielleicht doch nicht so häufig sind, haben wir uns für die Variante entschieden, die konzeptuell am klarsten scheint (syntaktische Konstituente).

(19) Sie ging in [den Wald]_{LOC}. [Der Dunkle]_{LOC} kam ihr bedrohlich vor.

1.2 Teil- vs. Vollannotation

Generell streben wir Vollannotationen an, d.h. jede Referenz auf eine Entität der u.g. Klassen wird annotiert, unabhängig von ihrem Status, ihrer Wichtigkeit oder interpretatorischer Relevanz. Für die Kategorien Ereignisse und abstrakte Konzepte birgt eine Vollannotation die Gefahr, dass hinterher jede NP des Textes annotiert ist, was uns ja auch nicht weiterhilft. Für diese beiden Kategorien sollten also die Ereignisse (oder Klassen von Ereignissen) und Konzepte (oder Klassen von Konzepten) vor der Annotation definiert werden, die zu annotieren sind. Maßstab dafür sollte sein, welche Entitäten für eine Interpretation/Analyse des Textes relevant sein werden oder sein können. Die Häufigkeit, mit der auf ein Ereignis oder ein Konzept Bezug genommen wird, könnte dafür ein Indikator sein. Die Definition der abstrakten Konzepte kann und sollte natürlich Beispiele enthalten, um sie den AnnotatorInnen verständlich zu machen. Es ist aber dann die Aufgabe der AnnotatorInnen, mögliche Verbalisierungen der Konzepte im Text zu identifizieren².

1.3 Verschachtelungen

Referenzausdrücke können verschachtelt auftreten, werden aber nur getrennt annotiert, wenn sie sich auf unterschiedliche Entitäten beziehen.

In (20) und (21) beziehen sich die gesamte NP und der eingebettete Referenzausdruck auf unterschiedliche Entitäten, daher werden beide annotiert. In (22) sind beide Bestandteile referent auf die gleiche Entität³ und werden daher nicht separat annotiert.

(20) [Der EU-Gipfel in [Spanien]] war ein voller Erfolg.

(21) [Gahmuret von [Anshevin]] trank ein Bier.

(22) [Bundeskanzler Schröder] hat “Basta” gesagt.

(23) [Die Kollegen [Merkel] und [Schröder]] ...

1.4 Status der Entitäten

Wir unterscheiden bei der Annotation nicht, in welchem Textabschnitt eine Entität erwähnt wird (Figurenrede, Protokollanmerkung, Diegese). Da sich alle Annotationen

²Ob für die weitere Verarbeitung von Teilannotationen ein maschinelles Lernverfahren dann das Richtige ist, sei im Moment dahingestellt. Denkbar wäre, aus den Annotationen Wortlisten zu generieren, die man ggf. um weitere Synonyme (automatisch) ergänzen lassen kann. Dahinter steckt die Frage, ob es denn – vor allem bei den teilannotierten Kategorien – überhaupt möglich ist, linguistische Kontextindikatoren zu finden, die bei der Erkennung helfen.

³Das gilt nur zu dem Zeitpunkt, in dem sich die Welt zum Zeitpunkt der Äußerung befindet. Um die Problematik, dass sich die von “Bundeskanzler” referierte Entität gelegentlich ändert, kümmern wir uns später.

auf Textstellen beziehen, können wir eine (auch spätere) Analyse von Textabschnitten leicht mit der Entitätenannotation übereinanderlegen. Die Unterscheidung nach “Status” der Entität kann dann davon abgeleitet werden.

1.5 Gruppenbildungen

Wenn Entitäten in Gruppen auftreten, werden sie als Referenzausdruck annotiert (24). Falls ein Zahlwort vorhanden ist, gehört das ebenfalls zur Nominalphrase (25). Kollektive werden als Gruppenreferenzen annotiert (26), wenn sie nicht generisch gemeint sind (27)⁴.

(24) [Die CDU-geführten Bundesländer] haben den Gesetzesentwurf blockiert.

(25) [Sieben Zwerge] kamen einst zurück in ihr Haus.

(26) [Die Deutschen]_{PER} fahren gerne in den Süden.

(27) Der Elefant ist das größte lebende Landsäugetier. (keine Entität)

1.6 Zeitliche Veränderungen

Einige Entitäten unterliegen zeitlichen Veränderungen. Maßgeblich ist dabei der Äußerungszeitpunkt, d.h. der Zeitpunkt, zu dem ein Satz/Text geschrieben oder gesagt wurde.

(28) [Die Stadt Rumbek] liegt [im Südsudan].

Südsudan in diesem Beispiel würde sich bei einer Äußerung vor 2010 auf eine Region beziehen, nach ca. 2011 jedoch auf einen Staat⁵. Für die hier beschriebenen Annotationen ist das nur für die Kategorisierung wichtig – in jedem Falle wäre es ein Referenzausdruck.

1.7 Kompositionen

Sind Referenzausdrücke Teil eines zusammengesetzten Nomens (“Koran” und “Vers” in 22), und stehen die Teile in einem semantischen Verhältnis, wird auch die NP mit zusammengesetztem Nomen annotiert.

(29) ... machte [die Koranverse]_{WRK} dafür verantwortlich.

⁴Ob definite Gattungsnomen im Singular als generisch oder kollektiv verstanden werden, ist vielleicht in Fällen schwer zu entscheiden. Zu analysieren ist hier vorerst, inwieweit das doch durch den Kontext bestimmbar ist.

⁵Natürlich ist es auch nach 2011 möglich, sich mit “Südsudan” auf eine Region zu beziehen, aber es ist anzunehmen, dass Sprecher das dann extra kenntlich machen.

1.8 Adjektivierungen

Adjektivierungen eines Nomens sind nur als Referenzausdruck zu behandeln, wenn sie zusammen mit einem Nomen einen Referenzausdruck darstellen (24, 25).

- (30) Die europäische Zustimmung zum Abkommen ließ auf sich warten. (‘europäisch’ wird nicht als Referenz auf Europa annotiert)
- (31) [Die Afrikanische Union]_{ORG} beriet sich in Kairo.
- (32) [Die europäischen Werte]_{CNC} befinden sich in der Krise.

In (31) wird auf eine Organisation referiert (mit der ganzen NP), und in (32) auf ein abstraktes Konzept⁶.

2 Klassifizierung

Ist ein Referenzausdruck im Text entdeckt, wird er einer der im folgenden erklärten semantischen Klassen zugewiesen. Referenzausdrücke außerhalb dieser Klassen lassen wir un-annotiert. Die Zuweisung eines Referenzausdrucks zu einer Klasse ist nicht immer eindeutig. In Fällen systematischer Ambiguität (z.B. Staaten) versuchen wir, die Klasse zu identifizieren, die im Kontext überwiegt/vorherrscht. Fälle, in denen das nicht feststellbar ist, werden als klassenambig gekennzeichnet.

Metonymie In Fällen metonymer Relationen wird die Kategorie der *gemeinten* Entität annotiert. In (33) wäre das Person, in (34) ein Werk, in (35) die Organisation Bundesregierung⁷.

- (33) Ich hole mir [den Shakespeare]_{PER} in die Kneipe.
- (34) Ich hole mir [den Shakespeare]_{WRK} aus dem Schrank.
- (35) [Die Deutschen]_{ORG} zwingen Griechenland einen Sparkurs auf.

Klassenambiguität

Wir nehmen an, dass in vielen Fällen eine zumindest überwiegende Zuordnung zu einer Klasse unter Berücksichtigung des Kontextes möglich ist. Die Fälle, in denen das nicht entscheidbar ist, markieren wir zunächst als ambig und setzen sie auf eine “Später-Nochmal-Anschauen”-Liste⁸. Eine Analyse und Diskussion der gesammelten Beispiele en bloc könnte uns bei einer Zuordnung helfen und erlaubt die Identifikation systematischer Ambiguitäten.

⁶Wenn europäische Zustimmung auch ein solch interessantes Konzept wäre, würde in (30) natürlich auch annotiert.

⁷Ob ein Sprecher in (35) tatsächlich nur die Bundesregierung *meint*, wird sich in vielen Fällen nicht feststellen lassen. Dieses Problem sollten wir im Auge behalten.

⁸Im Annotationswerkzeug wird es eine Markierungsfunktion dafür geben und natürlich die Möglichkeit, dass AnnotatorInnen freie Kommentare/Notizen speichern.

Beispiele: - ORG vs. LOC: Ein zu erwartendes Problem ist die Abgrenzung von Organisationen (Europa als politische Entität) und Ortsbezeichnungen (Europa als Kontinent). - EVT vs. LOC: Betrachtet man lediglich die einzelne Entität, ist eine Ambiguität zwischen Ereignis und Ort zu erwarten. Ähnlich wie bei ORG vs. LOC würden wir aber erwarten, dass der Kontext die Ambiguität in den meisten Fällen auflöst. *Srebrenica* in (51) ist klar eine Referenz auf ein Ereignis, das in Srebrenica stattfand. - PER vs. CNC: *Der Unglückliche* in einem möglichen unentscheidbaren Textzusammenhang, wenn das abstrakte Konzept für Krankheit für die Annotation festgelegt ist. - LOC vs. CNC: *Der Wald* als Referenz auf das Konzept Furcht, falls dieses Konzept für die Annotation festgelegt wurde.

2.1 Personen/Figuren (PER)

Referenzausdrücke, die sich auf Personen, Figuren oder figurenähnliche Entitäten beziehen (z.B. Tiere). Gruppen werden normal mitannotiert.

- (36) [Frau Wedemeier]_{PER} hat heute ein Bankkonto eröffnet.
- (37) Leb Wohl, [mein Engel]_{PER}, leb Wohl, [Lotte]_{PER}!
- (38) So hat [George]_{PER} in dem Gedicht "Der Teppich" ...
- (39) [Drei Ritter]_{PER} gingen über die Brücke. [Parzival]_{PER} kam hinzu. [Zwei Ritter]_{PER} fielen in den Fluss.
- (40) [Emilia Galotti]_{PER} bittet [ihren Vater]_{PER} flehentlich, ...

Für die Annotation des Referenzausdruckes ist es unerheblich, auf wen genau sich ein Ausdruck bezieht (39). Referenzen auf fiktive Personen in anderen Werken werden in der gleichen Weise annotiert (40), das gleiche gilt bei toten Personen.

2.2 Orte (LOC)

Ausdrücke, die sich auf Städte, Länder oder Gebiete beziehen.

- (41) In [Paris]_{LOC} steht der [Eiffelturm]_{LOC}.
- (42) Als Hannibal über [die Alpen]_{LOC} zog, ...
- (43) Schornsteinfeger aus [dem EU-Ausland]_{LOC} können ab sofort hier tätig werden.

Auch nicht namentlich bekannte Orte können Ortsreferenzen sein:

- (44) [Im Universitätsgebäude]_{LOC} ist es im Sommer warm und im Winter kalt.
- (45) Drei Ritter gingen über [die Brücke]_{LOC}.

2.3 Organisationen (ORG)

Ausdrücke, die Organisationen bezeichnen.

- (46) werden wir die Forderungen [der EU]_{ORG} im Auge haben müssen

- (47) [Die NATO]_{ORG} hat [zwölf Kompanien]_{ORG} in Marsch gesetzt. [Das Bündnis]_{ORG} verfolgt damit das Ziel ...
- (48) [Die NATO]_{ORG} reagiert damit auf Truppenbewegungen [des Warschauer Pakts]_{ORG}.

2.4 Ereignisse (EVT)

Referenzen auf Ereignisse, die für eine Textanalyse relevant sind.

- (49) [Der 11. September]_{EVT} hat alles verändert.
- (50) [Der letzte Gipfel]_{EVT} liegt bereits knapp sieben Jahre zurück.
- (51) [Srebrenica]_{EVT} hat mein Leben zerstört.

2.5 Werke (WRK)

Referenzen auf kulturelle Artefakte ⁹ in einem relativ weiten Sinn.

- (52) So hat George in [dem Gedicht “Der Teppich”]_{WRK} es gelehrt.
- (53) [Die Römischen Verträge]_{WRK} wurden 1957 in Rom unterzeichnet.
- (54) [Die Goldene Bulle]_{WRK} ist das wichtigste Verfassungsdokument des mittelalterlichen Reiches.
- (55) wenn man aus [der Bibel]_{WRK} vorliest und behauptet, es sei [der Koran]_{WRK}
- (56) [Emilia Galotti]_{WRK} ist ein bürgerliches Trauerspiel

2.6 Abstrakte Konzepte (CNC)

Referenzen auf Konzepte, die für die Analyse wichtig sind. Das sind oft Fachkonzepte oder solche Konzepte, die unter Diskussion stehen. Referenzen in dieser Kategorie können sowohl auf Klassen als auch auf Instanzen referieren¹⁰.

- (57) Lebendig ist [ästhetische Erfahrung]_{CNC} vom Objekt her, in dem [die Kunstwerke]_{CNC} unter ihrem Blick selbst lebendig werden.
- (58) ... wir nennen das [eine Krankheit zum Tode]_{CNC}.

⁹Und damit verlassen wir endgültig das etablierte Terrain.

¹⁰Erwogen haben wir hier, eine feinere Unterteilung in “Referenz auf Klassen” und “Referenz auf Instanzen” vorzunehmen. Wir haben das jedoch verworfen, weil sich gerade bei Abstrakta (“Macht”, “Islam”) oft gar nicht klar sagen lässt, ob es eine Klasse oder Instanz wäre bzw. sich gute Gründe für beide Lesarten finden lassen. Für die Forschungsfragen, die wir untersuchen wollen, scheint uns der Unterschied auch gar nicht so wesentlich.

3 Einzelfälle

3.1 Umgang mit ‘Gott’

Wir unterscheiden verschiedene Arten der Erwähnung von “Gott”. Die Verwendung in idiomatischem Kontext (59) wird *nicht* als Referenzausdruck annotiert. Hat Gott in einer fiktiven Welt einen figurenähnlichen Status (wie z.B. die griechischen Götter bei Homer oder in 50), werden sie als PER-Referenz klassifiziert. Je nach Fragestellung kommt Gott auch als abstraktes Konzept vor (61).¹¹

(59) “Oh Gott, das ist eine Überraschung!” (keine Entität)

(60) Sie wartete auf ein Zeichen [des Herrn]_{PER}. Doch antwortete er nicht.

(61) Ein Gefühl wie [Gott]_{CNC}über allen Dingen.

4 Workflow beim Annotieren allgemein

Große Menge eindeutiger Annotationen erzeugen Grundsätzlich sollten *vorerst* vor allem viele der entscheidbaren, eindeutigen Textstellen annotiert werden, damit Daten für das Maschinelle Lernen bereitgestellt werden können (das dazu benutzt werden soll, Vorschläge beim Annotieren zu machen und die Annotationsarbeit zu erleichtern). Fälle, die schwer zu entscheiden sind, werden zuerst nur zur Wiedervorlage markiert und gesammelt, um zu aufzudecken, ob es Einzelfälle sind, oder ob ähnliche Fälle öfter auftreten, so dass es sich lohnt, bald Zeit in deren Klärung zu investieren. Durch diesen Ablauf gewinnen die Annotierer auch schneller einen Überblick über und ein Gefühl für die Eigenheiten der Texte im Hinblick auf die Annotationsaufgabe.

Zweifelsfälle und Entscheidungen dokumentieren Fälle, die nicht eindeutig von den Hilfskräften entscheidbar sind, sollten (mit der “Wiedervorlage”-Kategorie markiert und) gesammelt und in *regelmäßigen Sitzungen* den wissenschaftlichen Mitarbeitern vorgestellt werden. Können die wissenschaftlichen Mitarbeiter mit wenig Zeitaufwand darüber entscheiden, soll diese Entscheidung zur Orientierung in die Richtlinien im Wiki eingetragen werden. Immer noch zweifelhafte Fälle werden auch ins Wiki eingetragen (lieber erst die Zeit für mehr eindeutige Annotationen und den Überblick verwenden), damit die anderen Fachgruppen, die Fachgruppenleiter und die Projektleitung einen Überblick bekommen und ggf. gesammelt darüber beraten werden kann.

Konsistenz Die konsistente Annotation der schon entschiedenen Fälle ist wichtig (auch, wenn zwischendurch die Entscheidung angezweifelt wird), um später bspw. eine durch die Anders-Entscheidung notwendig gewordene Änderung der Kategorie systematisch vornehmen zu können (was dadurch hoffentlich auch halb-automatisch geschehen kann).

¹¹Wichtig ist für uns ja hier nicht der theologische Status von Gott, sondern die Tatsache, dass in verschiedener Weise auf Gott referiert werden kann. Wir hoffen damit eine pragmatische Vorgehensweise gefunden zu haben, die für die verschiedenen Textsorten funktioniert.

5 Offene Fragen / Baustellen / schwierige Fälle

(... bitte nach Fachgruppen aufgeteilt)

Zusatzrichtlinien Plenardebatten Stand 28.07.2016

1. Entscheidung LOC vs. ORG

1.1 Städte, Länder und Regionen als Organisationen (ORG)

Wenn mit der Äußerung von Entitäten wie z.B. Berlin, Europa oder der Westen eindeutig politische Institutionen (z.B. Bundesregierungen, internationale Organisationen usw.) gemeint sind, dann werden sie als ORG annotiert. Hinweise dafür sind z.B.:

Modalverben, also solche Verben die eine Notwendigkeit oder Möglichkeit ausdrücken (dürfen, können, mögen, müssen, sollen, wollen) - „Die Türkei muss endlich die gefangen genommenen Journalisten freilassen.“ - „Europa darf hier nicht klein beigeben.“ - „Brüssel kann unsere Armee nicht herunkommandieren.“ - „Der Westen sollte seinen militärischen Verpflichtungen nachkommen.“

Tätigkeitsverben, also solche Verben, die ein allgemeines Tun bzw. eine zweckgerichtete Handlung beschrieben (wie z.B. entscheiden, machen, bringen, fahren) - „Auf dem Christenverfolgungsindex 2011 rangiert die Türkei auf Platz 30, noch vor Weißrussland und dem Sudan.“ - „Teheran weigert sich, die inhaftierten Demonstranten auszuliefern.“

Die gilt analog für substantivierte Tätigkeitsverben - „Die Entscheidung Brüssels zur Aufnahme von Beitrittsverhandlungen zwischen der Türkei und Europa ist von historischer Bedeutung“ - „Die Haltung in Brüssel gegenüber Moskau halte ich für hoch problematisch.“

Eindeutige Bezüge auf die politische Dimension einer Entität (in diesem Fall Europa als EU bzw. der Westen als NATO): - Wir haben ein gutes, freundschaftliches und enges Verhältnis mit der Türkei in den letzten Jahrzehnten. Die Türkei ist ein assoziiertes Mitglied der Europäischen Union. - Sie werden die Aufnahme von Beitrittsverhandlungen mit der Türkei beschließen und Sie werden unseren Vorschlag eines dritten Weges, einer privilegierten Partnerschaft, ausschlagen. - „Wir werden versuchen, anhand konkreter Stadien der Verhandlungen in Europa diese Grundsätze durchzusetzen.“ - „Das reicht nicht aus, im Gegenteil, das gefährdet die anstehende notwendige Integration Europas und gerade das Nahebringen dieser Entwicklung in der Innen- und Justizpolitik in Europa.“ - „Die Einführung des Euro war ein außerordentlich wichtiger und markanter Punkt in der gemeinschaftlichen Politik in Europa.“ - „Es stellt sich eine ganz andere Frage: Welche Kraft auf sicherheitspolitischem Gebiet hat der Westen und insbesondere die Europäische Union?“ - Darin wurden soziale Mindeststandards für Europa

1.2 Städte, Länder und Regionen als Orte (LOC)

Lokale Präpositionen (an, auf, hinter, in, neben, über, unter, vor, zwischen), die im Zusammenhang mit Länder- oder Regionen-Namen genannt werden, sind häufig ein

Hinweis dafür, dass es sich um eine Ortsangabe (=LOC) handelt. - „Erst vor wenigen Wochen wurden Journalisten in der Türkei festgenommen wegen angeblicher Mitgliedschaft in einer Organisation Ergenekon, die sich gegen Ministerpräsident Erdogan wendet.“ - „Noch immer können Kirchen und christliche Glaubensgemeinschaften kein Eigentum erwerben, noch immer ist die Ausbildung für Priester und Ordensleute in der Türkei nicht möglich.“ - „Die Enteignungen beim 1 600 Jahre alten Kloster Mor Gabriel sind ein unglaubliches Zeugnis dafür, dass die Religionsfreiheit, insbesondere die Religionsfreiheit der Christen, in der Türkei geradezu mit Füßen getreten wird.“ - „Wenn wir Bewusstsein für Grundrechte schaffen wollen, dann müssen wir dafür begeistern, dann müssen die Bürger in Deutschland erkennen, was wir damit verbinden und wie wir Europa in dieser Hinsicht als Wertegemeinschaft gestalten wollen. [Europa ist hier als ORG zu annotieren, denn es geht um die EU]“ - „Wenn es um ‘Flexicurity’ geht, wird gerne auf den geringen Kündigungsschutz in Dänemark“ - „Die Lösung, die die Bundeskanzlerin Angela Merkel auch in der Türkei sehr offen vertreten hat, ist daher richtig.“ - „Soll das etwa eine Situation wie in Deutschland sein: hohe Arbeitslosigkeit bei großen Exportüberschüssen?“ - Ich hoffe, dass die im internationalen Vergleich immer noch viel zu hohe Zahl von 17 Millionen Arbeitslosen in Europa in der Debatte über den Stabilitätspakt und über die Koordination der Europäischen Zentralbank etwas bewirkt. - Doch sehr ernst scheinen es Bundesregierung und EU-Kommission nicht zu meinen, wenn es um die soziale Gestaltung der Globalisierung geht, und dies trotz der massiven Zunahme der Armut in Europa. - Deshalb haben wir als Opposition beantragt, darüber zu debattieren, weil wir aus vielen Umfragen und Gesprächen wissen, wie sehr dieses Thema die Gemüter in Deutschland

ACHTUNG! Nicht immer wird durch lokale Präpositionen eine LOC-Annotation ausgelöst. Wenn mit der geäußerten Entität eindeutig eine politische Organisation gemeint ist, so wird diese als ORG annotiert. - „In der Europäischen Union [ORG] wurde heute über einen neuen Verfassungsantrag verhandelt.“ - „Innerhalb der Regierung [ORG] gab es einen Streit darüber, wie man mit der neuen Situation umgehen sollte.“

Wenn sich Ausdrücke eindeutig auf eine Stadt, ein Land bzw. ein Gebiet beziehen und wenn hiermit nicht eine politische verfasste Entität gemeint ist - Aber ich bezweifle, dass man argumentieren kann, die dadurch entstandene Situation hinsichtlich der Beziehung zwischen der Europäischen Union [ORG] und der Türkei [ORG] sei vergleichbar mit der Situation in Europa nach dem Zweiten Weltkrieg. - An diesen Entwicklungen in Europa bis zurück ins Mittelalter hat die Türkei keinen Anteil gehabt. - Seitdem die privilegierte Partnerschaft als ein dritter Weg im Raum ist, ist die Diskussion in Europa [à Europa als LOC, denn es geht um die räumliche Begrenzung, wo die Diskussion stattgefunden hat] - Aber ich bezweifle, dass man argumentieren kann, die dadurch entstandene Situation hinsichtlich der Beziehung zwischen der Europäischen Union [ORG] und der Türkei [ORG] sei vergleichbar mit der Situation in Europa nach dem Zweiten Weltkrieg. - Im südöstlichen Europa herrscht hohe Arbeitslosigkeit. - Im westlichen Europa ist die Säkularisierung weitestgehend abgeschlossen. - In ganz Europa herrscht eine unguete Stimmung. - In Europa sollte man über solche Dinge heutzutage nicht mehr streiten müssen. - Von Europa bis nach Moskau reicht heutzutage der Erfolg populistischer Bewegungen. - In Europa kann man heutzutage nicht mehr von einer gemeinsamen

Idee sprechen. - In Europa dürfe so etwas nicht mehr geschehen. - In Europa habe die gleichgeschlechtliche Ehe überall einen verfassungsrechtlichen Status zu besitzen. - Im neuen Europa werde so etwas der Vergangenheit angehören. - In Europa könnte man bestimmte Probleme nicht öffentlich aussprechen. - Über Europa würde viel zu wenig diskutiert. - In Europa würde diesem Problem viel zu wenig Aufmerksamkeit geschenkt. - In der heutigen Türkei wäre so etwas heutzutage unvorstellbar. - Im Westen braucht man sich über so etwas heutzutage zum Glück keine Sorgen mehr machen.

Referenzen, die auf Vorstellungen, Ideen oder Visionen für Europa hinweisen. Beispielreferenzen sind: Ein vereintes, friedliches, sicheres, demokratisches, stabiles usw. Europa

- „Natürlich brauchen wir für ein Europa der Bürger, für ein Europa des Rechtsraums der Freiheit, der Sicherheit einen leichteren und unmittelbareren Zugang der Bürger zu Gerichten und zum Rechtsschutz.“
- Ein christliches Europa kann unmöglich in einem politischen Gebilde mit der Türkei zusammenhängen.
- Europa ist diesbezüglich geteilt.

Weitere Beispiele für Referenzen, die auf einen weiteren geographisch-historisch-politischen Raum hinweisen

- Im europäischen Kontext / Raum wird aktuell heiß darüber debattiert, wie es mit Europa weitergehen soll.
- An der Peripherie Europas brodeln es mächtig.
- Im Herzen Europas streitet man über die Zukunft der sozialen Demokratie.
- Auf europäischem Boden grassiert die Fremdenfeindlichkeit.
- Die europäischen Kernländer sind christlich-abendländisch geprägt.
- Der europäische Kontinent muss sich zukünftig erheblichen politischen Umwälzungen stellen.
- Die europäischen Metropolen sind vor diesen Gefahren nicht gefeit.

Wenn Länder bzw. Regionen eindeutig als geographische oder kulturelle und nicht als politische verfasste Entitäten genannt werden: - Die Türkei ist nach geographischer Ausdehnung, Bevölkerungszahl, nationaler und kultureller Identität, ökonomischer und politischer Struktur von einer Bedeutung und Eigenart, die die Frage nach dem Konzept, der finalité der europäischen Einigung unausweichlich macht. [hier werden viele Aspekte in Bezug auf die Türkei angesprochen, die über die rein institutionelle Dimension hinausgehen, daher LOC] - „Dieses Argument ließe sich, wenn man konsequent zu Ende denkt, auch auf andere Länder übertragen, indem man sagt, die Türkei sei kein typisch arabisches und islamisches Land.“ - „Die kulturellen, die geistigen und die historischen Wurzeln Europas sind nicht die gleichen wie die der Türkei.“ - „Tatsache ist, dass - daran kommt man nach einem Blick auf die Landkarte nicht vorbei - erstmals in der Geschichte des europäischen Einigungsprozesses Verhandlungen mit einem Land begonnen werden, dessen Grenzen weit über Europa hinausgehen; sie reichen bis zum Iran, zum Irak und nach Syrien.“ [Europa als Kontinent. Eindeutige geographische Marker in diesem Satz: Landkarte, Grenzen] - „Sie werden weiterhin sagen, dass der Türkei [ORG] die Zusage

auch deshalb gegeben worden sei, weil sie historisch und politisch zu Europa gehöre.“ [Europa als LOC, da der Bezug auf den Raum, der historisch und politisch gedeutet wird, besteht]. - Es ist auch richtig, dass der 11. September des Jahres 2001 die Welt verändert hat. [da die Welt keine politische Organisation ist à LOC]

1.3 Schwierige Fälle:

Was ist mit solchen Ortsreferenzen, die nicht auf eine konkrete Organisation hinweisen, gleichwohl sie keine Ortsbeschreibung sind? - „Die Europäische Union ist entstanden aufgrund der Lehren aus dem Zweiten Weltkrieg und aus der Schreckensherrschaft des Nationalsozialismus sowie der festen Überzeugung der europäischen Länder, nie wieder Krieg gegeneinander zu führen.“ [ORG? oder LOC, weil hiermit keine konkrete Institution gemeint ist? Andererseits sind es schon Regierungen [ORG], die Kriege gegeneinander führen] - „Als Opposition können wir diese Entscheidung nicht verhindern; wir werden aber mit der Bevölkerung in diesem Lande über die Folgen einer solchen Entscheidung für Europa [für die EU?] und für die Gesamtlage weitersprechen.“ - Darin wurden soziale Mindeststandards für Europa [d.h. für alle Länder innerhalb EU; daher = ORG?] - Schon Jahre vorher, durch eine gemeinsam ausgerichtete Politik bei Kohle und Stahl, zusammengekommen, wurde 1957 in Rom [LOC] manifestiert und fortgeschrieben, dass man sich künftig in Europa auf eine gemeinschaftliche Wirtschaftspolitik einigen wolle. [à Europa als LOC, da man sich in Europa und nicht innerhalb einer konkreten Organisation auf einigen wolle; andererseits geht es schon um die Bestrebung von Nationalstaaten, sich auf eine gemeinschaftliche Wirtschaftspolitik zu einigen]

Außerdem: was machen wir mit Ausdrücken, die auf eine Organisation hinweisen, aber eigentlich auf die Menschen innerhalb der EU-Mitgliedstaaten hinweisen?

- Die Zunahme der Armut unter Kindern und älteren Menschen in der EU [ORG? oder LOC?] ist ein Skandal.

2. Sonstiges

[diese stehen z.T. auch in den generellen Annotationsrichtlinien, aber zur Wiederholung] - Die Anrede ganz zu Beginn einer Rede („Herr/Frau PräsidentIn“), die offensichtlich reine Formalität und nicht von inhaltlicher Relevanz ist, wird nicht annotiert. - Konkrete Personen, wie z.B. die Bundeskanzlerin, der Ministerpräsident, der EU-Kommissar werden als PER annotiert. - EU-Beitritt wird nicht annotiert. Weil "Beitritt" ist keine Entität. - Personalpronomen (wir, er, sie usw.) werden nicht annotiert. - „Staats- und Regierungschefs der Parteienfamilie der Europäischen Volkspartei“: Den ganzen Satz als PER und "der Parteienfamilie der Europäischen Volkspartei" zusätzlich als ORG - Opposition als ORG, da ein Zusammenschluss mehrerer Organisationen (i.e. Parteien) - Wenn man die eckigen und runden Klammern bei der Annotation vermeiden kann, dann raus. Wenn nicht, dann werden sie mitannotiert. - Wenn von „den Deutschen“ oder „den Franzosen“ als Gruppe gesprochen wird, dann = PER.

ABER: Wenn mit dem Ausdruck hingegen die politische Komponente („die Deutschen zwingen den Griechen einen rigiden Sparkurs auf“) gemeint ist, dann = ORG.

- Substantivierte Adjektive, wie z.B. „die Identität Europas“, die wiederum auch in die Form „die europäische Identität“ umgewandelt werden könnten, werden nicht annotiert, es sei denn, das zentrale Substantiv fällt unter eine der kodierrelevanten Entitäten [in diesem Fall ist das Substantiv „Identität“ und da wir keine abstrakten Konzepte annotieren, ist diese Textstelle auch nicht kodierrelevant]