

Richtlinien für die Annotation von Begründungen

Andreas Peldszus

Manfred Stede

1 Einleitung

1.1 Zielsetzung

Einer Reihe theoretischer Beschreibungsansätze (LDM, SDRT, RST, etc.) zufolge lässt sich Texten eine Baum- oder Graphstruktur zuschreiben, die daraus entsteht, dass *Kohärenzrelationen* zwischen (zumeist benachbarten) Segmenten angesetzt werden, die den inhaltlichen Zusammenhang zwischen diesen Segmenten charakterisieren. In dieser Annotationsaufgabe betrachten wir nur ein Teilproblem davon, nämlich die *kausalen* Relationen, mit denen Ursachen für Sachverhalte geschildert, oder Gründe für Handlungen bzw. Überzeugungen genannt werden. Momentan interessieren wir uns dabei nur für *lexikalisch markierte* Relationen, bei denen ein Konnektor (*da, weil, denn, ...*) den Zusammenhang markiert.

In der vorliegenden Fassung der Aufgabe ist die Identifikation der signalisierenden Lexeme bereits fertiggestellt. Sie bekommen daher in MMAX2 alle zu behandelnden Lexeme angezeigt und müssen nun entscheiden,

- welches genau die Segmente sind, die von dem Lexem verbunden werden (es sind meist zwei, können aber auch mehrere sein, dann sprechen wir von *Verkettung*),
- welchen illokutiven Beitrag jedes Segment hat.

Beides wird in diesen Richtlinien beschrieben und durch Beispiele illustriert.¹ Bitte versuchen Sie sich möglichst genau an die Vorgaben zu halten. Sollten diese Vorgaben Ihnen an bestimmten Stellen aber zu ungenau erscheinen, um eine Entscheidung zu treffen, machen Sie bitte eine Notiz, damit wir die Richtlinien ggf. weiter verbessern können.

Ziel des gesamten Annotationsprojekts ist, einerseits überprüfte und nachvollziehbare Richtlinien für die Identifikation und Markierung kausaler Relationen zu erhalten (solche gibt es unseres Wissens bisher nicht), andererseits auch ein Korpus aufzubauen, das gemäß dieser Richtlinien annotiert ist und für verschiedene Forschungszwecke verfügbar gemacht werden soll. Ihr Beitrag besteht jetzt primär darin, uns die Evaluation der Richtlinien zu erlauben; Ihre Annotationen fließen möglicherweise aber auch in das Korpus ein (nachdem wir alles nochmals überprüft haben).

Unsere Textsorte sind Beschreibungen/Rezensionen von Hotelbesuchern aus einem WWW-Forum. Das bedeutet, dass wir es mitunter mit eigenwilligem Stil und auch mit grammatischen Fehlern zu tun haben. Versuchen Sie, dem gegenüber möglichst tolerant zu sein.

1.2 Notationskonventionen

Zur Veranschaulichung der Richtlinien werden Beispiele gegeben. Dabei stehen eckige Klammern für Segmentgrenzen und Unterstreichungen für Konnektoren.

¹Dies ist die finale, korrigierte Version der Richtlinien.

- (1) [Man wird jedoch in klimatisierten Reisebussen abgeholt] , so dass [die lange Fahrt doch recht angenehm ist] . 048

Am Ende jedes Beispiels ist ein Hinweis auf die Quelle. Einfache Zahlen verweisen dabei auf die Nummer des Hotelreviews, das Leere-Mengen-Zeichen \emptyset zeigt ausgedachte Beispiele an.

- (2) [Dieses Beispiel ist ausgedacht] , weil [es im Korpus nicht vorkommt] . \emptyset

Manche Segmente sind diskontinuierlich, d.h. sie werden zwischendurch von Worten unterbrochen, die nicht diesem Segment zugehörig sind. Die dadurch entstehenden Teile eines diskontinuierliche Segments werden durch ein numerisches Subskript eindeutig zugeordnet.

- (3) Im Sand liegen häuft Samen herum , [die dummerweise Stachel mit Widerhaken versehen sind] , und [₁ die] deshalb [₁ lästig bis schmerzhaft werden können] ! 242

Falls im Fließtext auf bestimmte Segmente eines Beispiels Bezug genommen wird, können aber auch ununterbrochene Segmente nummeriert sein.

Fokuspartikel, die sich auf den Konnektor beziehen, werden kursiv hervorgehoben.

- (4) *Nur wegen* [des schlechten Wetters] [läßt man doch keinen Ausflug ausfallen] . \emptyset

1.3 Konnektoren, Segmente und Fokuspartikel

Für die Annotation von Begründungen gibt es drei Sorten von Markables: *Konnektoren*, *Segmente* und *Fokuspartikel*. Das Zusammenspiel dieser unterschiedlichen Typen von Markables ist in Abbildung 1 schematisch dargestellt. Hierbei stehen die Knoten des Graphens für die Markables des jeweiligen Typs, die Pfeile für Relationen zwischen den Markables. In den folgenden Untersektionen wird auf jeden dieser Markable-Typen exemplarisch eingegangen werden.

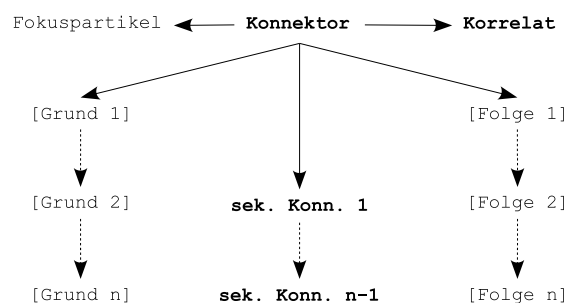


Abbildung 1: schematische Darstellung der Konnektor-Annotation

1.3.1 Konnektoren

Der Konnektor ist das zentrale Markable jeder Begründung. Er verbindet die verschiedenen Segmente, in denen die Gründe und die Folgen aufgeführt sind, er verweist auf Fokuspartikel, auf eventuelle Sekundärkonnektoren oder –sofern vorhanden– auf sein Korrelat (mehr dazu siehe unten). Konnektoren sind in MMAX Markables des CONJUNCTION-Levels.

Da der Konnektor hier eine solch zentrale Rolle spielt, werden Begründungen, die sprachlich nicht durch einen Konnektor markiert sind, auch nicht annotiert. Beispiel (5) wird also nicht annotiert, Beispiel (6) aber enthält einen Konnektor und muss annotiert werden.

(5) Wir waren im Pool nicht schwimmen . Er war sehr dreckig . \emptyset

(6) [Wir waren im Pool nicht schwimmen] , denn [er war sehr dreckig] . \emptyset

Eine vollständige Liste der Lexeme, die kausale Relationen markieren, sei hier nicht angegeben, zumal in der gestellten Annotationsaufgabe diese schon vormarkiert sind. Allerdings sei noch auf eine Ausnahme hingewiesen, bei der der vormarkierte Konnektor für sich allein eigentlich keine Kausalität signalisiert, die Konstruktion insgesamt trotzdem markiert werden soll: In *so...*, *dass*-Konstruktionen (Es war so laut, dass ich nicht schlafen konnte.) ist der Subjunktor *dass* als Konnektor markiert, während das *so* Teil des Grund-Segments sein soll.

1.3.2 Segmente: Gründe und Folgen

Eine der Hauptaufgaben in der Annotation wird die Identifikation des begründenden und des begründeten Materials im Text sein. Im folgenden wird Ersteres immer *Grund*, das Letztere immer die *Folge* genannt werden. Diese werden in MMAX als Markables des Typs CAUSE und CONSEQUENCE markiert. Der genauere Umgang mit diesen Segmenten wird in den folgenden Sektionen erläutert werden. Sektion 2 erklärt, welche Ausdehnung die Segmente haben sollen, Sektion 3 geht genauer auf die Verknüpfung von Segmenten ein und Sektion 4 beschreibt, auf welche Weise das jeweilige Segment zu klassifizieren ist.

1.3.3 Fokuspartikeln

Fokuspartikel wie *nur*, *sogar*, *auch*, *nicht*, *vor allem* etc. werden ausschließlich dann markiert, wenn sie sich auf den Konnektor der Begründung beziehen. In diesem Fall werden sie als Markable des Typs FOCUSPARTICLE annotiert. Außerdem muss der Konnektor auf den Fokuspartikel zeigen (siehe Abbildung 1).

(7) *Sogar* wegen [des Wecker-Klingelns] [beschwerte sich der Nachbar] . \emptyset

Es kann vorkommen, dass es syntaktisch nicht eindeutig ist, ob sich der Fokuspartikel auf den Konnektor oder auf eine andere Konstituente bezieht.

(8) Wegen [des schlechten Wetters] [bin ich jedenfalls nicht abgereist] . \emptyset

Zum Beispiel hat (8) zwei Lesarten. In der Lesart, wo sich der Fokuspartikel (*nicht*) nicht auf den Konnektor bezieht sondern auf *abgereist*, wird das schlechte Wetter als Grund für die ausgefallene Abreise angegeben. In der Lesart, wo er sich aber auf den Konnektor bezieht, wird gesagt, dass etwas anderes als das schlechte Wetter der Grund für die (gelungene) Abreise war.

Der Annotator muss sich in einem solchen Fall für eine mit dem Kontext vereinbare Lesart entscheiden und den Fokuspartikel dementsprechend markieren, oder nicht.

1.3.4 Sekundäre Konnektoren

Es kann manchmal vorkommen, dass eine Begründung, die mehrere Gründe oder Folgen anführt (siehe dazu Sektion 3), auch mehrmals den gleichen Konnektor enthält. Wenn der zweite Konnektor tatsächlich nicht eigenständig, sondern nur eine Wiederholung des ersten ist, dann wird dieser als sekundärer Konnektor markiert.

(9) [All Inclusive] , da [man mit einem Sohnebaby nicht immer die angegebenen Zeiten einhalten kann] und *vor allem* da [man eh keine Zeit hat sich all zu viel vorzunehmen] 251

Die Begründung des obigen Beispiels enthält zweimal den Konnektor *da*, doch nur der erste wird als zentraler markiert, während der zweite als sekundär verstanden wird. Der sekundäre Konnektor unterscheidet sich in der Annotation in MMAX von einem normalen nur dadurch, dass er nicht auf Gründe oder Folgen verweist, ist aber sonst genauso ein Markable des CONJUNCTION-Levels. Er wird nur der Vollständigkeit halber mitmarkiert, und weil ein Fokuspartikel sich noch auf ihn beziehen könnte (wie im Beispiel *vor allem*). Entsprechend dem Schema (siehe Abbildung 1) verweist der für die Begründung zentrale Konnektor auf seinen sekundären Konnektor und dieser eventuell weiter auf den nächsten sekundären Konnektor.

Ein sekundärer Konnektor muss aber nicht immer string-identisch sein. Mehrere Gründe oder Folgen können auch durch unterschiedliche Konnektoren und mit unterschiedlichen syntaktischen Segmenttypen angegeben werden. Ausschlaggebend ist nur, dass jeder weitere Konnektor wirklich die zu annotierende Begründung weiterführt und nicht eigenständig eine andere Begründung aufspannt.

(10) Aufgrund [des unerträglichen Gestanks], wegen [der schlechten Lage] und weil [ich dort zweimal bestohlen wurde], [werde ich dieses Hotel nicht mehr besuchen] . \emptyset

1.3.5 Korrelate

Neben den Sekundärkonnektoren sind auch Korrelate als Konnektoren vormarkiert. Im folgenden Beispiel finden sich in einer Begründung mit nur einem Grund und einer Folge zwei Vormarkierungen, *deswegen* und *weil*.

(11) [₁ Ich habe sie] *nur* deswegen [₁ geheiratet] , weil [sie so viel Geld hatte] . \emptyset

Die Korrelate, die uns während der Annotation begegnen werden, sind Platzhalter-Elemente, die unbesetzte syntaktische Positionen füllen und meistens kataphorisch auf ein Segment verweisen. Ihre Leistung ist also vornehmlich syntaktisch, sie tragen semantisch nicht eigenständig zur Bedeutung des Satzes bei. Im o.g. Beispiel ist *weil* der zentrale Konnektor der Begründung, währenddessen *deswegen* redundant ist. Um diesem Unterschied Rechnung zu tragen, sollen Korrelate als solche markiert werden, nämlich indem der eigentliche Konnektor auf sein Korrelat mit der Korrelat-Relation verweist. Bezieht sich auf das Korrelat noch ein Fokuspartikel (wie im obigen Beispiel *nur*), so sollte das Korrelat darauf zeigen. Wie auch schon die Sekundärkonnektoren sind also auch Korrelate technisch gesehen, d.h. in der Implementation des Schemas, Konnektoren, verweisen aber selbst nicht auf Segmente.

2 Segmentierung

Wenn im Text eine Begründung und der erforderliche Konnektor ausfindig gemacht worden ist, stellt sich die Frage, welche genaue Ausdehnung die Segmente für Gründe und Folgen haben sollen. Dabei steht der Annotator oft im Konflikt zwischen der Möglichkeit größere ununterbrochene Segmente zu erstellen, die aber Textmaterial enthalten, daß ihm nicht als begründungsrelevant erscheint, oder andererseits ein insgesamt kürzeres, dafür aber unterbrochenes, diskontinuierliches Segment zu erstellen. Als allgemeine Regel für solche Abwägungen soll folgendes gelten:

1. Minimiere die Außengrenzen des Segments.
2. Vermeide diskontinuierliche Segmente.

Damit soll sichergestellt sein, dass die Segmente bis auf bestimmte Ausnahmen immer kontinuierlich, dennoch aber möglichst minimal sind.

2.1 Syntaktischer Typ der Segmente

Im einfachsten Fall sind die Segmente vollständige Haupt- und Nebensätze, die weder Einschübe noch weitere untergeordnete Nebensätze enthalten.

(12) Da [das Hotel auch eine eigene Kinderanimation bietet] , [findet man viele Familien in der Anlage] . 001

Neben satzförmigen Segmenten finden sich zudem Nominalisierung, vornehmlich bei Konnektoren wie *wegen, halber, aufgrund* etc..

(13) [Sehr zu empfehlen ist dieses Hotel] *vor allem* wegen [der tollen Animation] . 049

(14) [Der Bequemlichkeit] halber [habe ich nach einer All Inclusive Anlage gesucht] und ... 052

Ebenfalls Nominalisierungen nimmt die Präposition *zu*. Sie tritt aber nur dann als Konnektor für Begründungen auf, wenn sie der Angabe eines Zweckes dient. Folglich wird Beispiel (15) als Begründung markiert und Beispiel (16) nicht.

(15) Zur [Erfrischung] [holten wir uns einen Cocktail] . \emptyset

(16) Zum Erschrecken aller wurde ein wildes Krokodil am Strand gesichtet . \emptyset

Außerdem können Infinitive als eigenständiges Segment fungieren. Häufig findet sich das bei Zwecksetzungen mit *um...zu*-Infinitiven:

(17) Um [uns etwas abzulenken] [gingen wir erstmal eine Runde im Meer baden] . \emptyset

Von diesen drei grundlegenden syntaktischen Typen von Segmenten (Satzartige, Nominalisierungen und Infinitiv) abgesehen, trifft man häufig auch auf Prädikationen die durch Ellipsen entstehen. Ein immer wieder zu findendes Schema ist die Konjunktion von zwei Zuschreibungen, die in einem begründenden Verhältnis stehen, das meistens durch Konnektoren wie z.B. *deswegen, deshalb* oder *daher* angezeigt wird.

(18) [Die Hitze ist trocken , nicht schwülwarm] und deshalb [gut zu vertragen] . 035

(19) [Regen ist in Ägypten normalerweise eine Seltenheit] und deshalb [etwas , worüber sich die Leute freuen] . 026

Eine extreme Form von elliptischer Formulierung mit *weil* gibt sich mal satzförmig, mal nominal:

(20) [Prima] , weil [billig] ! \emptyset

(21) Die [unangenehme] , weil [etwas übersteuerte] Musik bei der Animation ... \emptyset

In den folgenden Untersektionen soll anhand von Beispielen gezeigt werden, wie die linke und rechte Segmentgrenze in typischen Situationen gezogen werden soll, und wie mit eingeschlossenem, aber nicht begründungsrelevantem Material umgegangen werden soll.

2.2 Linker Rand von Segmenten

Matrixsätze von Relativsätzen ausschließen.

(22) und außer dem gab es jeden Tag vertrocknete Nudeln , [die wirklich teilweise wieder hart wurden] weil [sie Stunden in der Wärme lagen] ! 248

Freie Relativsätze, die als Subjekt oder Objekt zum Verb im Segment gehören, einschließen.

(23) [Wer will kann über den Felsen von San Agustin noch weiter] , da [hinter diesem Ort schon der nächste Strand beginnt] . 001

2.3 Rechter Rand von Segmenten

Parenthesen ausschließen.

(24) [Leider gibt es zu wenige Kaffeekannen] (die meisten sind auch noch kaputt) , so dass [nicht jeder Tisch eine bekommen kann] . 246

Restriktive Relativsätze einschließen.

(25) [Auf der Terrasse gibt es viele Vögel die beim Essen stören] , deshalb [empfehle ich nur im Restaurant zu essen] . 247

Nicht-restriktive Relativsätze aber ausschließen.

(26) [Zwischen der Hotelanlage und dem Strand liegt eine Düne] , die mit niedrigen Bäumen bewachsen ist (Strandweiden-Bäume) . dadurch [beträgt der Weg zwischen Pool und Strand etwa 150m] . 246

Verbkomplemente einschließen.

(27) Da [hier für nichts weiter zu zahlen ist] , [erübrigt es sich eigentlich besonders zu erwähnen , das keine Gebühr erhoben wird] . 002

(28) [Das war mir persönlich lieber] , da [ich sicher sein wollte , dass er es auch wirklich selbst bekommt] . 017

(29) [unsere Reisebegleitung von ITS hat uns gleich am Anfang den Tip gegeben die Getränke in geschlossenen Taschen mitzunehmen] , da [hier keiner nachschauen darf] . 042

2.4 Segmentinternes

Wie am Anfang des Abschnitts beschrieben, sollen zwar die Außengrenzen des Segments minimal sein, alles, was innerhalb des Segments auftritt, bleibt aber zur Wahrung der Kontinuität eingeschlossen. Das gilt auch für Parenthesen und Relativsätze. In solchen Fällen wird also begründungsirrelevantes Material eingeschlossen um die Kontinuität des Segments sicherzustellen.

(30) [Aber sie müssen sich beeilen] , denn [auch der materielle Schaden (Waschbecken wurden ausgerissen) scheint bald zu gross] . 096

(31) [angeblich arbeitet das Management bereits an einer Lösung] , denn [die deutschen Touris , die das ' echte ' Geld bringen werden rar] . 096

Das gilt auch für in ein Segment eingebettete Begründungen.

(32) [Es kann wegen der Einflugschneise zu Fluglärm kommen] , da [Tags über das eine oder andere Flugzeug direkt über das Hotel fliegt] . 017

Die eingebettete Begründung muss für sich selbst natürlich auch noch annotiert werden (siehe unten).

2.5 Diskontinuität

Diskontinuität eines Segmentes ist dann und nur dann erlaubt, wenn der zugehörige Konnektor (eventuell inklusive dessen Fokuspartikel und anderem Segment) oder sein Korrelat interveniert.

(33) [₁ Es kann] wegen [der Einflugschneise] [₁ zu Fluglärm kommen] , ... 017

(34) [Zu seiner gewählten Essenszeit hat man jedoch immer einen reservierten Tisch] . [₁ Leider hat man] dadurch [₁ nicht immer denselben Kellner] . 246

(35) [₁ man sollte] wegen [Durchfallgefahr] [₁ nicht das Leitungswasser zum Waschen benutzen] ! 019

(36) [₁ Wir haben] , da [wir nur eine Woche da waren] [₁ nur die Stadtrundfahrt mit dem Minizug in Mahdia und den Ausflug zum Wochenmarkt gemacht] . 042

Ein Beispiel für ein Korrelat, das ein Segment teilt, ist folgendes:

(37) [₁ Er hatte] nur darum [₁ noch einmal den Kellner gerufen] , weil [er noch einen Kaffee wollte] . 0

2.6 Sonstige Fälle

Es kann vorkommen, dass nur einer der Teilsätze der Konjunktion begründungsrelevant ist. Sollten sich beide Sätze ein Subjekt teilen, darf man sich nicht davor scheuen, ein unvollständiges Segment zu markieren.

(38) Der Strand wird täglich gereinigt und [wird fast nur von Hotelgästen besucht] , da [hier die Hoteleigenen Wachleute Patrolieren] 251

3 Verknüpfung von Segmenten

Wie zu Beginn schon beschrieben und in Abbildung 1 dargestellt, sollen komplexere Begründungen, die mehr als einen Grund oder eine Folge anführen, so annotiert werden, dass die zugehörigen Segmente miteinander verkettet werden. Ermöglicht wird dies in MMAX durch die Relation CHAIN-RELATION. Sie erlaubt die Verknüpfung von zwei Markables gleichen Typs. Der Konnektor zeigt demnach auf drei Ketten, die der Gründe, die der Folgen und evtl. auf die Kette der Sekundärkonnektoren. Wovon die o.g. Darstellung bewusst abstrahiert, ist die Tatsache, dass die Segmente im Textfluss linear angeordnet sind. Die Reihenfolge der Verkettung bestimmen folgende Regeln:

- Mehrere begründungsrelevante Segmente eines Typs werden von links nach rechts verkettet.
- Der Konnektor zeigt jeweils auf das erste Glied der Kette, also auf das am weitesten links stehende Segment der Kette.

Beispiele für komplexe Begründungen sind diese:

(39) ...mussten wir ein Stück laufen , [was uns jedoch nichts ausmachte] , da [unser Gepäck später gebracht werden würde] und [die Gartenanlage sehr schön war] ! 007

(40) Jeden Abend war live Programm , [was wir uns aber nie angesehen hatten] , da [es doch ein etwas älteres Publikum ab 40 ansprach] und [Gran Canaria soviel ausserhalb zu bieten hatte] . 111

(41) ...und da [es so eine tolle Erinnerung ist] und [der Funfaktor so groß war] , [mußte dies in unser Erinnerungsrepertoire] . 063

(42) [Erholung hatten wir auch wenig] da [man mittags am Strand mit lauter Animations Musik belästigt wird] und [ebenso nachts bis 2 Uhr die laute Disco Musik hört] 072

Nominalisierungen, infinitivische und elliptische Konstruktionen sind bei komplexen Begründungen natürlich genauso möglich.

(43) Wegen [dessen chlorigen Geschmacks] und [der akuten Durchfallgefahr] [sollte man besser Trinkwasser kaufen] . 0

(44) Um [den Magen zu schonen] und [Durchfall zu vermeiden] [empfehle ich , das Leitungswasser nicht zu trinken] . 0

(45) [Ich würde davon nicht trinken] , weil [das Wasser verschmutzt ist] und [eklig schmeckt] . 0

Nicht in allen Fällen von komplexen Begründungen werden die Gründe oder Folgen durch *und* oder Interpunktion verknüpft. Neben Konjunktionen gibt es Disjunktionen und andere Konstruktionen, wie *sowohl...als auch*, *weder...noch* oder *nicht...sondern* etc.. Alle diese inhaltlichen Zusammenhänge zwischen den einer Begründung angehörigen Segmenten gehen in der hier beschriebenen Annotation ersteinmal verloren und müssen in einem folgenden Interpretationsschritt rekonstruiert werden. An dieser Stelle begnügen wir uns aber damit, alle diese Komplexe vereinfacht durch die oben beschriebene Verkettung von Gründen und Folgen zu markieren. Beispiele:

(46) Es gibt teilweise Aktionen mit den Kindern ...[die mir sehr gut gefallen haben] , weil [sie die Kinder so nicht von den Erwachsenen abgrenzen] , sondern [sie etwas in Mitten aller präsentieren] . 052

(47) [Es ist ein lustiges System] , denn [man bestellt sein Gericht nicht von der Karte] , sondern [hat die Beilagen bereits auf dem Tisch] und [das Fleisch wird von den Kellner im Tournus an grossen Spiessen angeboten] . 052

Man sollte sich nicht davon irritieren lassen, dass die Segmente im Satz so durcheinander aufgereiht sein können. Auch im folgenden Beispiel, wo die Folge zwischen zwei Gründen steht, folgt man den o.g. Regeln: Der erste Konnektor verweist auf den ersten Grund (dieser wiederum auf den Zweiten), außerdem auf die Folge mittendrin und auf den sekundären Konnektor:

(48) Sowohl wegen [der ruhigen Lage] [kann ich das Hotel empfehlen] , als auch weil [man dort behandelt wird wie ein König] . 0

Ähnlich hier:

(49) *Nicht* wegen [des schlechten Wetters] [bin ich wieder abgereist] , sondern weil [der Service einfach schlecht war] . 0

Ein Wort zur Distanz, die zwischen zwei verknüpften Segmenten liegt. Wenn eine komplexere Begründung nicht in einem Satz formuliert ist, sondern sich über mehrere Sätze verteilt, überbrückt die Verknüpfung natürlich jeweils eine Satzgrenze. Segmente, die über mehrere Satzgrenzen voneinander

entfernt sind, sollten allerdings nicht verknüpft werden, es sei denn, bei dem dazwischenliegenden Satz handelt es sich eindeutig um eine Parenthese oder einen Gedankeneinschub.

In seltenen Fällen kann es vorkommen, dass einem Konnektor wie *daher* oder *deswegen* kein einzelnes, in der Nähe liegendes Segment als Grund zugeordnet werden kann, sondern dass er sich stattdessen (im Sinne von *aus diesen o.g. Gründen*) auf ganze Paragraphen und die darin formulierten Erwägungen bezieht. Auch dann sollte der Annotator versuchen, in linker Nähe einen halbwegs passenden Kandidaten zu finden. Für diese speziellen Fälle sollte beim Konnektor die Eigenschaft `LARGE-SCOPE` gesetzt werden. Weiter greifende Verknüpfung von ganzen Gedankensträngen im Sinne einer Argumentationsstruktur soll an dieser Stelle nämlich *nicht* geleistet werden. Die Minimalität und Lokalität der annotierten Begründungen ist das angestrebte Ziel.

3.1 Zwei komplexere Beispiele

Abschließend für diesen Abschnitt soll anhand zweier komplexerer Beispiele noch einmal veranschaulicht werden, wie das angedachte Zusammenspiel von Konnektor, sekundären Konnektoren, Segmenten und Fokuspartikeln funktionieren soll.

(50) Ein Kurzbesuch in Berlin lohnt sich immer , stand auf meinem handlichen Reiseführer . Und wenn ich so auf meinen Kurz-Trip zurückschaue , [kann ich mich dem definitiv anschließen] . *Nicht nur* , weil [Berlin einfach eine beeindruckende Stadt ist] , wegen [der ganzen Sehenswürdigkeiten] oder weil [dort immer was los ist] , sondern *vor allem* wegen [des ganz speziellen Flairs dieser Stadt] . ∅

Im ersten Beispiel liegt eine Begründung vor, die erläutert, warum der Autor sich dem Urteil seines Reiseführers anschließt. Diese besteht aus fünf Segmenten, eines beschreibt die Folge, die anderen vier geben die Gründe an. Sie werden jeweils mit sekundären Konnektoren angegeben, sind aber alle Gründe ein und derselben Begründung. Die Grund-Segmente werden links-linear verkettet, genauso wie die entsprechenden sekundären Konnektoren. Die stehen teilweise sogar noch im Skopus einer Fokuspartikel. Der erste und deswegen zentrale Konnektor verweist auf seinen Fokuspartikel und auf die Ketten von Folgen, Gründen und sekundären Konnektoren. Das Verknüpfungsverhältnis wird noch einmal graphisch in Abbildung 2 dargestellt.

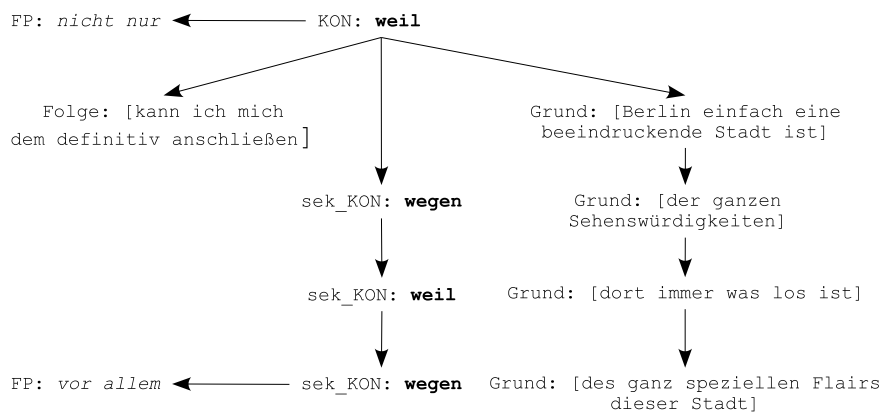


Abbildung 2: schematische Darstellung des Beispiels

Wiederum sei darauf hingewiesen, dass der inhaltliche Zusammenhang der *nicht nur ... sondern*-Konstruktion in dieser Annotation unberücksichtigt bleibt. Dies ist explizit vorgesehen. Die argumentative Struktur, in der die gegebenen Gründe oder Folgen verstanden werden können, wird in diesem Annotationsschritt ausgeblendet. Wichtig ist nur, dass die verschiedenen einzelnen Gründe im Kontext

einer Begründung identifiziert und durch die Verkettung assoziiert werden. Jede weitere Strukturierung kann in einem darauf aufbauenden Interpretationsschritt erfolgen.

Während im ersten Beispiel vor allem gezeigt werden sollte, wie mehr als zwei Segmente in einer Begründung zusammenspielen, soll im zweiten Beispiel die Entscheidung problematisiert werden, ob mehrere Konnektoren eben nun einer oder mehreren Begründungen angehören.

(51) Alles war voll mit Algen , Pferde- und Kamelkacke . Deshalb waren wir auch nur einmal am Strand , denn uns war es wirklich zu eckelhaft darin zu baden . 120

In diesem Textstück sind zwei kausale Konnektoren vormarkiert, *deshalb* und *denn*. Es gäbe unterschiedliche Möglichkeiten ihr Zusammenspiel zu verstehen. Eine Möglichkeit besteht darin, dass hier eine einzige Begründung vorliegt, nämlich die, warum der Autor nur einmal am Strand war, deren zentraler Konnektor *deshalb* ist, deren erster Grund die spezifische Unreinheit des Wassers, und deren zweiter Grund die Tatsache ist, dass in solchem Wasser nicht gebadet werden kann. Nach dieser Variante verstünde man den zweiten Konnektor *denn* als Sekundärkonnektor. Allerdings sollte man in solchen Situationen immer überprüfen, ob die Segmente (im akuten Fall die beiden Gründe) wirklich unabhängig voneinander zur Begründung beitragen, oder ob sie sich nicht auch gegenseitig stützen. Es stellt sich schnell heraus, dass die Verunreinigung des Wassers den Ekel der Urlauber stützt und somit müssen wir die These verwerfen, dass hier nur eine komplexere Begründung mit Sekundärkonnektor vorliegt, und stattdessen mit zwei Begründungen rechnen.

Die zweite Möglichkeit wäre also davon auszugehen, dass zwei Begründungen vorliegen, eine mit dem Konnektor *deshalb* und eine mit *denn*. Die mit dem Konnektor *deshalb* verknüpft die Unreinheit des Wassers mit der Tatsache nur einmal am Strand zu sein. Die mit dem Konnektor *denn* verknüpft andererseits die Feststellung, dass Baden in dreckigem Wasser eklig ist, mit dem Ergebnis nur einmal am Strand gewesen zu sein. In diesem Fall hätten beide Begründungen unterschiedliche Gründe, aber dieselbe Folge. Jedoch steht *denn* syntaktisch im Skopus von *deshalb*.

Die dritte Variante und die einzige, die letztlich in Frage kommt, ist also eine untergeordnete Begründung innerhalb einer allgemeineren. Eine allgemeinere Begründung:

(52) [Alles war voll mit Algen , Pferde- und Kamelkacke .] Deshalb [waren wir auch nur einmal am Strand , denn uns war es wirklich zu eckelhaft darin zu baden] . 120

und eine, in der ein bestimmter Teilaspekt noch elaboriert wird:

(53) [waren wir auch nur einmal am Strand] , denn [uns war es wirklich zu eckelhaft darin zu baden] 120

Es sollte deutlich geworden sein, dass die Entscheidung, welche Konnektoren welchen Begründungen angehören, d.h. wieviele eigenständige Begründungen vorliegen, nicht immer trivial ist, aber anhand einer überlegten Interpretation meist zufriedenstellend getroffen werden kann.

4 Klassifikation von Segmenten

4.1 Illokutionszuweisung

Ein wichtiger, aber nicht immer einfacher Teil der Annotation ist die Zuweisung einer Illokution jedes Segments. Das hierfür angenommene Inventar von Illokutionen wird in den entsprechenden Untersektionen erläutert. Danach sollen strittige Fälle unter die Lupe genommen werden.

4.1.1 report

Ein Segment sollte als Reportivum markiert werden, wenn es eine Aussage über Geschehnisse in der Welt, über den Zustand derselben oder über darin ausgeführte Handlungen trifft.

(54) [Zum Strand ging es durch Dünen über einen ca . 100m langen Holzsteg] , da
[sich das Hotel oberhalb einer Ausgrabungsstätte befand] . 250

(55) [Der Kellner brachte uns einen neuen Cocktail] , da ... ∅

4.1.2 report-author

Dieser Typ ist speziell für Reportiva vorbehalten, in denen sich der Author sprachlich in den Vordergrund der Beschreibung bringt, also typischerweise beim Vorliegen des Subjekts in der 1. Person singular oder plural. Das muss nicht nur der Fall sein, wenn der Author selbst der Handelnde ist, wie in folgenden Beispielen.

(56) Da ... , [machten wir uns zum ABENDESSEN auf] . 007

(57) [Spontan beschlossen wir aber doch , dieses Hotel zu wählen] , weil ... 007

sondern auch, wenn gar nicht eine Handlung sondern nur ein Geschehniss vorliegt,

(58) ... , weswegen [ich viel Flüssigkeit verloren hatte] . ∅

oder wenn der Autor sich bei der Beschreibung von Handlungen anderer z.B. durch eine Passiv-Konstruktion in den Vordergrund rückt.

(59) [Ich wurde von einem kleinen Hund gebissen] , weil ... ∅

4.1.3 ident

Wenn der Autor sich in einem Segment über seine eigenen Gefühle, Wünsche oder Grundwerte äußert, klassifizieren wir es als Identifikativum. Es geht dabei immer um den bewussten, geistigen Zustand des Autors.

(60) ... , denn [ich hatte große Angst vor dem Krokodil, welches am Strand gesehen worden ist] . ∅

(61) da ... , [freuten wir uns sehr über das Indoor-Schwimmbad des Hotels] . 242

(62) [Mir war es lieber, die Kinder um mich zu haben] , denn ... ∅

4.1.4 ident-judge

Segmente, in denen der Autor eine Wertung vornimmt und ausdrückt, dass er etwas gut oder schlecht findet, werden als Evaluativa klassifiziert. Das kann explizit durch Formulierungen wie

(63) [Besonders gut gefallen hat mir die Animation] , da ... ∅

(64) [Positiv zu bewerten ist , dass das Hotel über einen umzäunten Strand verfügt] , weil ... ∅

geschehen, aber auch durch Zuschreibungen, die ganz eindeutig wertende Konnotation haben. Entscheidend dabei ist, dass diese wertende Zuschreibung im Verb realisiert ist, also wie unten *sympathisch*

sein, einen Pluspunkt verdienen:

(65) [Die Anlage ist sympathisch] , da ... 246

(66) Schon allein durch ... [verdient es einen fetten Pluspunkt] . 245

Sollte ein Segment aber eine wertende Zuschreibung enthalten, die nicht vom Verb herrührt, dann soll das Segment nicht als Evaluativum klassifiziert werden, sondern gemäß der oberflächlich vorrangigen Illokution (siehe dazu auch den Abschnitt 4.1.10 mit Grenzfällen).

4.1.5 ident-assume

Behauptungen, die der Autor in seiner Formulierung als nicht bewiesen oder gesichert herausstellt, also vor allem Vermutungen und Prognosen, werden als Estimativa klassifiziert.

(67) [₁ Vielleicht hat sie] deswegen [₁ auch so gut geschlafen] . 242

(68) ... denn [anscheinend sind wir nicht die einzigen gewesen , denen es so erging] . \emptyset

(69) [Das soll wohl schon öfter vorgekommen sein] , da ... \emptyset

4.1.6 commit

Wenn sich der Autor in einem Segment dazu verpflichtet, etwas zu tun, klassifizieren wir dies als Kommissivum.

(70) ... Deswegen [wird dies auch nächstes Jahr wieder unser Wunschziel sein] . \emptyset

(71) [Nie mehr werde ich in diesem Pool baden gehen] , da ... \emptyset

4.1.7 direct

Wird der Leser vom Autor zu einer Handlung aufgefordert, wird das entsprechende Segment als Direktivum klassifiziert.

(72) [Unbedingt einen eigenen Fön mitbringen] , da ... 244

(73) ... [₁ Getränke] deshalb [₁ außerhalb kaufen] . 112

Auch Aufforderungen, die weniger als Befehl, eher höflich als Ratschlag oder Empfehlung formuliert sind, sollte als vom Typ DIRECT klassifiziert werden.

(74) ..., deshalb [empfehle ich nur im Restaurant zu essen] . 247

(75) [Am besten vor Ort Medikamente beschaffen] da [unsere nicht wirklich hilft] ! 030

(76) [Es empfiehlt sich ein Zimmer im oberen Stock zu nehmen] , da ... 244

4.1.8 present-hypothetic-situation

Handelt es sich bei der Begründung um die Angabe eines Zweckes, wird das Segment, welches den Zweck formuliert, als vom Typ PRESENT-HYPOTHETIC-SITUATION markiert.

(77) Um [uns zu erfrischen] [holten wir uns einen Cocktail] . \emptyset

(78) Zur [Erfrischung] [holten wir uns einen Cocktail] . \emptyset

(79) Damit [niemand eine trockene Kehle bekommt] , [stehen an jeder Ecke
Trinkwasserspender] . \emptyset

Eine Anmerkung zur Richtung des Kausalverhältnisses: Das Segment, in dem die zu erreichende Situation präsentiert wird, soll der Grund sein; das Segment, welches die Umsetzung formuliert, die Folge. Markiert werden soll also das Kausalverhältnis zwischen Zwecksetzung und Umsetzung, nicht aber das zwischen Umsetzung und Zielerfüllung.

In seltenen Fällen kann auch mit *wegen* eine hypothetische Situation ins Spiel gebracht werden, wie das folgende Beispiel zeigen soll:

(80) Wegen [eines Cocktails] [musste ich 20 Minuten anstehen] . \emptyset

Das Cocktail-Segment kann hier verstanden werden als Präsentation einer zu erreichenden und daher noch hypothetischen Situation (nämlich dem Erhalt eines Cocktails).

4.1.9 unknown

In seltenen Fällen kann es vorkommen, dass der Annotator sich nicht in der Lage sieht, dem Segment eine Illokution zuzuweisen, weil er z.B. aufgrund von verwirrenden oder ungrammatischen Formulierung des Autors, oder wegen widersprüchlicher Argumentation kein ausreichendes Verständnis für eine Interpretation erreicht. In diesen Fällen sollte das entsprechende Segment als vom Typ UNKNOWN markiert werden. Es sollte jedoch vorher genau überprüft werden, ob das Segment nicht noch seinem Surface-Value nach typisiert werden kann (siehe nächster Abschnitt).

4.1.10 Diskussion von Grenzfällen

Bei der Zuweisung von Illokutionen kann es sehr schnell zu strittigen Fällen kommen. Die hier angestrebte Lösung von Streitfällen richtet sich nach folgender Regel:

Entscheide im Zweifelsfall immer für die oberflächlich vorrangige Illokution.

Im strittigen Fall wird also die indirektere Illokution, das heißt die, die mehr Interpretation bedarf, verworfen.

Hier einige Beispiele: Häufig muss sich der Annotator zwischen REPORT und IDENT-JUDGE entscheiden:

(81) [Es gibt leider nicht mal Kerzen auf den Tischen] , weshalb ... 140

Durch das *leider* haben wir zwar eine Bewertung der Tatsache, daß es keine Kerzen gibt, aber die Hauptillokution scheint eher das Berichten dieses Umstandes zu sein. Im Zweifel ist nach unserer Regel die oberflächlich eindeutiger Illokution zu wählen, also aufgrund der Satzform und der Abwesenheit eines explizit wertenden Verbs der Typ REPORT.

(82) [Am besten einen Sonnenschirm mitnehmen] , da [diese sehr teuer dort
sind] . 243

Auch in Beispiel (82) ist mit dem Ausdruck *sehr teuer* zwar eine Zuschreibung mit negativ wertender Konnotation gegeben, aber es liegt insgesamt eher eine Beschreibung der Welt, als eine Bewertung vor. Deswegen muss auch hier REPORT gewählt werden. Ein ähnlicher Fall ist der folgende:

(83) [Die Treppe ist für Kleinkinder hochgefährlich] , da [teilweise nur an einer
Seite ein Geländer angebracht ist] 152

Man könnte behaupten, die Eigenschaft der Treppe, hochgefährlich zu sein, wäre negativ genug konnotiert, um das Segment als Bewertung zu verstehen. Allerdings ist sie das nicht von vornherein. Es könnte auch Kontexte geben, wo die Gefahr erstrebenswert und gut ist. Im Zweifelsfall gilt also auch hier: REPORT.

Häufig weiss man nicht, ob die Verallgemeinerungen des Autors gültige Beschreibungen der Welt sind, oder nur rhetorisch verpackt seine eigenen Annahmen oder Handlungen widerspiegeln. Auch hier sollte man sich klar an die oberflächlichen Merkmale der Segmente halten:

(84) da [man die meiste Zeit der Woche mit Tauchen verbringt] [bekommt man nicht viel mit] , und [will Abends oft nur ins Bett fallen] . 140

Zwar läßt sich bezweifeln, dass alle Touristen in diesem Hotel die meiste Zeit tauchen, aber nach der o.g. Regel muss dieses Segment –genauso wie die anderen im Beispiel– als REPORT klassifiziert werden, nicht als REPORT-AUTHOR oder IDENT-ASSUME. Schließlich wird bei der Annotation nicht über die Wahrheit der vom Autor aufgestellten Behauptungen geurteilt.

Ein anderes interessantes Beispiel:

(85) ... Daher [ist es gut ein Auto dabei zu haben] . \emptyset

Das Segment kann durchaus als Ratschlag verstanden werden. Der surface-value aber ist der einer Bewertung, das Hauptprädikat ist *gut sein*, weswegen als Typ hier nur IDENT-JUDGE infrage kommt. Würde das Beispiel stattdessen aber

(86) ... Daher [ist es empfehlenswert ein Auto dabei zu haben] . \emptyset

heißen, müßte es natürlich als vom Typ DIRECT annotiert werden.

4.2 Performative

Manchmal wird die Illokution eines Segmentes durch ein Performativ oder eine performative Formel explizit gemacht. Vorallem bei Aufforderungen und Ratschlägen, aber auch bei Vermutungen kommt das häufig vor.

(87) [Ich empfehle, Sandalen mitzunehmen] , weil [es sonst am Strand unangenehm werden kann] . \emptyset

(88) ... Daher [nehme ich an , dass der Weg von der Küche zum Speiseraum ungeheuer lang ist.] . \emptyset

Die Eigenschaft PERFORMATIVE soll aber nicht nur positiv markiert werden, wenn ein eindeutig performatives Verb vorliegt, sondern auch wenn Umschreibungen wie *ich kann empfehlen*, *es ist empfehlenswert* oder *es ist zu empfehlen* benutzt werden. Gleiches gilt für *raten/ratsam* etc..

(89) [Ich kann aber es jedem nur empfehlen] , denn [hier bekommt man wirklich einen traumhaften Urlaub für sein Geld geboten] . 037

(90) [Es empfiehlt sich ein Zimmer im oberen Stock zu nehmen] , da [es dort keine oder weniger Tiere (z.B . Geckos oder Frösche) im Zimmer hat] . 244

4.3 Meta-Bezüge

Für jedes Segment soll zusätzlich annotiert werden, ob es Meta-Bezüge enthält, oder nicht. Typischerweise wird in den Hotelreviews, da sie aus Bewertungsforen im WWW stammen, auf die Struktur oder die Handlungsmöglichkeiten im Forum verwiesen.

(91) [Jetzt gibt es endlich die Rubrik Hotel Miramar] und deswegen [poste ich das jetzt hier rein] . 248

Auch das Einfügen von URLs oder das Hinweisen auf die Begleithandlungen (eintippen) ist zu finden.

(92) Nun gut , [etwas spezifischer auf Hotel , Strand und Leute möchte ich an dieser Stelle nicht eingehen] , da [ich auf www.oberthal-online.de/hurghada.html bereits alles wichtige niedergeschrieben habe] und [mir ein doppeltes Eintippen ersparen möchte] . 029

Häufig trifft man auch Segmente, die sich auf die Texthandlung selber –das Bewerten eines Hotels– unabhängig vom Medium beziehen.

(93) Da [wir keine Kinder haben] , [kann ich den Mini Kids Club nicht bewerten] , ... 019

In allen diesen Beispielen sollte im entsprechenden Segment die Eigenschaft META positiv markiert werden.