

Diskurssegmentierung

Manfred Stede

Hinweise:

- Alle in diesem Kapitel genannten Textbeispiele ohne Quellenangabe sind fiktiv.
- Die vorliegenden Richtlinien verwenden eine Notation, wie sie "auf Papier" oder im Text-Editor verwendet werden kann. Für das praktische Vorgehen mit einer speziellen Annotationssoftware werden separate Hinweise für deren Benutzung ausgegeben.

3.1 Hintergrund

Ein großer Teil diskursbezogener Analyse und Annotation bezieht sich auf zugrundeliegende "minimale Einheiten", also auf eine Gliederung des Texts in *Diskurssegmente*. Im einfachen Fall hat man es mit einer Sequenz von Hauptsätzen zu tun, die trivialerweise dann auch die Segmente (gekennzeichnet durch eckige Klammern) bilden:

(3.1) *[Letzte Woche sind wir nach Chemnitz gefahren.] [Dort haben wir das Museum Gunzenhauser besucht.] [Es hat uns sehr gut gefallen!]*

Es bleibt allerdings nicht so einfach. In Texten begegnen wir im Hinblick auf Segmentierung zwei zentralen Komplikationen, die wir in den nachfolgenden Abschnitten besprechen werden.

- **Zusammengesetzte Sätze:** Wenn mehrere Hauptsätze nebenordnend verbunden sind, oder wenn zum Hauptsatz noch Nebensätze hinzutreten, muss entschieden werden, ob diese jeweils eigene Diskurssegmente bilden.
- **Fragmentarische Einheiten:** Fragmente unterschiedlicher Komplexität (ein einzelnes Wort, eine Phrase, etc.) können im Text die Position von Sätzen

einnehmen, oder auch als Parenthesen bzw. Appositionen in Sätze eingefügt sein. Ihre Behandlung muss für die unterschiedlichen Fälle ihres Auftretens geklärt werden.

Wegen dieser Komplikationen teilen wir den Annotationsprozess in eine Sequenz aufeinander folgender Schritte der Identifikation von Sinneinheiten und der anschließenden feineren Einteilung in strukturelle Einheiten. Diesen Einheiten wird dann jeweils ein *Typ* zugewiesen. Anhand der vorgenommenen Typisierung können die in anderen Kapiteln behandelten Annotationsebenen ihren eigenen Begriff von "minimalen Analyseeinheiten" definieren, indem sie ggf. Segmente, die qua ihres Typs für die Analyse-Ebene zu klein sind, an benachbarte größere Segmente anschließen.

3.2 Annotationsvorgehen: Übersicht

Die Annotation verläuft in drei Schritten, die jeweils den Text komplett behandeln:

1. Im ersten Schritt werden satzbeendende Interpunktionszeichen daraufhin geprüft, ob sie eine vollständige Sinneinheit abschließen, oder nur eine kaum interpretierbare, unvollständige Einheit. Es werden diejenigen Zeichen markiert, die dann insgesamt den Text in eine Folge von in sich geschlossenen Sinneinheiten gliedern.
2. Im zweiten Schritt wird jede einzelne Sinneinheit — soweit nötig — ihrerseits in eine Folge von strukturellen Einheiten gegliedert, die entweder Hauptsätze oder fragmentarische Einheiten darstellen.
3. Im dritten Schritt werden alle Hauptsätze daraufhin geprüft, ob sie durch Unterordnung (Nebensätze), Nebenordnung, oder Einbettung parenthetischer Einheiten noch einmal zu untergliedern sind.

Die in Schritt 1 nach semantisch/pragmatischen Kriterien markierten Sinneinheiten werden also in Schritt 2 und 3 ggf. weiter aufgegliedert.

Am Ende ist der Text vollständig in eine Sequenz von Sinneinheiten aufgeteilt. Jede Sinneinheit ist ihrerseits vollständig zerlegt in strukturelle Einheiten: Hauptsätze (HS) oder Fragmente (FR). Jedes token des Textes muss also einem HS oder FR zugeordnet sein. In HS können andere Einheiten rekursiv eingebettet sein (Hauptsätze, Nebensätze, Parenthesen). Grundsätzlich unzulässig sind *Überschneidungen* zwischen Segmenten: Es darf nicht sein, dass der Anfang eines Segments A innerhalb von Segment B liegt, das Ende von A aber außerhalb von B. Mit anderen Worten, unsere Notation der eckigen Klammerung ist stets (durch Zählen der Klammern) eindeutig interpretierbar.

Im Folgenden besprechen wir die drei Schritte im Detail. Danach gibt 3.6 eine Übersicht über alle besprochenen Typ-Bezeichner.

3.3 Schritt 1: Teilung in Sinneinheiten

“Satzartige” Einheiten enden mit satzbeendenden Interpunktionszeichen. Aber nicht jedes Auftreten eines solchen Zeichens beendet auch eine vollständige “Sinneinheit”: Manche Einheiten sind oftmals in sich nicht interpretierbar, sondern nur im Verbund mit davor oder dahinter stehenden Einheiten. Darum ist ein Interpretationsschritt erforderlich, der diejenigen Interpunktionszeichen identifiziert, die tatsächlich eine vollständige Sinneinheit abschließen.

Identifikation einer Sinneinheit: Betrachte die Einheit zwischen zwei satzbeendenden Interpunktionszeichen (. / : / ! / ?). Enthält sie einen Konnektor (Konjunktion, verbindendes Satzadverb), ignoriere diesen. Enthält sie anaphorische Pronomen oder Elipsen, ersetze diese durch ihr Antezedens. Ist die Einheit im Skopus einer zuvor eröffneten Modalität (wie in/direkte Rede, hypothetische Situation), füge diese Modalität der Einheit hinzu.

Prüfe nun, ob eine vollständige, in sich abgeschlossene Information vorliegt.

- Wenn ja, prüfe ob das nachfolgende Material eine neue Sinneinheit beginnt, oder aber der vorliegenden Sinneinheit angefügt werden muss (weil es sich um unselbständige, ergänzende Information handelt). Füge solche nachfolgenden unselbständigen Einheiten der aktuellen Sinneinheit hinzu.
- Wenn nein, prüfe in entsprechender Weise, ob durch Anfügung der nachfolgenden Einheit (das Material bis zum nächsten satzbeendenden Interpunktionszeichen) eine vollständige Sinneinheit entsteht; wenn nötig, setze diesen Prozess mit nachfolgenden Einheiten fort.

Wenn eine nach obiger Anweisung gefundene Sinneinheit zwei (oder mehr) durch Semikolon oder Komma getrennte Hauptsätze (oder nahezu vollständige Hauptsätze) enthält, dann prüfe ob an diesen Stellen auch eine Sinneinheitengrenze vorliegt. Wenn ja, markiere sie.

In diesen Richtlinien markieren wir solche Grenzen mit XX. Am Textanfang und am Textende wird kein XX gesetzt. Es folgen einige Beispiele:

(3.2) *Bube, Dame, König, As: Das sind die beliebten Spielkarten.*

Die Aufzählung vor dem Doppelpunkt wird erst durch den nachfolgenden Satz in eine Aussage eingebettet, daher liegt am Doppelpunkt keine Grenze vor.

(3.3) *Das ist das wichtigste: Man muss immer die Augen offenhalten.*

Hier steht vor dem Doppelpunkt zwar ein syntaktisch vollständiger Satz, aber "das" ist kataphorisch, der Sachverhalt erschließt sich nur durch Vorgriff auf den nachfolgenden Satz. Daher liegt am Doppelpunkt keine Grenze vor.

- (3.4) *Meier sagt immer, man soll die Augen offenhalten. XX Und gerade bei schlechtem Wetter immer vorsichtig fahren. XX Damit hat er recht!*

Der erste Satz ist eine vollständige, interpretierbare Einheit. Der zweite Satz ist eliptisch, wir ergänzen ihn gemäß obiger Anweisung mit *man soll*. Außerdem ergänzen wir die indirekte Rede, erhalten dann *Und Meier sagt, man soll gerade bei schlechtem Wetter immer vorsichtig fahren* – dies ist eine vollständige Einheit, so dass sie vom ersten Satz getrennt wird. Im dritten Satz verschiebt sich die Perspektive wieder zum Autor, die Anaphern *damit* und *er* füllen wir mit ihren Antezedenten, und wieder resultiert eine eigenständige Sinneinheit.

- (3.5) *Gestern hat mein Freund Hans mich wieder besucht; XX er war ja immer schon eine treue Seele.*

Anstelle des Semikolons könnte auch ein Punkt stehen – das Thema wechselt, und wir bekommen eine neue, separate Information. Darum wird hier eine Abtrennung vorgenommen.

- (3.6) *Der Kreisel ist Asbest verseucht. Nicht nur hier und da, sondern durch und durch. XX Zwar könnte man, wie beim Palast der Republik, den Bau bis aufs wacklige Stahlskelett entkleiden und neu aufbauen. XX Aber das würde mindestens 84 Millionen Euro, vielleicht auch das Doppelte kosten. (procon-kreiselpro)*

Der erste Punkt (nach *verseucht*) beendet einen Hauptsatz, der auch inhaltlich vollständig ist. Die nachfolgende Einheit jedoch ist nur in Verbindung mit diesem Satz verständlich, so dass der erste Punkt keine Sinneinheit beendet. Mit *zwar* beginnt ein neuer Gedanke, so dass davor die erste Sinneinheit-Grenze gesetzt wird. Die beiden Konnektoren "zwar..aber" sind semantisch sehr eng verbunden (ein "zwar"-Satz muss fast immer von einem "aber" oder "doch" Satz gefolgt werden). Unsere Definition besagt jedoch, dass wir Konnektoren ignorieren, daher ist zu prüfen, ob *Man könnte, wie beim Palast ... und neu aufbauen* einerseits und *Das würde mindestens ... kosten* andererseits jeweils eine vollständige Einheit bilden; mit Einsetzung des Antezedens (etwa: *der Umbau*) für die Anapher *das* ist das der Fall, so dass auch vor *Aber* eine Sinneinheit endet.

- (3.7) *Die Topographie des Terrors, eben ausgegraben und freigelegt, stand für sich. XX (...) XX Dort hingen Fotos, XX dort gab es einen Katalog XX – mehr als genug, um wenigstens zu ahnen, welcher Terror von diesem Ort ausgegangen ist. (procon-topographiecon)*

Nach dem Bindestrich gibt der Autor eine Bewertung, die sich offenbar auf *Fotos* und *Katalog* bezieht, so dass sich die Frage stellt, ob alles zu einer umfangreichen Sinneinheit zusammengefasst werden sollte; wir behandeln solche Fälle aber analog zu den oben besprochenen von indirekter Rede oder anderen satzübergreifenden Modalitäten (die keine Zusammenfassung zu einer Gesamteinheit auslösen) und trennen die einzelnen Aussagen daher in Sinneinheiten auf.

3.4 Schritt 2: Unterteilung der Sinneinheiten in Segmente

In diesem Schritt werden jeder Sinneinheit ein oder mehrere strukturelle Segmente zugeordnet, die jeweils durch [] abzugrenzen und mit einem *Typ* zu etikettieren sind (in diesen Richtlinien durch Tieferstellung markiert). Im einfachsten Fall bestehen die aus Schritt 1 resultierenden Einheiten jeweils aus einem einzelnen Hauptsatz. Das ist im eingangs besprochenen Beispiel 3.1 gegeben: Jeder HS stellt auch eine eigene Sinneinheit dar. An den Satzgrenzen ist also jeweils noch ein XX zu ergänzen, zudem ist jede schließende Klammer mit einem _{HS} zu versehen. Das komplett annotierte Beispiel ist also:

(3.8) *[Letzte Woche sind wir nach Chemnitz gefahren.]_{HS} XX [Dort haben wir das Museum Gunzenhauser besucht.]_{HS} XX [Es hat uns sehr gut gefallen!]_{HS}*

Oftmals wird eine Sinneinheit hingegen komplex sein, weil sie mehrere Hauptsätze oder auch Fragmente enthält. Das entscheidende Merkmal für die Unterteilung sind wieder die satzbeendenden Interpunktionszeichen. Jede von einem solchen Zeichen gebildete Einheit muss auf ihren Hauptsatz-Status geprüft werden.

Identifikation eines Hauptsatzes: Betrachte das Segment zwischen zwei satzbeendenden Interpunktionszeichen (. / : / ! / ?), bzw. das Segment, das durch eine mit Semikolon oder Komma markierte Sinneinheitsgrenze entstanden ist.

Weise ihm den Typ HS (Hauptsatz) zu, wenn es sich um einen strukturell vollständigen Satz (Aussagesatz, Fragesatz, Aufforderungssatz) handelt.

Besteht eine Sinneinheit nach dieser Erkennung allein aus (einem oder mehreren) HS, ist Schritt 2 für sie beendet.

Anderenfalls ist fortzufahren:

- Befindet sich in der Sinneinheit zumindest ein HS, sind die übrigen als fragmentarische Hauptsätze (HSF), einleitende Fragmente (FRE) oder beendende Fragmente (FRB) zu klassifizieren; siehe unten.

- Befindet sich in der Sinneinheit kein HS, ist darauf zu achten, dass zumindest ein fragmentarischer Hauptsatz zugewiesen wird, nämlich demjenigen Segment, das am ehesten als vollständige Informationseinheit bzw. Illokution interpretiert werden kann. Sind weitere Segmente vorhanden, sind sie als HSF, FRE oder FRB zu klassifizieren; siehe unten.

Identifikation eines Einleitenden Fragments: Betrachte das Segment zwischen zwei satzbeendenden Interpunktionszeichen (. / : / ! / ?).

Weise ihm den Typ FRE (einleitendes Fragment) zu, wenn es keine vollständige Information übermittelt und keinen eigenen Sprechakt darstellt, sondern die Funktion hat, das nachfolgende Segment vorzubereiten. Ein FRE ist nur in Verbindung mit dem nachfolgenden Segment (in Ausnahmefällen: mit den nachfolgenden Segmenten) interpretierbar.

Die "prototypischen" FRE sind recht kurz, doch die letzten beiden Beispiele illustrieren, dass es auch relativ komplexe Einheiten (auch Nebensatzstrukturen, die nicht in einen HS eingebettet sind) sein können::

- (3.9) [*Und dann:*]_{FRE} [*Ist der Preis nicht sowieso unbezahlbar?*]_{HS}
- (3.10) [*Klar ist:*]_{FRE} [*Die Situation wird sich nicht verbessern.*]_{HS}
- (3.11) [*Werner Meier.*]_{FRE} [*Das ist der neue Star im deutschen Fußball.*]_{HS}
- (3.12) [*Vielen Ostdeutschen wurde wendebedingt jedoch eingeredet:*]_{FRE} [*Ihr habt alles falsch gemacht!*]_{HS} (maz-9852)
- (3.13) [*Und selbst wenn schon 2012 eine europäische Metropole Ausrichter werden sollte:*]_{FRE} [*Man muss die Muskeln früh spielen lassen, um...*]_{HS}
(procon-olympiapro)

Identifikation eines Beendenden Fragments: Betrachte das Segment zwischen zwei satzbeendenden Interpunktionszeichen (. / : / ! / ?).

Weise ihm den Typ FRB (beendendes Fragment) zu, wenn es nicht satzwertig ist und keine vollständige Information übermittelt und keinen eigenen Sprechakt darstellt; sondern die Funktion hat, das vorangehende Segment zu ergänzen, darin eine Leerstelle zu füllen.

Beispiele für FRB:

- (3.14) [*Wer wäre wohl bereit zu zahlen:*]_{HS} [*die Lehrer, die Eltern, die Opernfreunde?*]_{FRB}
(procon-olympiacon)

(3.15) *[Diepensee zieht um.]_{HS} [Ohne Wenn und Aber.]_{FRB}* (maz-6993)

Relativ häufig treten knappe Fragmente auf, mit denen der Autor das vorangehende Segment kommentiert oder bewertet:

(3.16) *[Und dann kam der Rektor auf mich zu.]_{HS} [Na prima.]_{HSF}*

(3.17) *[Meier hielt es für einen schlechten Scherz.]_{HS} [Kann man wohl sagen.]_{HSF}*

Diese fassen wir *nicht* als Fortsetzung des vorangehenden Segments auf, denn es handelt sich um eigenständige Sprechakte; sie sind der nachfolgend beschriebenen Kategorie HSF zuzuordnen. Beachte den Unterschied zu 3.15: Auch dies könnte man als "Kommentierung" verstehen, doch wird hier eine genauere Charakterisierung des im HS genannten Ereignisses vorgenommen – es findet keine Perspektivenwechsel zwischen den beiden Segmenten statt.

Identifikation eines Hauptsatzfragments: Betrachte das Segment zwischen zwei satzbeendenden Interpunktionszeichen (. / : ! / ?) bzw. das Segment, das durch eine mit Semikolon oder Komma markierte Sinneinheitengrenze entstanden ist.

Weise ihm den Typ HSF (Hauptsatzfragment) zu, wenn es zwar kein strukturell vollständiger Satz ist, aber

- durch Rückgriff auf den vorangehenden Kontext zu einem solchen ergänzt werden kann (Auffüllen von Elipsen),
- oder einen eigenständigen Sprechakt darstellt.

Alle zuvor nicht als FRE oder FRB klassifizierten Segmente sollten nun der Kategorie HSF zugehörig sein. Bei der Interpretation kann der Rückgriff auf den vorangegangenen Kontext auch über Sinneinheit-Grenzen hinweg erfolgen, wie die ersten beiden der folgenden Beispiele zeigen. Zudem illustrieren einige der Beispiele, dass HSF nicht zwingend ein Verb enthalten (satzwertig sein) müssen: Semantisch "leere" Verbkomplexe wie *es ist* oder *dies gilt* dürfen bei der Interpretation ergänzt werden.

(3.18) *[Klaus fährt nach Berlin.]_{HS} XX [Kauft sich ein Motorrad.]_{HSF}*

(3.19) *[Der Kreisel ist nicht schön.]_{HS} XX [Ein Symbol der West-Berliner Filzwirtschaft in den späten Sechziger Jahren.]_{HSF}* (procon-kreiselpro)

(3.20) *[Der Bürgermeister verlas den Antrag.]_{HS} XX [Kichern und Prusten auf dem Fraktionsgestühl.]_{HSF}*

(3.21) *[Die Kahlschlagmentalität breitet sich aus.]_{HS} [In den Mietskasernen-Bezirken zum Beispiel.]_{HSF}* (procon-bäumecon)

3.5 Schritt 3: Unterteilung zusammengesetzter Sätze

Im letzten Schritt sind alle bisher gefundenen (durch satzbeendende Interpunktionszeichen markierten) Segmente darauf hin zu prüfen, ob sie selbst weiter untergliedert werden müssen (wodurch eine rekursive Einbettung entsteht). Im Wesentlichen handelt es sich dabei um die Behandlung zusammengesetzter Sätze.

Prinzipiell gilt in diesem Schritt, dass wir für aneinandergereihte Segmente keine (semantisch oder anders motivierte) Hierarchie etablieren wollen, sondern die Strukturen so "flach" wie möglich halten. Schematisch bedeutet das:

- NICHT: [[[Erst geschah X,] [dann geschah Y,]] [[aber ich habe P,] [und dann habe ich Q]]]
- SONDERN: [[Erst geschah X,] [dann geschah Y,] [aber ich habe P,] [und dann habe ich Q]]

Den erste Fall der zu besprechenden Untergliederung stellt die strukturelle *Unterordnung* (Subordination) dar, wie sie prototypisch bei Hauptsatz/Nebensatz-Strukturen vorliegt. Ein Beispiel zur Illustration:

(3.22) [*Letzte Woche sind wir nach Chemnitz gefahren, [was im Süden von Sachsen liegt.]*_{ANR}_{HS} XX [*Dort haben wir [(das war der Anlass der Reise)]*_{HS} *das Museum Gunzenhauser besucht.*]_{HS} XX [*Wie nicht anders zu erwarten:*]_{FRE} [*Es hat uns sehr gut gefallen!*]_{HS}

Der zweite Fall der Unterteilung zusammengesetzter Sätze betrifft die *Nebenordnung* (oft durch Konjunktionen) von Hauptsätzen oder Nebensätzen. Es wird sowohl der (in Schritt 2 behandelte) gesamte HS annotiert als auch die eingebetteten:

(3.23) [[*Meier öffnete die Flasche,*]_{HS} [*und er trank sie in einem Zug leer.*]_{HS}]_{HS}

Der dritte Fall sind (ähnlich zu Schritt 2) Fragmente, die in Haupt- oder Nebensätzen eingebettet sein können; das betrifft vor allem Parenthesen und ähnliche Einschübe. Oben ist im Beispiel 3.22 bereits eine solche Parenthese annotiert.

Nachfolgend besprechen wir diese drei Fälle der Reihe nach.

3.5.1 Unterordnung: Nebensätze

In diesem Abschnitt behandeln wir Segment-Einbettungen, die durch verschiedene Arten von *Nebensätzen* entstehen, d.h. durch Teilsätze, die "hinsichtlich Wortstellung, Tempus- und Moduswahl (...) abhängig vom übergeordneten Hauptsatz"

(Bußmann, 2002, p. 460) sind. Wir annotieren stets den *Gesamtsatz* als HS, und etwaige Nebensätze sind darin eingebettet. Also:

- NICHT: [Weil die Sonne scheint,] [legen wir uns an den Strand.]
- SONDERN: [[Weil die Sonne scheint,] legen wir uns an den Strand.]

Die Nebensatzbeendenden Interpunktionszeichen sind dem Nebensatz-Segment zugeordnet, denn es gilt (wie implizit schon bei den Hauptsätzen angenommen) die Grundregel, dass Interpunktionszeichen nicht von einem unmittelbar adjazenten Wort abgetrennt werden sollen, d.h., es darf keine Grenze zwischen *scheint* und dem Komma liegen. Analog bilden wir keine Grenze zwischen Wort und dem Satzbeendenden Punkt (oder anderem Interpunktionszeichen) – alle schließenden Klammern folgen also dem Satzende-Zeichen:

- NICHT: [Wir legen uns an den Strand, [weil die Sonne scheint].]
- SONDERN: [Wir legen uns an den Strand, [weil die Sonne scheint.]]

Für unser Typ-Inventar der Nebensätze übernehmen wir in diesem Abschnitt weitgehend die Klassifikation von Bußmann (2002). Neben einigen fiktiven Beispielen und denen aus PCC verwenden wir im Folgenden auch einige von canoo.net.¹

2.5.1.1. Erweiterung von Satzgliedern

Wir unterscheiden Subjektsatz, Objektsatz, Prädikativsatz und Adverbialsatz, die jeweils in verschiedenen Formen auftreten können. Die ersten drei ersetzen ein Satzglied bzw. ergänzen es zu einem satzwertigen Element; es handelt sich um Komplemente (also grammatisch notwendige Einheiten). Demgegenüber stellt der Adverbialsatz ein Adjunkt dar, das in einer semantischen oder pragmatischen Relation zum gesamten übergeordneten Satz steht.

Subjektsatz (Kürzel: SUB)

- dass-Satz: [[Dass ihr hier seid,]_{SUB} freut uns sehr.]_{HS} (canoo)
- ob-Satz: [[Ob ich den Schlüssel wiederfinde,]_{SUB} ist fraglich.]_{HS} (canoo)
- w-Satz: [[Wann wir ankommen werden,]_{SUB} ist ungewiss.]_{HS} (canoo)
- Infinitivkonstruktion, auch mit resumptivem Pronomen:
 [[Ihr zu verzeihen,]_{SUB} fiel ihm schwer.]_{HS} (canoo)
 [[Dass ihr hier seid,]_{SUB} das freut mich sehr.]_{HS}

¹<http://www.canoo.net/services/OnlineGrammar/Satz/Komplex/Funktion/Uebersicht.html>

- Uneingeleitet: [*Es ist besser, [du kommst noch einmal zurück.]_{SUB}*]_{HS} (canoo)

Bei den ersten vier Formen kann der Subjektsatz auch hinten stehen, wenn der Hauptsatz durch ein *es*-Korrelat eingeleitet wird. Beispiel:

[*Es ist ungewiss, [wann wir ankommen werden]*]_{SUB}]_{HS}

Objektsatz (Kürzel: OBJ)

- dass-Satz: [*Wir verstehen, [dass ihr uns nicht begleiten könnt.]_{OBJ}*]_{HS} (canoo)
- ob-Satz: [*Sie wissen nicht, [ob sie auch eingeladen sind.]_{OBJ}*]_{HS} (canoo)
- w-Satz: [*Ich wundere mich, [warum du erst so spät kommst.]_{OBJ}*]_{HS} (canoo)
- Infinitivkonstruktion: [*Wir bedauern, [sie enttäuschen zu müssen.]_{OBJ}*]_{HS} (canoo)
- Uneingeleitet: [*Er sagt, [er habe keine Zeit.]_{OBJ}*]_{HS} (canoo)
- wie-Satz: [*Sie fühlte, [wie ihre Kräfte nachließen.]_{OBJ}*]_{HS} (canoo)
- Akkusativ mit Infinitiv: [*Er hörte [ihn die Treppe hinaufkommen.]_{OBJ}*]_{HS} (canoo)

Auch beim Objektsatz sind oftmals verschiedene Abfolgen der Konstituenten möglich. Entscheidend für die Annotation ist die Identifikation der grammatischen Rolle (hier: Objekt).

Prädikativsatz (Kürzel: PRD)

Das typische Schema des Prädikativsatzes ist 'Subjekt + Kopulaverb + Konjunktion'.

[*Diese Wohnung ist, [was ich schon immer suchte.]_{PRD}*]_{HS} (canoo)

[*Die Frage ist, [ob die Sonne scheinen wird.]_{PRD}*]_{HS}

[*Was mir dann klar wurde, war, [dass die Sonne scheinen würde.]_{PRD}*]_{HS}

Adverbialsatz (Kürzel: ADV)

Für alle nachfolgend genannten Unterkategorien gilt, dass stets ein bestimmter semantischer oder pragmatischer Zusammenhang zwischen zwei Sachverhalten/Propositionen ausgedrückt wird (einer im Nebensatz, einer im Hauptsatz). In der Annotation wird die Art des Zusammenhangs aber nicht vermerkt; wir verwenden nur den allgemeinen Typ ADV.

- Satz mit adverbialer Konjunktion wie *während, indem, weil, wenn, obwohl, so-dass, damit* usw.
[*Er ging schwimmen, [obwohl er eine leichte Grippe hatte.]_{ADV}*]_{HS} (canoo)

- Verberstsatz (temporale, konditionale Lesart)
[[Konnte der Kosovo-Krieg noch mit dem Kampf für die Menschenrechte in Europa begründet werden,]_{ADV} geht es nun um die militärische Solidarität mit den USA in einem weit entfernten Dritte-Welt-Land.]_{HS} (maz)
[[Wäre ich Millionär,]_{ADV} würde ich sofort Dein Taschengeld erhöhen.]_{HS}
- Erweiterter-Infinitiv-Konstruktion
[Sie beeilte sich, [um den Zug zu erreichen.]]_{ADV}_{HS}
[Sie beeilte sich, [ohne spürbar in Aufregung zu geraten.]]_{ADV}_{HS}
[Sie beeilte sich [anstatt den Tag geruhsam ausklingen zu lassen.]]_{ADV}_{HS}
 (Zu unterscheiden von modal eingebetteten Infinitiven, die nicht abgetrennt werden: *[Sie wünschte sich den Zug zu erreichen.]_{HS}*)
- Partizipialkonstruktion (den HS-Sachverhalt ergänzend)
[[In Paris angekommen,]_{ADV} suchten wir ein Hotel.]_{HS} (cano)

2.5.1.2. Attributsatz

Der Attributsatz ersetzt nicht ein Satzglied, sondern er modifiziert ein overt gegebenes Satzglied, auf das er also explizit Bezug nimmt.

- Restriktiver Relativsatz (Kürzel ARR): Der Relativsatz beschreibt die Kopf-NP näher, damit sie identifizierbar wird
[Das Haus, [das dort drüben steht,]_{ARR} gehört meinem Onkel.]_{HS} (cano)
- Nicht-restriktiver Relativsatz (Kürzel ANR): Der Relativsatz übermittelt weitere Information zur Kopf-NP, die aber nicht primär die Identifikation erlauben soll
[Wohin mit den vom Urlaub übrig gebliebenen Münzen, [die bald sowieso nichts mehr wert sind?]_{ANR}]_{HS} (maz-8838)
- Partizipialkonstruktion (Kürzel APK): ein einzelnes Satzglied modifizierend; im Gegensatz zur o.g. Adverbialsatz-Partizipialkonstruktion
[Spät werden jetzt vereinzelt Proteste laut, [angeregt und unterstützt allerdings auch nur von Einzelpersonen.]]_{APK}_{HS} (maz-9612)
- Sonstiger Attributsatz (Kürzel ATT): Es wird ein Satzglied modifiziert, aber nicht im Sinne der obigen drei Kategorien
[Die Idee, [einen Kaffee zu trinken,]_{ATT} ist prima.]_{HS}
Es stellt sich die Frage, [ob die Preiserhöhung gerechtfertigt ist.]]_{ATT}_{HS}

2.5.1.3. Weiterführender Nebensatz (Kürzel: WEI)

Der weiterführende Nebensatz nimmt Bezug nicht auf ein einzelnes Element des vorangehenden (Teil-)Satzes, sondern auf dessen gesamte Aussage.

[Er hat uns eingeladen, [was uns sehr freut.]]_{WEI}_{HS} (cano)

2.5.1.4. Annotation unklarer Nebensätze (Kürzel: UNS)

Wenn ein Nebensatz vorliegt, die Typ-Bestimmung anhand der o.g. Kategorien aber nicht gelingt, ist das Kürzel UNS für “unklarer Nebensatz” zu verwenden.

2.5.1.5. Im Zweifel: Verbalkomplexe abtrennen

Mitunter können Zweifel aufkommen, ob ein bestimmter Verbalkomplex im Sinne einer ‘clause’ als Segment abzutrennen ist oder nicht. Etwa in diesem Fall:

(3.24) *Nein, man sollte mal schön alles so lassen wie es ist.*

In solcherlei Zweifelsfällen gelte die Präferenz für eine Abtrennung. Wenn nicht klar ist, welche Art von Nebensatz vorliegt, ist wie eben gesagt das Kürzel UNS zu verwenden..

3.5.2 Nebenordnung

Sind Hauptsätze oder andere satzwertige Einheiten durch eine nebenordnende Konjunktion (n.K.) verbunden oder durch Komma/Semikolon getrennt, so werden – zusätzlich zur übergeordneten HS/HSF-Annotation – alle als eigenständige Segmente annotiert. Als n.K. gelten prototypisch *und*, *oder* und *denn*. *Aber* kann ebenfalls als n.K. gebraucht werden. Zudem fassen wir *sondern* als n.K. auf. Die n.K. wird jeweils in das nachfolgende Segment aufgenommen (d.h., sie leitet stets ein Segment ein, analog zur unterordnenden Konjunktion).

Beispiele für Hauptsatz(fragment)-Nebenordnung:

(3.25) *[[Die erste Sammelbox war schneller als gedacht prall gefüllt und schwer,]_{HS} [und sie konnte bereits an den Initiator der Sammelaktion, den Internationalen Bund (IB), übergeben werden.]_{HS}]_{HS} (maz-8838)*

(3.26) *[[Heute scheint die Sonne,]_{HS} [morgen soll es angeblich auch schön sein,]_{HS} [und übermorgen ist mir das Wetter egal.]]_{HS}]_{HS}*

(3.27) *[[Der Mann bezahlte viele Handwerker nicht]_{HS} [und wurde voriges Jahr zu einer mehrjährigen Haftstrafe verurteilt.]]_{HSF}]_{HS} (maz-6193)*

(3.28) *[Berlin braucht Hilfe.]]_{HS} XX [[Braucht finanzielle Unterstützung]_{HSF} [– und bittet darum nicht nur,]_{HSF} [sondern klagt sie ein.]]_{HSF}]_{HS} (procon-olympiacon)*

(3.29) *[[Bei schlechtem Wetter die Frei- und Sommerbäder zumachen]_{HSF} [und alle Hallen öffnen?]]_{HSF}]_{HSF}*

Beispiele für Nebensatz-Nebenordnung:

- (3.30) [*Leute, [die auf Kellermauern und alte Kacheln starren]*_{ARR} [*und ihren Gedanken nachhängen,*]_{ARR} *gelten als unfähig zu erfassen, [was geschehen ist.]*_{OBJ}]_{HS}
(procon-topographiecon)
- (3.31) [*Was dort fehlt, ist ein Dach unter dem freien Himmel.*]_{HS} [*Eine Gelegenheit, [sich auszuruhen,]*_{ATTJ} [*sich zu sammeln,*]_{ATT} [*in Büchern nachzulesen]*_{ATT} [*und das Gesehene mit Hilfe von Bild und Ton zu vertiefen.*]_{ATT}]_{HSF}
(procon-topographiepro)

Die n.K. sind oft fakultativ; auch Reihungen von HS/HSF ohne eine solche Konjunktion sind zu annotieren, wenn einzelne Teile des Gesamtsatzes die Kriterien von HS bzw. HSF erfüllen:

- (3.32) [*[Der Kreisel ist auch innen häßlich,]*_{HS}, [*zudem zugig]*_{HSF} [*und Energie verschleudernd.*]_{HSF}]_{HS} (procon-kreiselpro)

Umgekehrt ist, wie eingangs erwähnt, die Präsenz einer n.K. kein *hinreichendes* Kriterium für zu annotierende Nebenordnung. Sie sind nur zu annotieren, wenn sie satzwertige Einheiten (Verbalkomplexe) verbinden. Durch verschiedene Grade der Elision kann es — mit oder ohne n.K. — Zweifelsfälle geben, als Faustregel gilt dann: eher keine Trennung vornehmen.

- (3.33) [*Das Auto war schick und auch teuer.*]_{HS}

- (3.34) [*Blauer Himmel, aber nur 17 Grad.*]_{HS}

3.5.3 Parenthetische Einschübe

Parenthetische Einschübe können durch Kommata, ein Paar von Bindestrichen oder Klammern begrenzt sein. Sie treten in vielen Formen auf, und wir annotieren hier ausschließlich *satzwertige* bzw. illokutionär-vollständige Einschübe: Nach Prüfung der Kriterien (vgl. Schritt 2) erhalten Hauptsätze und Hauptsatzfragmente wie gewohnt das Kürzel HS bzw. HSF.

Hat die Parenthese syntaktisch die Gestalt eines Nebensatzes, wählen wir ebenfalls das passende Kürzel gemäß der in Abschnitt 3.5.1 beschriebenen Klassifikation. Die Interpunktionszeichen selbst behandeln wir so, dass Klammern bzw. das Bindestrich-Paar stets dem Einschub zugeordnet werden.

Das erste der folgenden Beispiele illustriert das HS/HSF-Kriterium "Illokution": Mit dem Einschub nimmt der Autor eine persönliche Bewertung des berichteten Sachverhalts vor, was einem eigenständigen Sprechakt gleichkommt.

- (3.35) [*Gestern hat der Lehrer [– ganz schön lächerlich! –]*_{HSF} *mit blauen Briefen für verspätete Schüler gedroht.*]_{HS}

Hauptsatz	vollständig unvollständig	HS HSF
Nebensatz	Satzgliederweiterung Subjektsatz Objektsatz Adverbialsatz Prädikativsatz Attributsatz restr. Relativsatz nichtrestr. Relativsatz Partizipialkonstruktion sonstige weiterführend unklar	SUB OBJ ADV PRD ARR ANR AKP ATT WEI UNS
Fragment	einleitend beendend	FRE FRB

Tabelle 3.1: Typen der Diskurssegmente

(3.36) [*Heute war, [so soll es ja auch sein,]_{HS} das Kind pünktlich in der Schule.*]_{HS}

(3.37) [*Heute war das Kind [(das öfters mal Probleme mit dem Aufstehen hat)]_{ANR} pünktlich in der Schule.*]_{HS}

Da wir lediglich “satzwertige” Einschübe behandeln, werden beispielsweise nominale Appositionen, die unmittelbar eine NP modifizieren, nicht annotiert. Sie verbleiben einfach als Teil des übergeordneten Satzes:

(3.38) [*Herder, der Dichter, wirkte unter anderem auch in Weimar.*]_{HS}

Eine weitere Konsequenz der Beschränkung auf satzwertige Einschübe in diesem Schritt ist, dass es keine einleitenden oder beendenden Fragmente (FRE, FRB) gibt, die in übergeordnete (Haupt-/Neben-) Sätze eingebettet sind. Tritt also an der Peripherie eines Satzes fragmentarisches Material auf, wird es nicht gesondert annotiert.

(3.39) [*Klar, Gedichte Auswendiglernen ist langwierig.*]_{HS}

3.6 Zusammenfassung der Typ-Bezeichner

Tabelle 3.1 vermittelt einen Überblick über alle oben eingeführten Typen von Segmenten und ihre Kürzel.