

Annotationsrichtlinien IGGSA-STEPS Shared Task 2016

IGGSA

April 20, 2016

1 Überblick

Diese Annotationsrichtlinien beziehen sich auf die Annotation eines Korpus bestehend aus Reden der Schweizer Bundesversammlung. Sie stellen eine Überarbeitung der Annotationsrichtlinien dar, mit denen der Goldstandard für Shared Task IGGSA-STEPS 2014 annotiert wurde. Diese Überarbeitung soll als Grundlage für den Goldstandard der Shared Task IGGSA-STEPS 2016 dienen.

Bei der Shared Task wird es 3 Varianten geben:

- Full task
 - subjektive Ausdrücke entdecken
 - Source und Target finden
- Source-only: für gegebene subjektive Ausdrücke soll das System die entsprechenden Sources finden
- Target-only: für gegebene subjektive Ausdrücke soll das System die entsprechenden Targets finden

Die Daten liegen im TIGER-xml Format vor. Wir annotieren mit Salto . Ein Beispiel einer einfachen Annotation ist in Abbildung 1 zu sehen.

Weitere Beispielsannotationen sind im letzten Abschnitt dieses Dokuments zu sehen. Die Details der Annotation sind aber nur im Salto-Annotationstool richtig ersichtlich. Die Annotationsdatei, aus der die screenshots stammen, heißt `beispiel.xml` und sollte zusammen mit diesem Dokument zur Verfügung stehen.

2 Hauptziel der Annotation: subjective expressions

Prinzipiell sind **subjektive Ausdrücke** zusammen mit den Sources (Meinungsträgern) und Targets (Meinungsgegenständen), die zu ihnen gehören, zu annotieren.

Wir definieren diese als Ein- oder Mehrwortausdrücke, die Folgendes zum Ausdruck bringen:

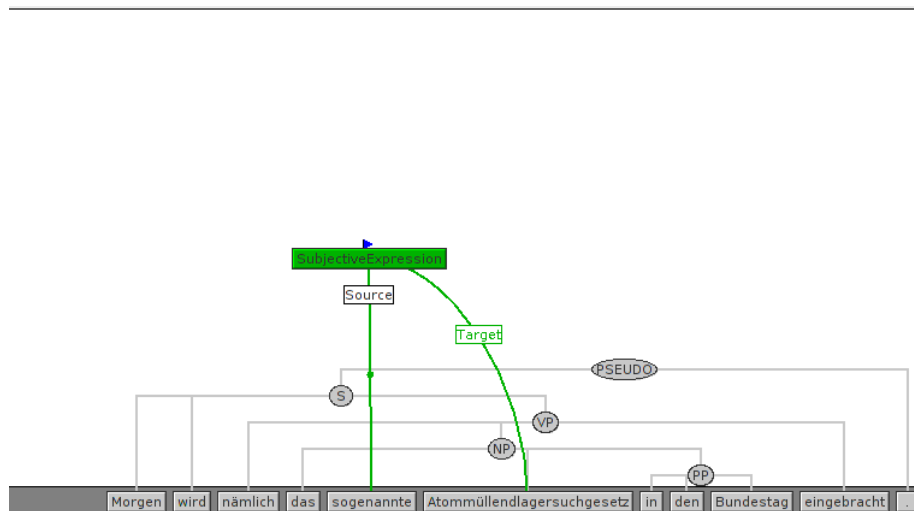


Figure 1: Annotationen eines Subjektiven Ausdrucks in Salto

- Wertungen (positiv oder negativ)
- (Un)Sicherheit
- Emphase / Nachdruck
- Sprechakte
- Mentale Vorgänge
- ...

Wir annotieren ausschließlich Subjektivität, die durch lexikalische Einheiten (sowohl Ein- als auch Mehrwortausdrücke) evoziert wird. (Satzzeichen werden nicht als Indikator von Subjektivität gewertet.)

Wir verwenden in Salto einen Annotationsframe vom Typ „SubjectiveExpression“ für subjektive Ausdrücke.

- Die einzelnen Untertypen werden von uns aber nicht unterschieden.
- Wir verlangen nicht, dass ein subjektiver Ausdruck eine Polarität haben muss.

Die Annotationsaufgabe selbst ist ebenfalls subjektiv und es wird Fälle geben, in denen sich der/die Annotator.in unsicher ist. Für diese Fälle gibt es die Möglichkeit, Aspekte der Annotation als unsicher zu markieren. So gibt es für den annotierten Frame eine Flag „Unsicher“ genauso wie für die beiden Elemente Source und Target.

Nachfolgend sind einige Beispielannotationen aufgeführt. Alle Beispiele, für die *keine* Quelle ausgewiesen ist, stammen aus einem Korpus von Bundestagsreden aus der Legislaturperiode 1994-1998.

Notation: In dieser Anleitung drucken wir subjektive Ausdrücke in Beispielssätzen fett. Die zu einem subjektiven Ausdruck gehörigen Elemente Source und Target werden in den Beispielen dadurch gekennzeichnet, dass eckige Klammern um die relevanten Phrasen/Spalten gesetzt werden. Einige Sonderfälle, die es zu beachten gilt:

- (1) Wenn die Source nicht lokal im (Teil)Satz zusammen mit dem subjektiven Ausdruck vorkommen, sie aber im vorhergehenden Text erwähnt werden, wie z.B. die Source in (8), dann wird die nächst liegende Vorerwähnung annotiert und mit einer Flag “Inferiert“ versehen. In Salto wird das Element Source oder Target dann mit einer Flag “Inferiert“ versehen. Die nächst liegende Vorerwähnung wird nur dann annotiert, wenn sie im selben Satz wie der subjektive Ausdruck vorkommt.
- (2) Wenn es sich bei der Quelle um den Sprecher einer Äußerung handelt, setzen wir in Salto auf dem Element Source eine Flag “Sprecher“; in dieser Anleitung werden die relevanten Fälle dadurch erfasst, dass wie in Beispiel (5) eine Annotation von der Form “[Source: Sprecher]” eingefügt wird.

Zu beachten ist, dass die Annotationen der Beispielssätze *nicht* den Anspruch haben, alle subjektiven Ausdrücke im Satz erschöpfend zu analysieren. Die Beispiele sollen üblicherweise nur die Behandlung des fett gedruckten subjektiven Ausdrucks und seiner Source und seines Targets illustrieren. Im Anhang findet sich aber die Analyse eines fortlaufenden Texts, in dem dann alle subjektiven Ausdrücke erschöpfend behandelt werden.¹

- (3) [Ich]_{Source} **bin gegen** [den Vorschlag]_{Target}. (konstruiert)
- (4) [Seine Untersuchung]_{Source} **sagt**, [daß bereits eine Veränderung um ein paar hundert Mark eine zweistellige Veränderung der projizierten Menge nach sich zieht]_{Target}.
- (5) **Leider** [ist sie nicht gekommen]_{Target}. [Source: Sprecher] (konstruiert)
- (6) Es ging darin u. a. um **interessante** [Fragen nach der Weiterentwicklung des Sozialstaates]_{Target} und um Deregulierung, Entbürokratisierung und grüne Mittelstandspolitik. [Source: Sprecher] (BUNDESTAG)
- (7) [Das Tun der Kandidaten]_{Target} ist **aller Ehren wert**. [Source: Sprecher] (adaptiert von http://www.nwzonline.de/oldenburg-kreis/lokalsport/tun-der-kandidaten-ist-aller-ehren-wert_a_1,0,549201959.html)
- (8) So sagte [er]_{Source:inferiert} gestern , daß [diejenigen , die über 55 Jahre alt sind , dies selbst bezahlen **sollen**]_{Target}.
- (9) ..., [es wird]_{Target} **zu Recht** [viel über die grundlegenden Werte für unsere Gemeinschaft nachgedacht und diskutiert]_{Target}. [Source: Sprecher]
- (10) [Es gibt dieses Engagement]_{Target} **Gott sei Dank** [in unglaublich reichhaltiger Form]_{Target}. [Source: Sprecher]

-
- (11) Im Gegensatz dazu [halten sich]_{Target} nach **Auffassung** [vieler Experten]_{Source} [insgesamt nicht mehr als 40000 vietnamesische Staatsbürger im Bundesgebiet auf]_{Target}.
- (12) [Wir]_{Source} **gehen** davon **aus**, [daß die derzeitige Dollarschwäche ein temporäres Phänomen ist]_{Target}.
- (13) Jetzt **gilt** es, [alle Kräfte anzuspinnen, ganz Deutschland fit zu machen für das nächste, das 21. Jahrhundert]_{Target}. [Source: Sprecher]

Wir annotieren Source oder Target nur, wenn sie im selben Satz wie ihr zugehöriger (evozierende) subjektive Ausdruck vorkommen.

Nicht zu den annotierbaren Ausdrücken gehören:

- Ausrufezeichen, z.B. bei Ausrufesätzen
- Rhetorische Stilmittel
 - Dazu zählen:
 - * Wiederholungen

(14) Und er **läuft und läuft und läuft**. (konstruiert)

(15) Ein Beschluss für Klimaschutz ist **an Deutschland gescheitert, an deutschen Abgeordneten, an Konservativen und Liberalen, die sich gegen einen ehrgeizigen Klimaschutz in Europa ausgesprochen haben**. (<http://dip21.bundestag.de/dip21/btp/17/17240.pdf>)
 - * Emphatisch geschriebene Wörter (Das hat 10 **MILLIONEN** Euro gekostet. (konstruiert))
 - * Rhetorische Fragen (**Und wer soll das bezahlen?** (konstruiert))

Da wir ausschließlich Ausdrücke annotieren, bei denen Subjektivität Teil der Wortbedeutung ist, werden sogenannte polar facts nicht annotiert (siehe folgenden Abschnitt 3).

Annotationstechnisch gilt folgendes: Wenn sich ein Mehrwortausdruck, den wir als subjektiven Ausdruck annotieren wollen, durch eine Phrase erfassen lässt, annotieren wir den entsprechenden Frame auf den nicht-terminalen Knoten. Wenn die zugehörigen Teile des Mehrwortausdrucks nicht durch einen Knoten abgedeckt sind, annotieren wir die minimale Menge an terminalen und nicht-terminalen Knoten, die nötig sind, um die relevante Spanne abzudecken. Dasselbe gilt auch für die Annotation der Elemente Source und Target.

Abschließend wollen wir festhalten, dass eine Annotation eines subjektiven Ausdrucks nur einer Source und einem Target zugeordnet werden kann. Dieser Punkt wird weiter diskutiert in Abschnitt 9.

3 Was sind polar facts?

Polar facts sollen im Rahmen unserer Shared Task nicht annotiert werden. Im folgenden werden wir polar facts von Subjective Expressions abgrenzen.

Der Grund, wieso diese Art von Sentiment von der Annotation ausgeschlossen wurde, ist, dass in unseren Pilotstudien vor allem bei diesem Sentiment eine nur sehr geringe Übereinstimmung zwischen unterschiedlichen Annotatoren erzielt wurde.

Wir unterscheiden zwei Hauptarten von polar facts:

Eine Gruppe von polar facts betreffen ganz neutrale Zustände, die typischerweise objektiv verifizierbar sind. Bei ihnen kann man sich sprachliche und außersprachliche Kontexte (also z.B. Weltwissen/Ideologien) vorstellen, in denen derselbe Ausdruck als Grundlage für eine Wertung mit gegenteiliger Polarität gemeint ist. Dies wird illustriert durch den Kontrast zwischen den möglichen Fortführungen a und b zu Beispiel 18.²

- (18) Im Auto ist viel Plastik verbaut. (konstruiert)
- a. Das ist toll, denn es spart Gewicht und damit Sprit.
 - b. Im Auto ist viel Plastik verbaut. Das wirkt sehr billig und unellegant.

Eine zweite Art von polar fact betrifft Ereignisse, von denen wir wissen, dass sie Teilnehmer positiv oder negativ betreffen. Betrachten wir dieses Beispiel, den ersten Absatz einer Meldung zu einem Verkehrsunfall (<http://www1.wdr.de/themen/panorama/unfallwuppertal100.html>). Zu beachten ist, dass in dem Text selbst nahezu keine explizite Wertung ausgedrückt wird. (“dramatisch” ist eventuell wertend, die Personen und Ereignisse werden aber sonst nicht direkt bewertet.)

1. **Der bei einem dramatischen Verkehrsunfall in Wuppertal schwer verletzte sechsjährige Junge ist außer Lebensgefahr.**
2. Das teilte die Polizei am Dienstag (04.09.2012) mit.
3. **Eine 80-Jährige hatte am Montag die Kontrolle über ihr Auto verloren und insgesamt elf Menschen angefahren.**

Hier spielt unser Weltwissen und die Empathie mit den Personen eine Rolle. Es ist klar, dass die elf angefahrenen Personen den Unfall negativ erlebt haben und wir Unfälle generell für negativ halten. Ebenso würden wir wohl sagen, dass das Außer-Lebensgefahr-sein positiv ist, da wir einem uns unbekannten 6-jährigen Jungen Gesundheit wünschen. Die Einstellung zu der 80-jährigen Fahrerin ist potentiell zwiespältig: wir könnten sie einerseits negativ beurteilen, wenn wir ihr Schuld am Unfall geben – sie hätte nicht fahren sollen oder die Kontrolle behalten müssen. Wir können aber auch positive Empathie für sie empfinden, weil ihr der Unfall wohl unabsichtlich passiert ist und wir uns vorstellen, dass sie selbst unter ihrer Schuld leidet.

² Im Gegensatz zu polar facts, kann bei echter, inhärenter Polarität nur Ironie oder explizite Negation oder ähnliches eine gegenteilige Polarität im Kontext bewirken:

- (16) Dieses **Genie** hat wieder das Auto nicht abgesperrt! (konstruiert)
- (17) Er ist **nicht der Klügste**. (konstruiert)

4 Prominenz/Salienz im Kontext

Wir behandeln Erwähnungen von subjektiven Zuständen, auch wenn diese pragmatisch nicht im Fokus stehen oder Details fehlen.

- Wir annotieren also Erwähnungen realer subjektiver Zustände, auch wenn lokal Source und/oder Target fehlen (“Diese **Prognose** war falsch”).
- Wir annotieren modalisierte Erwähnungen, also z.B. hypothetische und negierte Zustände.

(19) Möchten [Sie]_{Source} eine **Prognose abgeben**? (konstruiert)

(20) [Er]_{Source} hat keine **Prognose abgeben**. (konstruiert)

(21) Wenn [sie]_{Source} eine **Prognose abgibt**, liegt sie oft richtig. (konstruiert)

(22) Wenn [das]_{Target} wirklich [Ihre]_{Source} **Absicht** ist, Herr Minister, dann, so muß ich Ihnen hier sagen, ist das erneuter Wortbruch .

- Wir annotieren quantifizierte und generische Erwähnungen von subjektiven Zuständen.

(23) [Alle]_{Source} **wollen** [Maoam]_{Target}.

(24) [Niemand]_{Source} **will** [eine Mauer bauen]_{Target}.

(25) **Prognosen** sind oft falsch. (konstruiert)

Ebenso unterscheiden wir bei Ausdrücken, die (Un)sicherheit ausdrücken, nicht zwischen Fällen, in denen es sich um eine Expertenaussage handelt, und Fällen, bei denen es sich um Aussagen von Laien handelt.

- [Experten]_{Source} **schätzen**, [daß in jedem Jahr mehr als 15 Millionen Menschen auf der Erde an Hunger sterben]_{Target}. (konstruiert)
- [Ich]_{Source} **schätze**, [er kommt nicht mehr]_{Target}. (konstruiert)

Einfache Futurkonstruktionen betrachten wir als Tatsachen, wenn Sie als Vorhersagen über erwartbare zukünftige Ereignisse geäußert werden, auch wenn die Zukunft letztlich immer mit Unsicherheit behaftet ist. Wir würden folgende Beispiele also nicht annotieren:

(26) Thomas ist morgen in Bangladesh. (konstruiert)

(27) Die Gesetzesvorlage betrifft die Arbeitszeitregulungen für Krankenhauspersonal. Sie wird morgen verabschiedet. (konstruiert)

(28) Wir beginnen die Nachmittagssitzung um 13:00 Uhr. (konstruiert)

Von dem *werden*-Futur zu unterscheiden ist eine *werden*-Konstruktion, mit der man eine unsichere Einschätzung abgibt. Somit sind die folgenden werden-Konstruktionen subjektiv und sollen annotiert werden:

(29) [Sie **wird** nach Hause gegangen sein]_{Target}. [Source: Sprecher](konstruiert)

(30) [Ich]_{Source:inferiert} denke, [das **wird** ihr Bruder sein]_{Target}. (konstruiert)

5 Grenzen von annotierten Ausdrücken

5.1 Grenzen der subjektiven Ausdrücke

Wir klammern die detaillierte Behandlung von modifizierenden Ausdrücken, die subjektive Ausdrücke verstärken bzw. abschwächen (engl. *Intensifier*), und von Negation komplett aus und annotieren solche Ausdrücke auch nicht als Teil des subjektiven Ausdrucks.

- (31) [Das]_{Target} ist gar nicht **dumm**. [Source: Sprecher]

MEHRWORTAUSDRÜCKE (im weitesten Sinn) sind nicht immer klar von produktiven Bildungen zu unterscheiden. Wie wir Mehrwortausdrücke definieren und behandeln ist weiter beschrieben in Abschnitt 11.1.

REDEWENDUNGEN sind im typischen Fall komplette (Teil)Sätze, die keine / minimale Modifizierungen zulassen, bei denen Source und Target oft nicht ausdrückbar sind. Bei (33) und (34) ist zwar die Source genannt, aber es ist nur aus dem Vortext erschießbar, welchen konkreten Optionen der Spatz und die Taube entsprechen.

- (32) **Lieber den Spatz in der Hand als die Taube auf dem Dach.**
(z.B. <http://www.shz.de/nachrichten/lokales/stormarner-tageblatt/artikeldetails/artikel/begegnungen-von-mensch-zu-mensch.html>)
- (33) [Ihnen]_{Source} ist anscheinend **der Spatz in der Hand lieber, als die Taube auf dem Dach.** <http://www.np-coburg.de/lokal/lichtenfels/lichtenfels/Traum-Judentrasse-ist-geplatzt;art83428,2567997>
- (34) “[Uns]_{Source} ist **der Spatz in der Hand doch lieber, als die Taube auf dem Dach**”, meinte dazu Ortschaftsratsrat Thomas Bieger. <http://www.schwarzwaelder-bote.de/inhalt.haigerloch-lieber-den-spatz-in-der-hand-halten.bc548522-dbe1-4266-a01c-44ddc830de93.html>

Schwierigkeiten ergeben sich, wenn die (Teile der) Redewendungen wortspielerisch verändert werden. In solchen Fällen annotieren wir die erkennbaren Teile der Redewendung als eigenen subjektiven Ausdruck. Eventuell vorhandene Modifizierer der Redewendung, sofern sie nicht nur Intensivierer, sondern subjektiv sind, werden als eigene subjektive Ausdrücke annotiert.

- (35) Dass damit nicht die **Taube auf dem Dach**, sondern der finanziell verkraftbare **Spatz in der Hand** ausgewählt wurde, führte zu Turbulenzen und Negativschlagzeilen. <http://www.nzz.ch/meinung/kommentare/redlicher-einzelkaempfer-1.18080642>

Weitere klare Fälle sind echte IDIOME wie *aus der Hand fressen*. Hier gehören alle Teile zum subjektiven Ausdruck.

Bei STÜTZVERBGEFÜGEN behandeln wir die Stützverben als Teil des subjektiven Ausdrucks (41-43). Wir orientieren uns dabei an der Konvention von Wörterbüchern. Hier erscheinen im Regelfall ebenfalls diese Verben als Teil des Lexikoneintrags eines Mehrwortausdrucks.

Reflexivpronomina sind dann Teil des subjektiven Ausdrucks, wenn die Pronomina *echte* Reflexivität ausdrücken (36-37).

-
- (36) [Er]_{Source} **entschuldigt** sich. (konstruiert)
- (37) [Er]_{Target} **befindet sich auf dem Holzweg**. (konstruiert)

Da wir ausschließlich auf Wortebene annotieren, ist bei subjektiven Ausdrücken wie der Konstruktion *ist zu* (38) auch das ganze Verb zu annotieren, wenn das Partikel *zu* Teil desselben ist (39). Dadurch kann sich ergeben, dass der subjektive Ausdruck gleichzeitig auch Teil des Targets ist (39).

- (38) Im Berichtsjahr **ist** insbesondere [die Intensivierung der Beziehungen zu Italien]_{Target} **zu** [begrüssen]_{Target}.
- (39) Ferner **ist** [dieser Bericht in Bezug auf die aussenpolitische Gesamtsicht in eine ganze Reihe von komplementären Berichten **einzuordnen**]_{Target}.

5.2 Grenzen von Source und Target

Wenn Source oder Targets in einer Präpositionalphrase oder einem Komplementsatz vorkommen, annotieren wir die Präpositionen / Subordinatoren mit (41-44).

- (40) [Das]_{Target} war nicht so **gut**.
- (41) [Ich]_{Source} **habe Lust** [auf Tofu]_{Target}. (konstruiert)
- (42) [Ich]_{Source} **habe** keine **Lust** [auf Tofu]_{Target}. (konstruiert)
- (43) [Ich]_{Source} **habe** total **Lust**, [Tofu zu essen]_{Target}. (konstruiert)
- (44) [Er]_{Source} **versprach**, [daß er zurückkommen würde]_{Target}. (konstruiert)

Wie die Beispiele zeigen, achten wir bei der Annotation von Source und Target nach Möglichkeit darauf, diese Elemente entweder als valenzgeforderte, direkte syntaktische Abhängige des subjektiven Ausdrucks zu annotieren oder wenn es lexikalisch-syntaktische Kontexte gibt, die die „Entfernung“ von Source und „Target“ aus der lokalen Syntax erlauben, wie z.B. das sog. Raising- / Anhebungsverb *scheinen* in (45a).

- (45) scheinen
- a. [Sie]_{Source} scheint kein **Interesse** [an dem Buch]_{Target} zu **haben**. (konstruiert)
- b. [Sie **scheint** kein Interesse an dem Buch zu haben]_{Target}. [Source: Sprecher]

Wir annotieren ausschließlich auf Wortebene und nicht auf Subwortebene. Bei Komposita, die sowohl einen subjektiven Ausdruck darstellen als auch Information zu Source und/oder Target enthalten, annotieren wir Source und Target daher nicht auf dem Kompositum. Zum Beispiel ist das Substantiv **Steuerdebatte** ein subjektiver Ausdruck (evoziert durch *-debatte*) und gleichzeitig auch sein Target (es wird über *Steuern* debattiert), wir annotieren das gesamte Wort

aber nur als subjektiven Ausdruck. Parallel dazu würden wir bei **Staatstrauer** nur den subjektiven Ausdruck annotieren, aber keine Source.

Einen weiteren wichtigen Fall, den wir ähnlich behandeln, stellen die Pronominaladverbien *dafür* oder *dagegen* dar. Hier ist es möglich, das Element ‘da’ als anaphorischen Teil, der auf das Target verweist, zu behandeln. Aber weil wir wortintern keine Frame-Elemente annotieren, annotieren wir dieses Element nicht.

(46) [Die wenigsten Mitglieder]_{Source} **waren für** [den Vorschlag]_{Target}.

(47) [Die wenigsten Mitglieder]_{Source} **waren dafür**.

6 Adjektivische Sourcen

Bei nominalen subjektiven Ausdrücken kann es sein, dass die Source durch ein prämodifizierendes Adjektiv ausgedrückt wird.

(48) Widersinnigerweise nährten die Maßnahmen anscheinend die [deutsche]_{Source} **Begeisterung** [für amerikanische Portfolio-Investitionen]_{Target} ...

Evidenz dafür, dass das Adjektiv wirklich die Rolle der Source füllt ist, dass man kein weiteres Abhängiges hinzufügen kann, das die Rolle der Source füllen soll (*die deutsche Begeisterung der Banken für ...).

Allerdings gibt es Fälle wie (49), bei denen das Hinzufügen einer expliziten Source-Phrase möglich ist, obwohl ein Adjektiv selbst auch Aufschlüsse über die Source zu geben scheint. In solchen Fällen bevorzugen wir die nicht-adjektivischen Abhängigen als Source.

(49) Kissinger nannte zunächst bilaterale **Konsultationen** [zwischen den USA und der Bundesrepublik]_{Source}. Es scheint, als zeige die nationale **Begeisterung** [der Deutschen]_{Source} [für sich selbst]_{Target} bei Ausbruch eines Krieges, Wiedervereinigung oder Ausrichtung einer Fußballweltmeisterschaft keine Unterschiede.

7 Fehlen einer lokalen Source / eines lokalen Targets

Manche subjektive Ausdrücke bringen ihre Source nicht als syntaktisch Abhängiges mit. Oft handelt es sich bei einer fehlenden Source um eine zitierte Sprecherin:

(50) “**Furchtbar** sogar“, erinnert sich [sein Enkel]_{Source:inferiert}.

Zu beachten ist, dass, falls der Source-Referent mehrmals im Gesamtsatz erwähnt wird, die Phrase als inferierte Source annotiert werden soll, die in unmittelbarer syntaktischen Beziehung zu dem dem Prädikat steht, das den relevanten subjektiven Ausdruck einbettet. In (51) soll daher das Subjekt *er* von *sagte* annotiert werden statt zum Beispiel der Phrase *der Junge* oder des *er*, welches das Subjekt von *verschwand* ist.

(51) Der Junge ging weiter, doch bevor er endgültig verschwand, sagte [er]_{Source:inferiert} noch: “[Das **muss** wohl noch warten]_{Target}.”

Wenn die relevante Source die Textverfasserin ist, dann versehen wir das Element Source mit einer Flag „Sprecher“.

(52) [Das]_{Target} war **doof**. [Source: Sprecher] (konstruiert)

Viele subjektive Ausdrücke bringen ihre Source als syntaktisch Abhängiges mit. Dieses kann aber durch verschiedene syntaktische Prozesse unterdrückt werden und dann nur im Kontext aufzufinden sein. Betrachten wir als Beispiel die Äußerung in (53). Hier fehlt das Subjekt von *sagen* aufgrund des Imperativs.

Die Imperativkonstruktion selber spiegelt einen subjektiven Zustand der Sprecherin wieder, wie in (53a) gezeigt. Zu beachten ist, dass wir hier nicht einfach den Sprechernamen *Diller* annotieren, der im Transkript am Anfang der Zeile steht. Wir betrachten diesen als Metadatum und nicht als Teil des Textes.

Das vom Sprecher Karl Diller gewünschte Ereignis, dass Kurt etwas sagen soll, beinhaltet nun selbst einen Sprechakt. Dementsprechend muss dessen Subjektivität ebenfalls erfasst werden.³ In diesem Fall kann man das unausgedrückte Subjekt von *sag* im Vokativ genannt finden. Dieser wird auch annotiert, aber die Annotation wird mit einer Flag „inferiert“ gekennzeichnet.

(53)

- a. Karl Diller [SPD] : Ja , Kurt , [jetzt **sag** mal was zu den Westhilfen]_{Target} ! [Source: Sprecher]
- b. Karl Diller [SPD] : Ja , [Kurt]_{Source:inferiert} , jetzt **sag** mal [was zu den Westhilfen]_{Target} !

Wenn wir die Beispiele in (53) leicht abwandeln und eine Imperativkonstruktion mit einem expliziten Subjekt (*Du*) verwenden, würde sich die Annotation wie in (54) darstellen.

(54)

- a. Ja , Kurt , [jetzt **sag** Du mal was zu den Westhilfen]_{Target} ! [Source: Sprecher] (konstruiert)
- b. Karl Diller [SPD] : Ja , Kurt , jetzt **sag** [Du]_{Source} mal [was zu den Westhilfen]_{Target} !

8 Einbettung von Sourcen

Anders als z.B. das bekannte Schema des MPQA-Korpus modellieren wir die Einbettung von Sources nicht direkt. Z.B. notieren wir nicht, dass die Meinung von Doris nur über den Umweg von Peter (und dem Sprecher) vorliegt.

(55) Peter sagt, Doris hält Herrn Müller für ein Genie. (konstruiert)

³Wenn der Imperativ ‘Finden Sie sich um 14:00 Uhr zur Wiederaufnahme der Sitzung wieder hier ein’ lautete, wäre darin kein subjektiver Zustand der Adressaten enthalten und es würde folglich keine ‘b’-Annotation geben, sondern nur die ‘a’-Annotation für die Imperativkonstruktion mit dem Sprecher als Source.

Wir annotieren stattdessen alle subjektiven Ausdrücke unabhängig von einander. Einbettungen sind nur indirekt sichtbar, dadurch dass ein subjektiver Ausdruck im Target eines anderen subjektiven Ausdrucks liegt. Z.B. ist (57) in (56) eingebettet und (58) unmittelbar in (57) und indirekt in (56).

(56) [Peter]_{Source} **sagt**, [Doris hält Herrn Müller für ein Genie]_{Target}.

(57) Peter sagt, [Doris]_{Source} **hält** [Herrn Müller für ein Genie]_{Target}.

(58) Peter sagt, [Doris]_{Source:inferiert} hält [Herrn Müller]_{Target} für ein **Genie**.

9 Mehrfachannotation desselben Ausdrucks

Wir lassen Mehrfachannotation zu, sowohl für die Fälle echter Mehrebenen-subjektivität (*prahlen*, *angeben mit*) wie auch für Fälle wie *beschuldigen*, wo entweder eine Person oder ihre Handlung als Target verstanden werden kann. Ferner drücken einige Ausdrücke symmetrische Beziehungen aus, wie z.B. *zusammenarbeiten*.

(59) prahlen

- a. [Kim]_{Source} **prahlt** [mit ihrem Porsche]_{Target}. (konstruiert)
- b. [Kim]_{Target} **prahlt** mit ihrem Porsche. [Source: Sprecher]

(60) beschuldigen

- a. [Er]_{Source} **beschuldigte** [mich]_{Target} der Lüge. (konstruiert)
- b. [Er]_{Source} **beschuldigte** mich [der Lüge]_{Target}.

(61) zustimmen

- a. [Ich]_{Source} **stimme** Ihnen **zu** : [Wir hier machen die Gesetze, und die Gerichte sollen sie anwenden]_{Target} .
- b. Ich **stimme** [Ihnen]_{Source} **zu** : [Wir hier machen die Gesetze, und die Gerichte sollen sie anwenden]_{Target} .

(62) zusammenarbeiten

- a. [Die Regierung]_{Source} **arbeitet** [mit der Opposition]_{Target} **zusammen**. (konstruiert)
- b. [Die Regierung]_{Target} **arbeitet** [mit der Opposition]_{Source} **zusammen**. (konstruiert)

10 Zur Verwendung des “Sprecher”-Flags

Der Sprecherflag wird **nicht** dazu verwendet, um anzuzeigen, dass ein bestimmter subjektiver Ausdruck im vorliegenden Satz keine explizite Source besitzt.⁴

In den meisten Fällen kommt der “Sprecher”-Flag für Prädikate infrage, bei denen das Sentiment einer Person ausgedrückt wird, die nicht am Ereignis, das das Prädikat beschreibt, beteiligt ist. Infolgedessen ist die Source auch nicht Bestandteil des Valenzrahmens des Prädikats. Diese Art von Subjektivität wird in der Fachliteratur auch *expressive subjectivity* oder *speaker subjectivity* genannt (64). Typische Beispiele sind *betrügen*, *beschönigen*, *schaden*, *täuschen*, *verdienen*. Jedoch kann in solchen Fällen auch eine explizite Source vorhanden sein, wie in (65). Dann sollte diese (explizite) Source annotiert werden; der “Sprecher”-Flag entfällt.

- (63) [Er]_{Target} wurde **kritisiert**. (konstruiert)
- (64) [Er]_{Target} ist ein **Betrüger**. [Source: Sprecher] (konstruiert)
- (65) [Für die Öffentlichkeit]_{Source:inferiert} ist [er]_{Target} ein **Betrüger**. (konstruiert)

11 Spezielle Phänomene

11.1 Mehrwortausdrücke

Uns interessieren generell Mehrwortausdrücke, die semantisch und/oder pragmatisch keine einfachen Bildungen sind, und die aber häufig vorkommende, etablierte Formulierungen sind. Dazu gehören echte Idiome (66), aber auch die häufig gebrauchten Stützverbgefüge (67).

Generell versuchen wir, den subjektiven Frame nur auf den minimalen Bestandteilen des MWE zu evozieren. Entsprechend lassen wir in z.B. (68) das optionale Adjektiv “verschärfter” unannotiert und in 69 den Relativsatz.

- (66) Also wer das damals erfunden hat , [der]_{Source} hat wohl **einen an der Waffel gehabt** .
- (67) **Hat** [jemand]_{Source} eine **Meinung** [dazu]_{Target}?
- (68) [Das Segment der zuletzt heftig herunter geprügelten Goldminenaktion]_{Target} **steht** [bei Anlegern]_{Source} seit geraumer Zeit **unter verschärfter Beobachtung**.
- (69) [Er]_{Source} **hat** eine **Meinung**, die sonst niemand teilt.

⁴Wenn dies die Funktion des Flags wäre, so wäre das Setzen dieses Flags redundant – es könnte dann genauso gut automatisch ermittelt werden, ob für einen subjektiven Ausdruck eine Source zugewiesen wurde oder nicht. Der Flag soll ausdrücken, dass für den subjektiven Ausdruck der Autor die Source ist. Es kann also durchaus Sätze geben, in denen ein subjektiver Ausdruck keine explizite Source besitzt, aber dennoch nicht der Sprecher die Source ist. In (63) wäre zum Beispiel der “Sprecher”-Flag nicht zu verwenden. Es ist zwar keine explizite Source vorhanden, aber wer tatsächlich *kritisiert* wird nicht spezifiziert.

Viele Stützverbgefüge kommen in mehreren Varianten vor, die sich durch Aspekt oder Register unterscheiden. Auch wenn dies bedeutet, dass nicht immer das gleiche Verb ‘stützt’, gehören für uns die entsprechenden Verben doch jeweils mit zum Stützverbgefüge dazu. Wir annotieren also wie folgt:

- (70) [Peter]_{Source} **hat Angst** [vor Hunden]_{Target}.
 (71) [Susanne]_{Source} **bekam Angst** [vor Spinnen]_{Target}.
 (72) Und ich wollte weg, [ich _{Source}] **kriegte Angst** [vor ihm]_{Target}.
 (73) [Peter]_{Source} **verlor seine Angst** [vor Hunden]_{Target}.

Die Syntax des Deutschen lässt es zu, dass Stützverben nicht als finite Verben sondern als Prämodifizierer vorkommen. Wir annotieren diese dann dennoch als Stützverbfälle.

- (74) Der 1949 von Haifa aus **unternommene Versuch** [Josef S.s]_{Source}, [die Rückstellung der beiden Fabriken zu bewirken]_{Target}, blieb erfolglos
 (75) Der [von anderen Forschern]_{Source} **unternommene Versuch**, [eine Grammatik “der deutschen Jugendsprache” zu schreiben]_{Target}, erscheint aufgrund der Vermischung der unterschiedlichen Sprechergruppen, Kommunikationssituationen und Textsorten als relativ problematisch.

Ein Fall, der besondere Erwähnung verdient, ist das Verb *sein*. Hierzu halten wir Folgendes fest:

- Wir betrachten produktive Kombinationen von sein (oder werden) mit Adjektiven (76) nicht als Fälle von Stützvergebrauch.
- Ebenso wenig annotieren wir Verwendungen der Kopula mit einem nominalen subjektiven Ausdruck, wenn die Kopula dazu dient, das Target des subjektiven Ausdrucks grammatisch mit dem subjektiven Ausdruck zu verbinden (77-78). Wenn die Kopula aber Source und subjektiven Ausdruck verbindet, dann behandeln wir sie als Teil des Mehrwortausdrucks, weil diese Fälle lexikalisch idiosynkratisch sind.⁵

- (76) [Mein Onkel]_{Source} war **wütend/sauer/übergücklich/enttäuscht**.
 (77) [Meine]_{Source} **Meinung** ist es, [dass die Filme heutzutage viel zu aggressiv sind und viel zu wenig Spannung vorkommt]_{Target}.
 (78) [Das]_{Target} ist nur [Ihre]_{Source} **Meinung**.
 (79) [Ich]_{Source} **bin der Meinung/Ansicht**, [dass wir diese Politik beenden müssen]_{Target}.

Schließlich haben viele Nomen, wenn sie in Stützverbgefügen vorkommen, eine Präferenz für einen bestimmten Artikel. So kommt ‘Versuch’, wenn es als Teil des Gefüges mit ‘unternehmen’ verwendet wird, in den meisten Fällen mit dem definiten Artikel vor. Es gibt jedoch auch viele andere Varianten, allein schon durch Effekte von Quantifizierung (81) oder Verneinung (82).

⁵Z.B. gibt es “Unser Ziel ist, dass ...” aber nicht “Wir sind des Ziels, dass ...”.

-
- (80) Auf der einen Seite soll der **Versuch unternommen** werden, das theologische Profil des Konzils zu charakterisieren
- (81) Es sollte jeder **Versuch unternommen** werden , die Familie direkt und schnell zu informieren.
- (82) Die Regierung der RS hatte zunächst keinen **Versuch unternommen** , das Schicksal der Vermissten zu klären . . .
- (83) Die Projektingenieure des amerikanischen Marsprogrammes wollen erneut einen **Versuch unternehmen** , mit dem Polar Lander in Kontakt zu treten.

Wir lassen Artikel bei der Annotation von Stützverbgefügen unannotiert.

11.2 Adverbien als subjektive Ausdrücke

Wir annotieren auch Adverbien als subjektive Ausdrücke. Dabei unterscheiden wir zwei Fälle.

Bei den ‘normalen’ Adverbien, die tatsächlich ein Verb modifizieren, annotieren wir nur das Verb als Target (84). Bei Satzadverbien, die sich auf eine ganze Aussage beziehen, annotieren wir die Spanne des relevanten Satzes oder Teilsatzes als Target. Wir nehmen das Satzadverb dabei immer mit in die Targetspanne hinein, auch wenn es am Rand des Teilsatzes steht wie in (86).

- (84) Peter [zeichnet]_{Target} sehr **gut**.
- (85) [Peter kam gestern **leider** nicht]_{Target}.
- (86) [**Leider** kam Peter gestern nicht]_{Target}.

11.3 Anreden

Auch wenn Sie formelhaft gebraucht werden, annotieren wir Anreden, die wertende Adjektive enthalten.

- Die Wähler, **werte** [Kolleginnen und Kollegen]_{Target}, haben erneut der Koalition der Mitte die Mehrheit im Deutschen Bundestag übertragen . [Source: Sprecher]
- Meine sehr **verehrten** [Damen und Herren]_{Target}, **liebe** [Kolleginnen und Kollegen]_{Target} , es wird zu Recht viel über die grundlegenden Werte für unsere Gemeinschaft nachgedacht und diskutiert . (BUNDESTAG)

11.4 Aufforderungssätze

Da Aufforderungssätze Wünsche des Sprechers/der Sprecherin ausdrücken, auch wenn Sie vielleicht routinemäßig vorgebracht werden und keinen „Herzenswünsche“ sind, annotieren wir sie. Wir annotieren bei direkten Aufforderungen die relevante Verbform (bei formellen Imperativen auch die Anrede *Sie*) (vgl. 87-89). Bei indirekten Aufforderungen wie z.B. 90 annotieren wir das Sprechaktverb als subjektiven Ausdruck und seine Argumente als Source und Target.

-
- (87) [Stellen Sie dem Entschädigungsfonds die Hunderte von Millionen DM zur Verfügung, die Sie für eine solche Katastrophe zurückgelegt haben, und zwar sofort]_{Target} ! [Source: Sprecher]
- (88) [Lassen Sie uns in die differenzierte Einzeldiskussion eintreten]_{Target}. [Source: Sprecher]
- (89) [Versuchen Sie erst einmal vor der eigenen Tür zu kehren]_{Target} !
- (90) [Ich]_{Source} darf Sie **bitten**, liebe Kolleginnen und Kollegen, [wieder Platz zu nehmen]_{Target}.

Bestimmte idiomatischen Fälle behandeln wir aber separat (falls sie überhaupt vorkommen).

- (91) Ach **scheren** [Sie]_{Target} **sich doch zum Teufel!** (konstruiert)

11.5 Modalverben

Wir annotieren sowohl deontische Modalverben, die das Bestehen einer Verpflichtung / eines Verbots anzeigen (cf. 92–105), als auch epistemische, die eine Schlussfolgerung markieren (cf. 96–97).

- (92) [Ich]_{Source:inferiert} denke, [das **müssen** wir sorgfältig abwägen]_{Target}.
- (93) [Wir **sollten** hier nicht den Fehler machen, alles in einen Topf zu werfen]_{Target}. [Source: Sprecher]
- (94) [Die Kommunalen Wohnungsgesellschaften **dürfen** nicht auf den Altschulden für abgerissene Bauten sitzen bleiben]_{Target}. [Source: Sprecher]
- (95) [Morgen **brauchen** Sie nicht zu kommen.]_{Target} [Source: Sprecher]
- (96) [Max Strauß **muss** es gewusst haben]_{Target}. [Source: Sprecher] (<http://forum.mormonentum.de/3764.html>)
- (97) [Er **dürfte** nun zu Hause sein]_{Target}. [Source: Sprecher]

Bei *können* unterscheiden wir so gut wie möglich Fähigkeitsfälle wie 98, die wir nicht annotieren, von Möglichkeitsfällen wie 99.

- (98) Er **kann** Ski fahren.
- (99) [Es **könnte** heute noch regnen]_{Target}. [Source:Sprecher]

Wir annotieren auch Instanzen des Verbs *wollen*. Neben regulären Vorkommen von *wollen* wie in 100 annotieren wir auch Fälle wie den in 101, wo die Konstruktion mit dem Verb bedeutet, dass jemand eine Behauptung aufstellt. Wir annotieren nicht Verwendungen mit unbelebten Subjekten wie 102–103, in denen kommuniziert wird, dass sich trotz versuchen ein bestimmtes Ergebnis nicht einstellt. (Negation ist hier nötig) Anders gelagert sind Fälle wie 104, die Ausdrücken, dass etwas passieren soll, also eine Art Verpflichtung ausdrücken.

- (100) [Er]_{Source} **wollte** [gehen]_{Target}.

-
- (101) [Peter]_{Source} **will** [es gewesen sein]_{Target}.
- (102) Es **will** nicht klappen :(
- (103) Ich fummele am Türgriff herum, aber die Tür **will** nicht aufgehen.
- (104) der standort der melonenbirne sollte hell und sonnig sein. [sie **will** oft gegossen werden]_{Target} [Source:Sprecher] und verzeiht nicht so gern uerreperiode

Ausser den einfachen Modalverben gibt es auch andere Modalausdrücke und -konstruktionen, die wir erfassen möchten. Z.B. hat *sein*+VP_{zu} eine deontische Bedeutung. Und umgangssprachlich wird auch *gehören* (mit oder ohne Reflexivpronomen) in einer deontischen Bedeutung verwendet.

- (105) [In solchen Fällen **sind** die Eltern schnellstmöglich **zu** informieren]_{Target}.
[Source:Sprecher]
- (106) [Cannabis **gehört sich** in Fachgeschäfte mit Alterskontrollen , nicht auf den Schulhof]_{Target} ! [Source:Sprecher]

11.6 Verben der mentalen oder sprachlichen Kategorisierung

Wir annotieren diese Prädikate immer. Wenn sich im Target eines solchen Prädikates ein weiterer klarer subjektiver Ausdruck befindet, annotieren wir diesen auch (cf. 108, 110).

- (107) Einen formalisierten Bericht über Schlußfolgerungen aus diesen Krawallen hat die Bundesregierung hingegen weder angekündigt, noch **hält** [sie]_{Source} [ihn für angezeigt]_{Target}.
- (108) Einen formalisierten Bericht über Schlußfolgerungen aus diesen Krawallen hat die Bundesregierung hingegen weder angekündigt, noch hält [sie]_{Source:inferiert} [ihn]_{Target} für **angezeigt** .
- (109) Mit dem Eintritt der Volljährigkeit sollen sie sich selbst entscheiden, ob sie auf Dauer die deutsche Staatsangehörigkeit erwerben oder die Staatsangehörigkeit ihrer Eltern behalten wollen. [Das]_{Target} **bezeichnen** [Sie]_{Source} hier [als Lächerlichkeit]_{Target}.
- (110) Mit dem Eintritt der Volljährigkeit sollen sie sich selbst entscheiden, ob sie auf Dauer die deutsche Staatsangehörigkeit erwerben oder die Staatsangehörigkeit ihrer Eltern behalten wollen . [Das]_{Target} bezeichnen [Sie]_{Source:inferiert} hier als **Lächerlichkeit** .
- (111) Das heißt: Der Ministerpräsident von Nordrhein-Westfalen, Herr Rau, kündigt genau das an, was ich sage . Wenn ich es sage, **kritisieren** [Sie]_{Source} [es als Ankündigung]_{Target}.
- (112) [Uns daraus jetzt einen Vorwurf zu machen]_{Target}, **halte** [ich]_{Source} [für schlechterdings unverschämt]_{Target}, Herr Hofmann .
- (113) [Uns daraus jetzt einen Vorwurf zu machen]_{Target}, halte [ich]_{Source:inferiert} für schlechterdings **unverschämt**, Herr Hofmann .

11.7 Appositionen

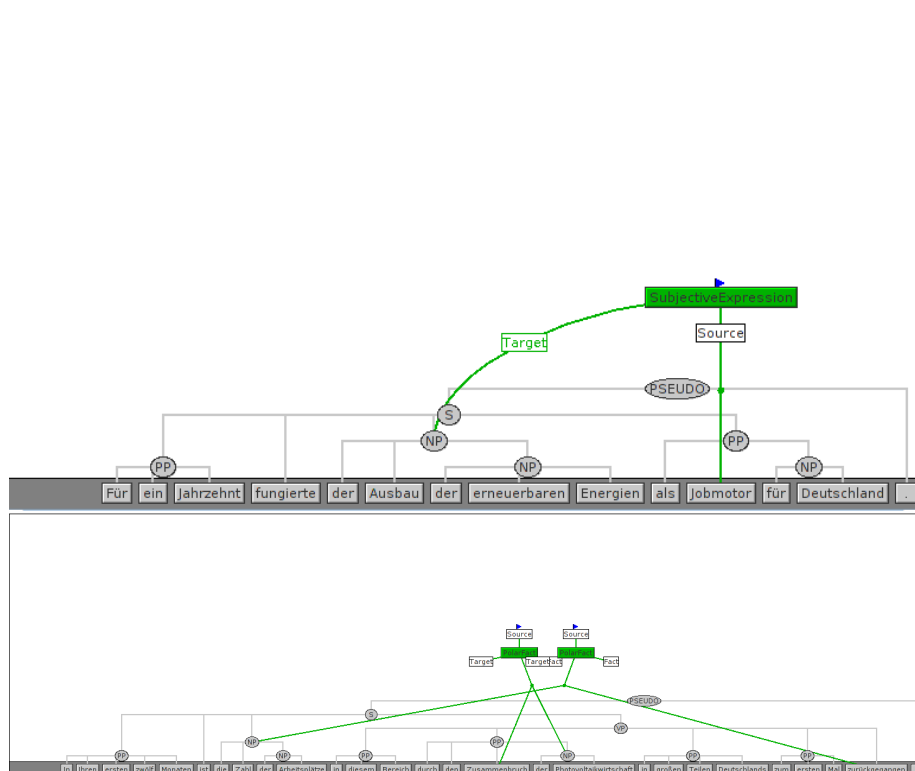
Wenn eine Nominalphrase der Appositionskonstruktion wertend ist und die andere das Target, dann annotieren wir die Teile entsprechend.

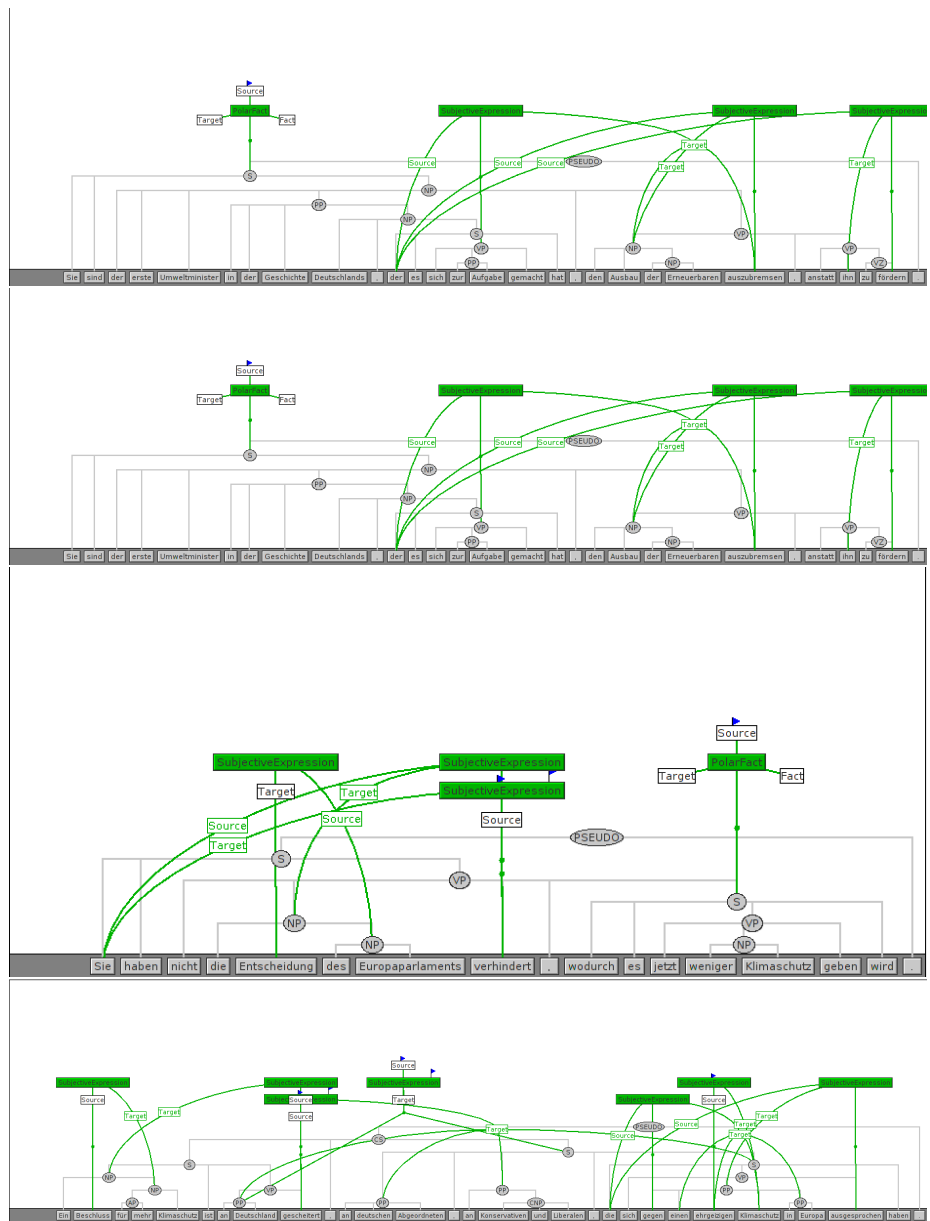
- (114) Arbeitslosigkeit zerstört jeden Tag aufs neue das **Wichtigste**, das Menschen besitzen, [nämlich ihre berufliche Qualifikation und Fähigkeit]_{Target}.
[Source: Sprecher]

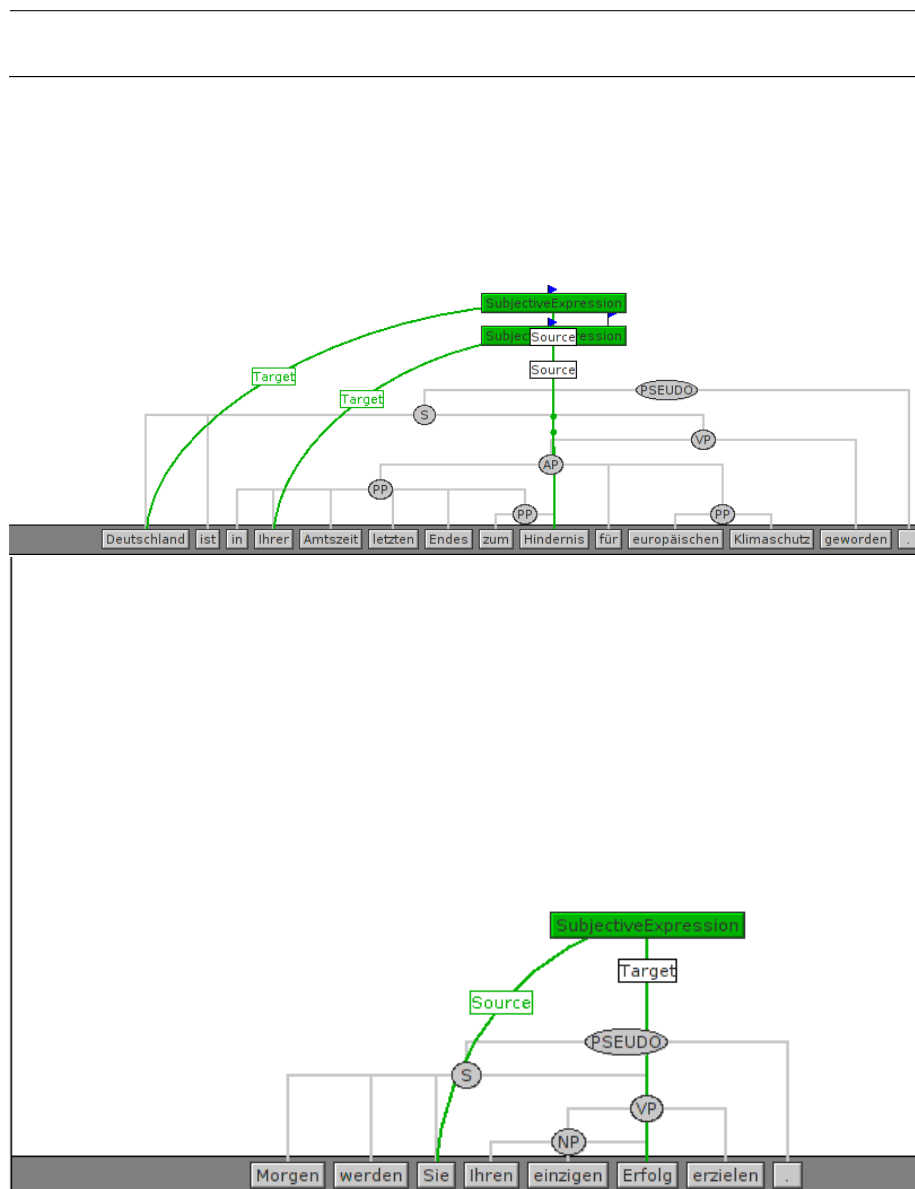
11.8 Spezifisch schweizerdeutsche Ausdrücke

Obwohl wir durch die Auswahl der Themen versucht haben den Ausdruck varietätenspezifischer Ausdrücke möglichst gering zu halten, werden mit Sicherheit immer wieder schweizerdeutsche Ausdrücke in den Daten vorkommen, die den Annotator_innen nicht vertraut sind. In diesen Fällen sind diese gehalten, diese Ausdrücke nachzuschlagen, um sie korrekt einschätzen zu können. Des Weiteren sollen die Annotator_innen die relevanten subjektiven Ausdrücke dann mit einer Flag „SchweizerDeutsch“ markieren.

12 Screenshots der Beispielsannotation eines zusammenhängenden Textabschnittes







Index

- Anreden, 14
- Appositionen, 16
- Aufforderungssätze, 14
- Ausrufezeichen, 4
- Beispiele, 2
 - eckige Klammern, 3
 - Fettdruck, 3
 - Format der, 3
- Emphatische Schreibung, 4
- Flag
 - FE
 - Inferiert, 3
 - Sprecher, 10
 - Unsicher, 2
 - Frame
 - Unsicher, 2
- Futur, 6
- Idiome, 7
- Imperativ, 10, 14
- Intensität, 7
- Ironie, 5
- Komposita, 9
- Mehrfachannotation, 11
- Mehrwortausdrücke, 7
- Modalverben, 15
 - deontisch, 15
 - epistemisch, 6, 15
- Polarität, 7
- Präpositionen, 8
- Pronominaladverbien, 9
- Redewendungen, 7
- Rhetorische Fragen, 4
- Salienz subjektiver Ausdrücke, 6
- Salto, 1
- Schweizer Bundesversammlung, 1
- Schweizerdeutsch, 17
- Source, 1
 - Einbettungen, 10
 - Fehlen einer, 9
 - Grenzen, 8
 - Sprecher als, 10
- Stützverbgefüge, 7
- SubjectiveExpression, 2
 - Grenzen, 7
- Subjektive Zustände
 - generisch, 6
 - hypothetische, 6
 - Laien vs. Experten, 6
 - negierte, 6
 - quantifiziert, 6
- Target, 1
 - Fehlen eines, 9
 - Grenzen, 8
- Verben der mentalen und sprachlichen Kategorisierung, 16
- werden
 - epistemisch, 6
 - Futur, 6
- Wiederholungen, 4