

**ANNOTATION UND ANALYSE ANAPHORISCHER
BINDUNG IMPLIZITER SEMANTISCHER
ROLLEN IN WEGINSTRUKTIONEN**

EINGEREICHT BEI: **PROF. DR. ANETTE FRANK**

INSTITUT FÜR COMPUTERLINGUISTIK
NEUPHILOLOGISCHE FAKULTÄT
RUPRECHT-KARLS-UNIVERSITÄT HEIDELBERG

EINGEREICHT AM: **21. JANUAR 2011**

EINGEREICHT VON: **MARIA SEMENCHUK**

Inhaltsverzeichnis

Abkürzungsverzeichnis	1
Einführung.....	2
1. Theoretische Grundlagen	4
1.1. Frame-Semantik	4
1.2. FrameNet.....	5
1.3. Null Instantiation.....	5
1.3.1. Interpretation	6
1.3.2. Licensor.....	7
1.4. Anaphern	7
2. Korpora.....	8
2.1. SemEval	9
2.1.1. Texteigenschaften.....	9
2.1.2 Annotation in SemEval	10
2.2. Wegbeschreibungskorpus.....	11
2.2.1. Texteigenschaften.....	11
2.2.2 Annotation im Wegbeschreibungskorpus	13
3. NI-Annotation im Wegbeschreibungskorpus	14
3.1. Annotationstool	15
3.2. Frame-Annotation	15
3.3. Ablauf der NI-Annotation	16
3.4. Annotationskriterien.....	17
3.5. Fehleranalyse bei der NI-Teilannotation im SemEval-Korpus	19
4. Auswertung der NI-Annotation im Wegbeschreibungskorpus	21
4.1. Parameter der Datenanalyse	21
4.1.1. Klassifikation	22
4.1.2. Antezedenten.....	22
4.1.3. Semantischer Typ der Antezedenten.....	22
4.1.4. Verteilung der NI-Fälle auf die Verben des Frames	22
4.1.5. Zusammenfassung der wichtigsten Beobachtungen	22
4.2. Self_Motion.....	23
4.2.1. Source.....	23

4.2.1.1. Klassifikation	23
4.2.1.2. Antezedenten	24
4.2.1.3. Semantischer Typ der Antezedenten	27
4.2.1.4. Verteilung der NI-Fälle auf die Verben des Frames	28
4.2.1.5. Zusammenfassung der wichtigsten Beobachtungen	30
4.2.2. Goal	31
4.2.2.1. Klassifikation	31
4.2.2.2. Antezedenten	32
4.2.2.3. Semantischer Typ der Antezedenten	35
4.2.2.4. Verteilung der NI-Fälle auf die Verben des Frames	36
4.2.2.5. Zusammenfassung der wichtigsten Beobachtungen	37
4.2.3. Self_mover	38
4.2.3.1. Klassifizierung	38
4.2.3.2. Antezedenten und ihr semantischer Typ	40
4.2.3.3. Verteilung der NI-Fälle auf die Verben des Frames	41
4.2.3.4. Zusammenfassung der wichtigsten Beobachtungen	43
4.2.4. Direction	44
4.2.4.1. Klassifikation	44
4.2.4.2. Antezedenten und ihr semantischer Typ	45
4.2.4.3. Verteilung der NI-Fälle auf die Verben des Frames	46
4.2.4.4. Zusammenfassung der wichtigsten Beobachtungen	47
4.3. Localisation	47
4.3.1. Klassifikation	47
4.3.2. Antezedenten	48
4.3.3. Semantischer Typ der Antezedenten	51
4.3.4. Verteilung der NI-Fälle auf die Verben des Frames	52
4.3.5. Zusammenfassung der wichtigsten Beobachtungen	53
4.4. Relative_location	53
4.4.1. Klassifikation	53
4.4.2. Antezedenten	54
4.4.3. Semantischer Typ der Antezedenten	56
4.4.4. Verteilung der NI-Fälle auf die Verben des Frames	56

4.4.5. Zusammenfassung der wichtigsten Beobachtungen	57
4.5. Visual_perception.....	57
4.5.1. Klassifikation	57
4.5.2. Antezedenten.....	58
4.5.3. Semantischer Typ der Antezedenten.....	61
4.5.4. Verteilung der NI-Fälle auf die Verben des Frames	61
4.5.5. Zusammenfassung der wichtigsten Beobachtungen	62
4.6. Zusammenfassung der Ergebnisse der quantitativen Analyse	63
4.6.1. Klassifikation	63
4.6.2. Antezedenten.....	63
4.6.3. Semantischer Typ der Antezedenten.....	66
4.6.4. Verteilung der nicht realisierten Core-Elemente auf die Verben der Frames	67
4.7. Vergleich der NI-Annotationsergebnisse zweier Annotatoren im Wegbeschreibungskorpus.....	73
4.8. Vergleich der Besonderheiten von NI-Annotation in SemEval und Wegbeschreibungskorpus.....	75
4.8.1. Klassifikation	76
4.8.2. Abstand zwischen der Anapher und dem Antezedenten.....	77
4.8.3. Semantischer Typ der Antezedenten.....	77
4.8.4. Verteilung der NI-Fälle auf Frame-einführende Elemente	78
4.8.5. Zusammenfassung der wichtigsten Beobachtungen	79
Zusammenfassung.....	81
Literaturverzeichnis.....	83

Abkürzungsverzeichnis

Folgende Abkürzungen werden in der Arbeit verwendet:

Antez.	Antezedent
constr.	constructional
DIR (VP)	DIRECTION des Frames VISUAL_PERCEPTION
DIR (SM)	DIRECTION des Frames SELF_MOTION
F:LS	LOCALISATION (Frame)
F:SM	SELF_MOTION (Frame)
F:RL	RELATIVE_LOCATION (Frame)
F:VP	VISUAL_PERCEPTION (Frame)
KA	kein Antezedent
LOC	LOCATION
LO (LS)	LOCATED_OBJECT des Frames LOCALISATION
LO (RL)	LOCATED_OBJECT des Frames RELATIVE_LOCATION
SE	SemEval
SM	SELF_MOVER
PO	PERC_OBJECT
PAREA	PATH:AREA
PCONT	PATH:CONTAINER
PERC	PERCEIVER
PLINE	PATH:LINE
PPOINT	PATH:POINT
REF	REFERENCE
UV	USE_VEHICLE
VP-K	VP-Koordination
WB	Wegbeschreibungskorpus

Einführung

Die vorliegende Arbeit basiert auf dem Projekt „SightSee“ (Roth, 2009), (Schuldes, 2009), in dessen Rahmen die automatische Generierung multimodaler (visueller und textueller) Wegbeschreibungen für Fußgänger entwickelt wurde. Eine der Ideen dieses Projekts, die noch nicht realisiert wurde, ist die Generierung der textuellen Weganweisungen auf eine natürliche, nicht monotone Weise, mit Benutzung der Phänomene, die Menschen normalerweise verwenden, um jemandem einen Weg zu erklären. Damit dies möglich ist, müssen zuerst diese Phänomene in der natürlichen Sprache untersucht werden. Es ist also notwendig, eine Studie durchzuführen, bei welcher ein Korpus mit von Menschen erstellten Weginstruktionen analysiert wird.

Im Rahmen dieser Arbeit wird das Phänomen der Null Instantiation (Ruppenhofer, 2010). in einem Frame-annotierten Wegbeschreibungskorpus untersucht. Es besteht in der Auslassung einiger Textteile (Wörter oder Phrasen), die durch den weiteren Kontext interpretierbar sind. Zum Beispiel lässt der Sprecher in (1) im zweiten Satz die Information über die Stelle, wo man abbiegen muss, aus. Der Hörer kann diese Information aber aus dem ersten Satz rekonstruieren, und zwar dahingehend, dass das Abbiegen gleich nach dem Erreichen von *McDonald's* passiert.

- (1) Laufen Sie bis zum McDonald's. Biegen Sie dann links ab.

Die Idee der Untersuchung von Null Instantiation im Kontext wurde dem Projekt „Linking Events and Their Participants in Discourse“ (Ruppenhofer, 2009) entnommen. Dessen Aufgabe war die Zuweisung der semantischen Rollen zu den Konstituenten des Textes (semantic role labelling) im SemEval-Korpus (Baker, 2007). Semantic role labelling ist eine Aufgabe, bei welcher semantische Rollen den syntaktischen Argumenten des Prädikats zugewiesen werden und die normalerweise auf Satzebene durchgeführt wird. Das besondere an der Annotation von SemEval ist, dass dabei versucht wurde, semantische Rollen in einem weiteren Kontext zu verlinken. D.h. wenn eine semantische Rolle lokal (auf der Satzebene) nicht realisiert werden konnte (Null Instantiation), wurde sie, wenn möglich, als Anapher an die Konstituenten des weiteren Kontextes verlinkt.

Die Annotation der Null Instantiation des Wegbeschreibungskorpus wurde nach dem gleichen Prinzip wie in SemEval durchgeführt. Die Aufgabe bestand aus folgenden Schritten: zuerst wurden im Frame-annotierten Korpus von Wegbeschreibungen anaphorische Bezüge von nicht realisierten semantischen Rollen annotiert, danach wurde eine qualitative und quantitative Auswertung der annotierten Daten gemacht. Es wurde dabei untersucht, wie oft das Phänomen der Null Instantiation vorkommt, wie es repräsentiert wird und dementsprechend, ob es beim Systemaufbau der automatischen Wegbeschreibung berücksichtigt werden muss. Die Ergebnisse der Untersuchung können als Grundlage zur Konzeption eines Lernverfahrens dienen, das in der Arbeit erstellte annotierte Korpus als Trainings- bzw. Evaluierungskorpus für einen entsprechenden Lernalgorithmus.

Die Auswertung des Korpus wurde als Vergleich zu SemEval durchgeführt, dessen Texte dem Genre Belletristik angehören. Dabei war es interessant zu prüfen, welche Ähnlichkeiten und Unterschiede es in der Repräsentation der Null Instantiation in unterschiedlichen Genres gibt. Außerdem wurde SemEval als Trainingskorpus für den Annotator benutzt. Es wurde am Beispiel von 50 Sätzen die NI-Annotation durchgeführt und mit der Annotation des SemEval-Korpus verglichen. Für die Qualitätskontrolle stand auch die teilweise durchgeführte NI-Annotation des Wegbeschreibungskorpus eines anderen Annotators zur Verfügung.

Die Arbeit ist wie folgt gegliedert: Im ersten Kapitel werden die in dieser Arbeit verwendeten Begriffe Frame-Semantik, Null Instantiation und Anapher erklärt. Im zweiten Kapitel werden das SemEval-Korpus und das Wegbeschreibungskorpus, sowie die Besonderheiten ihrer Textsorten vorgestellt. Im dritten Kapitel werden zuerst das Annotationstool, der Ablauf der Annotation und Annotationskriterien beschrieben. Danach wird die im Rahmen dieser Arbeit durchgeführte Teilannotation von SemEval mit der ursprünglichen Annotation im SemEval-Korpus verglichen und die Fehleranalyse gemacht. Kapitel vier stellt die Parameter der Datenanalyse und Resultate der NI-Annotation im Wegbeschreibungskorpus selbst sowie im Vergleich zu SemEval dar. Außerdem wird hier zur Qualitätskontrolle die Fehleranalyse bei der Annotation zweier Annotatoren des Wegbeschreibungskorpus durchgeführt. Abschließend werden die wichtigsten Beobachtungen der durchgeführten Arbeit zusammengefasst.

1. Theoretische Grundlagen

Dieses Kapitel erklärt in der vorliegenden Arbeit behandelte Begriffe. Zuerst werden die Theorie der *Frame-Semantik* (Abs. 1.1) bzw. die lexikalische Datenbank *FrameNet* (Abs. 1.2) vorgestellt. Danach wird der Begriff *Null Instantiation* sowie die Klassifikation von *Null Instantiation* erläutert (Abs. 1.3). Abschließend stellt Abschnitt 1.4 die Begriffe *Anapher* und *Antezedent* vor.

1.1. Frame-Semantik

Frame-Semantik ist eine semantische Theorie, die durch prototypische Situationen und dazugehörige Partizipanten lexikalische Bedeutung des Prädikats präsentiert. Diese Theorie wurde von Fillmore (Fillmore, 2003) entwickelt und später im FrameNet-Projekt (siehe Abs. 1.2) realisiert. In der Theorie der Frame-Semantik wird eine prototypische Situation als Frame und ihre Partizipanten als *semantische Rollen* oder *Frame-Elemente*¹ dargestellt. Frame-Elemente werden Konstituenten des Satzes zugewiesen. Ein Wort, das einen Frame auslöst, wird als *Frame-einführendes-* oder *frame-evoking Element* bezeichnet. So wird beispielsweise eine prototypische Situation für STUDYING durch das Verb *study* eingeführt (evoziert) und beinhaltet folgende Partizipanten: einen Student (STUDENT), einen Lehrer (TEACHER), eine Institution, an der studiert wird (INSTITUTION) und ein Fach, das studiert wird (SUBJECT). In Tabelle 1 sind Beispiele für die Frame-Elemente des Frames STUDYING (fettmarkiert) aufgeführt.

Frame-Element	Beispiel
STUDENT	She <u>studies</u> astrophysics under him at the university.
TEACHER	She <u>studies</u> astrophysics under him at the university.
SUBJECT	She <u>studies</u> astrophysics under him at the university.
INSTITUTION	She <u>studies</u> astrophysics under him at the university .

Tabelle 1: Frame-Elemente des Frames STUDYING aus FrameNet.

Frame-Elemente eines Frames sind entweder als *Core-* oder *Non-Core-Elemente* repräsentiert. Core-Elemente sind konzeptuell notwendige Elemente eines Frames. In Tabelle 1 sind alle Rollen Core-Elemente. Andere Frame-Elemente des Frames STUDYING, wie z.B. LEVEL (*a*

¹ Im Folgenden werden die Begriffe (*semantische*) Rolle und *Frame-Element* synonym verwendet.

graduate student) oder MOTIVATION (*He is studying for his Master's*) sind Non-Core Elemente, weil sie optional sind und nicht unbedingt realisiert werden müssen.

1.2. FrameNet

FrameNet ist eine lexikalische Ressource für das Englische, in der lexikalische Einheiten, vor allem Verben, Substantive und Adjektive, mit Frame-semantischer Beschreibung versehen sind (Baker, 1998).

FrameNet liefert ein Verzeichnis der lexikalischen Einheiten, die einen Frame einführen. Bei jedem Eintrag sind angegeben: Frames, die die jeweilige lexikalische Einheit einführen kann, sowie alle möglichen syntaktischen Realisierungen der Frame-Elemente dieses Frames und Frame-annotierte Beispielsätze. In einem anderen Verzeichnis werden Frames beschrieben. Die Beschreibung jedes Frames beinhaltet die Definition, alle zu ihm gehörenden Frame-Elemente, die in Core- und Non-Core-Elemente geteilt sind, die Liste der Wörter, denen dieser Frame zugewiesen werden kann und seine Beziehung zu den anderen Frames.

1.3. Null Instantiation

Wie in 1.1. erwähnt wurde, sind Core-Elemente konzeptuell notwendige Elemente eines Frames. Jedoch kann es vorkommen, dass diese Frame-Elemente in einem Satz nicht realisiert (ausgelassen) werden, wie z.B. der Ort des Ankommens, das durch die Rolle LOCATION des Frames LOCALISATION in (2) repräsentiert wird. In solchen Fällen spricht man von *Null Instantiation* (NI).

(2) Um 15 Uhr sind wir angekommen LOCATION 0.

Null Instantiation wird aus zwei Perspektiven betrachtet: dem Typ der *Interpretation* und dem Typ des *Licensors*. Bei der Interpretation wird, je nachdem, ob eine nicht realisierte Rolle aus dem Kontext rekonstruiert werden kann, zwischen *Definite Null Instantiation (DNI)* und *Indefinite Null Instantiation (INI)* unterschieden. Beim Licensor wird davon ausgegangen, dass entweder eine bestimmte lexikalische Einheit oder eine grammatische Struktur vorhanden sein muss, damit die Auslassung eines Core-Elements stattfinden kann. Dementsprechend

wird NI entweder als *lexical licensor* oder *constructional licensor* klassifiziert. Unten werden die Typen der Interpretation und des Licensors ausführlicher beschrieben.

1.3.1. Interpretation

Die Typen der Interpretation unterscheidet man, je nachdem, ob ein nicht realisiertes Frame-Element durch situativen Kontext oder ein Argument des linguistischen Kontexts abgedeckt werden kann. Bei *Definite Null Instantiation* geht es um Auslassungen, deren Interpretation aus dem Kontext ableitbar ist. Wird der Satz in (2) im weiteren Kontext betrachtet (3), ist ersichtlich, dass der Ort des Ankommens im vorhergehenden Satz als Ziel des Fahrens genannt wurde und deswegen bei der zweiten Erwähnung ausgelassen werden konnte.

- (3) Wir sind drei Stunden nach [Berlin GOAL] gefahren. Um 15 Uhr sind wir angekommen LOCATION 0.

Es ist aber nicht unbedingt notwendig, dass im linguistischen Kontext ein Element vorhanden sein muss, auf das ein nicht realisiertes Element referiert. Eine Entität kann auch im ganzen Text komplett ausgelassen und trotzdem als *DNI* klassifiziert werden, weil sie z.B. durch situativen Kontext rekonstruierbar ist. Beispielsweise ist die Angabe des Ortes in (4), von dem es weitergeht (SOURCE), überflüssig, weil sich die beiden Gesprächsteilnehmer an diesem Ort befinden. Außerdem kann in einer Gesprächssituation bestimmte Information allen Teilnehmern bekannt sein und muss deswegen nicht explizit erwähnt werden (in (5) die Frage zum Thema der früheren Konversation (TOPIC)).

- (4) Lass uns weitergehen SOURCE 0!

- (5) Was hat der Arzt gesagt TOPIC 0?

Im Fall von *Indefinite Null Instantiation* können fehlende Core-Elemente nicht durch den Kontext abgedeckt werden. So kommt beispielsweise *INI* in (6) aufgrund des Auslassens des Subjekts (AGENT) in einer Passivkonstruktion und in (7) aufgrund des fehlenden Objekts (INGESTIBLES) in einer generischen Aussage vor.

(6) Das Packet wurde eingeliefert AGENT 0.

(7) Abends viel essen INGESTIBLES 0 ist ungesund.

1.3.2. Licensor

Je nachdem, welches linguistische Phänomen (bestimmte lexikalische Einheit oder grammatische Konstruktion) die Auslassung einer Rolle beeinflusst, wird zwischen dem *lexical licensor* und *constructional licensor* unterschieden.

Lexical licensor kommt meistens in den Fällen vor, in welchen ein Element (Wort, Phrase oder Satz) bereits genannt wurde und kann bei seiner nächsten Erwähnung ausgelassen werden (in (8) die zweite Nennung des Verkehrsmittels (SALIENT_ENTITY)).

(8) Es wurde uns mitgeteilt, dass der Zug mit Verspätung kommt. Wir mussten eine Stunde warten SALIENT_ENTITY 0.

Constructional licensor wird durch eine bestimmte syntaktische Konstruktion ausgelöst, z.B. eine Passivkonstruktion, wie im oben betrachteten Beispiel (6). Andere Beispiele für *constructional licensor* sind die Nichtrealisierung des Subjekts im Imperativsatz (SELF_MOVER in (9)) oder des Subjekts im zweiten Teilsatz bei einer VP-Koordination (SPEAKER in (10)).

(9) Lauf SELF_MOVER 0 die große Straße entlang!

(10) Er setzte sich auf das Sofa und erzählte SPEAKER 0 mir diese Geschichte.

1.4. Anaphern

Unter *Anaphern* (Heim, 1993) versteht man in der Textlinguistik Elemente eines Textes, die auf andere bereits erwähnte Elemente im Kontext referieren. Man spricht in diesem Fall von *anaphorischer Verbindung* zwischen diesen Elementen. Ein Element, auf das referiert wird, heißt *Antezedent*. Das Ziel der Verwendung einer Anapher ist der Aufbau der semantischen Zusammenhänge im Text und die Vermeidung der Wiederholungen bei der Erwähnung der

bereits eingeführten Information. Anaphorische Verbindung besteht in (11) zwischen der Anapher *es* und dem Antezedent *das Buch*, auf den sie sich bezieht.

(11) **Das Buch** ist sehr interessant. Ich empfehle dir, **es** zu lesen.

Eine besondere Art der Anaphern bilden Nullanaphern. Diese sind lexikalische Einheiten, die komplett ausgelassen werden, aber aus dem Kontext durch den Bezug auf andere Elemente des Textes ableitbar sind. In (12) ist die Nullanapher die nicht realisierte zweite Erwähnung von Peter, die sich auf den Antezedenten *Peter* bezieht.

(12) **Peter** stand an der Apotheke und **0** wartete auf Maria.

Wie man aus dem Beispiel (12) sieht, sind Nullanaphern mit den Auslassungen der oben beschriebenen *Definite Null Instantiation* assoziiert. Bei dieser besteht eine anaphorische Verbindung zwischen dem nicht realisierten konzeptuell notwendigen Frame-Element und einem anderen identischen Element des Textes, auf das sich dieses Element bezieht. So können die Fälle der *Definite Null Instantiation* als Anaphern und Textteile, auf die sie sich beziehen, als Antezedenten betrachtet werden.

2. Korpora

In diesem Kapitel werden das im Rahmen der vorliegenden Arbeit annotierte Wegbeschreibungskorpus (Abs. 2.2) sowie das SemEval-Korpus (Abs. 2.1), welches als Vorbild dieser Annotation gedient hat, vorgestellt.

Bei den Texten des Wegbeschreibungskorpus und des SemEval-Korpus handelt es sich um zwei verschiedene Textarten. Im ersten Fall ist es eine Weginstruktion, im Zweiten belletristische Literatur. Im Folgenden werden Eigenschaften der Textarten der beiden Korpora sowie dadurch bedingte Unterschiede in der Frame-Annotation beschrieben.

2.1. SemEval

2.1.1. Texteigenschaften

SemEval ist ein frame-annotiertes Korpus, bestehend aus 2000 Sätzen. Es enthält Texte belletristischer Literatur, für welche die Verwendung einer großen Vielfalt unterschiedlicher sprachlicher Mittel und Satzstrukturen charakteristisch ist und die dementsprechend wenige Regularitäten aufweisen. Nachfolgend sind einige Besonderheiten der literarischen Texte genannt:

Die Satzlänge der belletristischen Texte kann von einem Wort (13) bis zu mehreren Zeilen variieren (14).

(13) What!

(14) It was certain from the first, however, that they would eventually be detected, as the cook, from the evidence of one or two tradespeople who have caught a glimpse of him through the window, was a man of most remarkable appearance - being a huge and hideous mulatto, with yellowish features of a pronounced negroid type.

Zu den zahlreichen Phänomenen, die verwendet werden, zählen z.B. Anreden (15), Wiederholungen (16), Ellipse (17), Passivkonstruktionen (18), Nebensätze (19), Partizipialsätze (*peering at this sinister relic* in (16)).

(15) Lord, sir, what a face it was!

(16) Very interesting - very interesting, indeed! said Holmes, peering at this sinister relic.

(17) Most interesting!

(18) The table was piled with half-eaten dishes and dirty plates.

(19) He says that they are not human.

Es wurde am Beispiel von 50 Sätzen manuell gezählt, welche Entitäten im SemEval-Korpus vorkommen. Die Ergebnisse sind in der Tabelle 2 dargestellt. Es ist ersichtlich, dass fast die Hälfte der Entitäten (49.6 %) aufgrund der Verwendung zahlreicher Dialoge und dementsprechend Anreden vom semantischen Typ *Person* sind. Die Kategorien *Körperteil* bzw. *Indoor-Objekt* treten in 15.4 % bzw. 13.7 % der Gesamtanzahl aller Entitäten auf. Andere Entitäten kommen eher unregelmäßig vor.

Entitäten des SemEval-Korpus		
Kategorie	gesamt	%
Person (<i>constable, Holmes, man</i>)	58	49.6
Körperteil (<i>face, eye, finger</i>)	18	15.4
Indoor-Objekt (<i>door, chair, candle</i>)	16	13.7
Objekt der Natur (<i>grass, shrubbery</i>)	5	4.4
Indoor-Ort (<i>room, kitchen, bedroom</i>)	4	3.4
Gebäude (<i>house</i>)	4	3.4
Kleidungsstück (<i>pocket, shoe</i>)	2	1.7
Zeit/Tageszeit (<i>hour, evening</i>)	2	1.7
Bewegungsprozess (<i>drive, walk</i>)	2	1.7
Firmenname (<i>Marx&Co</i>)	2	1.7
Outdoor-Ort (<i>wood</i>)	1	0.8
Musikinstrument (<i>guitar</i>)	1	0.8
Straße (<i>avenue</i>)	1	0.8
Waffe (<i>revolver</i>)	1	0.8
Σ	117	100

Tabelle 2: Entitäten des SemEval-Korpus (50 Sätze). **Gesamt:** Anzahl der Vorkommen jeder Entität (in %: relativ zu der Gesamtanzahl aller Entitäten). Σ : Gesamtanzahl der Entitäten.

2.1.2 Annotation in SemEval

SemEval beinhaltet außer Frame-Annotation, die Angaben zu Lemma, POS und syntaktische Dependenz der lexikalischen Einheiten. Für die Annotation von SemEval wurden mehr als 300 unterschiedlichen Frames verwendet. In (20) sieht man, dass praktisch jedes Wort ein Auslöser für einen Frame sein kann, was zu einer hohen Anzahl an Frames in einem Satz führt (hier acht).

- (20) A [cold AMBIENT_TEMPERATURE] and melancholy [walk SELF_MOTION] of a [couple CARDINAL_NUMBERS] of [miles MEASURE_CARDINAL_EXTENT] [brought BRINGING] us to a [high DIMENSION] wooden gate, [which RELATIVISATION] opened into a gloomy [avenue ROADWAYS] of chestnuts.

Zusätzlich zu der Frame-Annotation wurde in einem Teil von SemEval (englische Sherlock Holmes-Texte, bestehend aus 438 Sätzen) anaphorische Bindung der nicht realisierten semantischen Rollen annotiert. D.h. wenn eine konzeptuell notwendige Rolle lokal (auf der Satzebene) nicht realisiert werden konnte, also ein Core-Element konnte keiner Satzkonstituente zugewiesen werden, wurde es versucht diese Rolle im Diskurs zu verlinken. Sie wurde, wenn es möglich war, an ein Element des weiteren Kontexts, auf das sie sich bezieht, gebunden. Ausführliche Beschreibung der NI-Annotation wird in Kapiteln 3 und 4 am Beispiel der Annotation des Wegbeschreibungskorpus gegeben.

2.2. Wegbeschreibungskorpus

2.2.1. Texteigenschaften

Das Korpus von Wegbeschreibungen wurde im Rahmen des Projektes SightSee als Trainingskorpus für automatische Generierung der Weginstruktionen entwickelt. Es enthält 145 von verschiedenen Personen gemachte Routenbeschreibungen in Heidelberg. Das Korpus ist in Indoorkorpus, mit Routen innerhalb des Campus im Neuenheimer Feld, und Outdoorkorpus, mit Routen im Zentrum der Stadt, unterteilt.

Weginstruktionen verfügen in der Regel über eine relativ feste Struktur und weisen daher einige Gesetzmäßigkeiten auf. Dazu gehören beispielsweise einfache kurze Sätze, bestehend aus einem Prädikat, Subjekt und Objekten und/oder Adjunkten (21); Verknüpfung zweier kurzer Sätze durch VP-Koordination (22), Temporalsätze (23), Imperativ (24), Partizipialsätze (25).

- (21) Man folgt dem Weg geradeaus über einen Betriebsparkplatz. Rechts sieht man den botanischen Garten. Links befindet sich die Mensa. Man folgt nun dem Weg links.

- (22) Sie schauen nach rechts in den Gang und sehen Gebäude 346. Dann biegen Sie nach links ab und laufen Richtung Treppe. Sie gehen die Treppen hinunter und kommen in eine Aula.
- (23) Sie werden, wenn Sie weiter geradeaus gehen, auf die Mittermaierstraße stoßen. Sobald sie am Chinarestaurant Orchid Royal-Tai vorbeigekommen sind, gehen Sie noch etwas geradeaus.
- (24) Geh die Straße entlang, bis du ein großes Gebäude siehst.
- (25) Unten angekommen, gehst du nach links.

Genauso wie bei SemEval wurde im Wegbeschreibungskorpus am Beispiel von 50 Sätzen die Anzahl der vorkommenden Entitäten berechnet. Die Ergebnisse sind in Tabelle 3 zusammengefasst. Dort ist zu erkennen, dass in Weginstruktionen eine relativ übersichtliche Anzahl von Kategorien verwendet wird, wobei der größte Teil (44.8 %) auf die Kategorie *Person* entfällt (vgl. SemEval).

Entitäten des Wegbeschreibungskorpus		
Kategorien	gesamt	%
Person (<i>Sie, du</i>)	56	44.8
Straße (<i>Straße, Bergheimerstraße, große Straße</i>)	18	14.4
Indoor-Objekt (<i>Tür, Infotafel, Säule</i>)	7	5.6
Outdoor-Objekt (<i>Schild, Straßenbahnschiene</i>)	8	6.4
Indoor-Ort (<i>Raum, Flur, Fachschaft</i>)	10	8.0
Outdoor-Ort (<i>Platz, Bismarkplatz, Kreuzung</i>)	12	9.6
Gebäude (<i>Cafe Botanik, Restaurant, INF 325</i>)	14	11.2
Σ	125	100

Tabelle 3: Entitäten des Wegbeschreibungskorpus (50 Sätze). **Gesamt:** Anzahl der Vorkommen jeder Entität (in %: relativ zu der Gesamtanzahl aller Entitäten). Σ : Gesamtanzahl der Entitäten.

Eine Wegbeschreibung enthält die detaillierte Schilderung einer Route mit den Angaben zu dem Startpunkt, dem Ziel, der Richtung und zahlreichen Orientierungspunkten, die als Zwischenziele für die einzelnen einander abwechselnde Pfade innerhalb der Route auftreten.

Am besten lassen sich die Eigenschaften einer Weginstruktion anhand der vier am häufigsten vorkommenden Frames (insgesamt 99.3 %) ² des Wegbeschreibungskorpus erläutern, die das ganze Spektrum der typischen Beschreibungen eines Weges abdecken (siehe Abs. 2.2.2).

2.2.2 Annotation im Wegbeschreibungskorpus

Die Daten des Wegbeschreibungskorpus sind mit Informationen über Wortart, Lemma und Frames versehen. Die vier Frames, mit denen ca. 950 Sätze (vgl. SemEval) das Korpus annotiert wurden, sind: SELF_MOTION, LOCALISATION, RELATIVE_LOCATION, VISUAL_PERCEPTION (in der Regel ca. 1 bis 3 Frames pro Satz, unabhängig von der Länge (vgl. SemEval)).

SELF_MOTION ist mit 59.8 % (936 Mal) der am häufigsten vorkommende Frame im Wegbeschreibungskorpus. Er stellt den Prozess der Bewegung des Protagonisten dar und wird durch die Bewegungsverben wie z.B. *gehen, laufen, folgen, überqueren, vorbeilaufen* usw. evoziert. Typische Partizipanten dieser Situation sind: die laufende Person (SELF_MOVER), Anfangs- (SOURCE) und Endpunkt (GOAL) der Route und die Richtung, in die sich diese Person bewegt (DIRECTION), wie in (26) gezeigt.

- (26) [Vom Startpunkt aus SOURCE] gehst [du SELF_MOVER] [geradeaus DIRECTION] [in Gebäude 345 GOAL].

Zwischen den Bewegungsphasen kann auch ein Anhaltspunkt angegeben werden. Ein solcher Fall wird durch den Frame LOCALISATION repräsentiert und kommt im Wegbeschreibungskorpus in 21.7 % aller Frame-Vorkommen des Korpus (339 Mal) vor. Typische Verben dieses Frames sind *sich befinden, ankommen, sein, stehen, erreichen* usw. Prototypische Partizipanten sind ein Ort (LOCATION) und eine an diesem Ort befindliche Person oder auch ein Gegenstand (LOCATED_OBJECT) (27).

- (27) [Du LOCATED_OBJECT] stehst dann [vor einem grünen Schaukasten LOCATION].

² Die restlichen 0.7 % entfallen auf die in dieser Arbeit nicht betrachteten Frames USE_VEHICLE (10 Mal) und READING (1 Mal).

Der Frame `RELATIVE_LOCATION` beschreibt die Lage einer Person oder eines Gegenstandes (`LOCATED_OBJECT`) relativ zu einem bestimmten Objekt (`REFERENCE`). `RELATIVE_LOCATION` kommt im Wegbeschreibungskorpus in 13.5 % aller Frame-Vorkommen des Korpus (211 Mal) ausschließlich als eingebetteter Frame vor und wird immer durch lokale Präpositionen wie *vor*, *nach*, *in* etc. evoziert. So wird beispielsweise in (27) dieser Frame, eingebettet vom Frame-Element `LOCATION` des Frames `LOCALISATION`, dem Teil *vor einem grünen Schaukasten* zugewiesen.

Während der Beschreibung einer Route wird auch auf Objekte verwiesen, welche die Person, wenn sie der Route folgt, sehen wird. Die laufende Person wird in diesem Fall zum `PERCEIVER` und das Orientierungsobjekt zum `PERC_OBJECT`. `DIRECTION` zeigt die genaue Richtung, in der sich `PERC_OBJECT` befindet (28). Der Frame `VISUAL_PERCEPTION`, zu dem diese Rollen gehören, tritt im Wegbeschreibungskorpus in 4.3 % aller Frame-Vorkommen des Korpus (67 Mal). Er wird durch Verben wie *sehen*, *erkennen*, *erblicken* usw. evoziert.

(28) [Links `DIRECTION`] sehen [Sie `PERCEIVER`] [ein großes Gebäude `PERC_OBJECT`].

Im Vergleich zum SemEval-Korpus wurden semantische Rollen beim Wegbeschreibungskorpus nur lokal kodiert, also nur den Konstituenten des Satzes, in dem ihr Frame vorkommt, zugewiesen. War das nicht möglich, ist die Rolle unrealisiert geblieben.

In den nachfolgenden Kapiteln werden die im Rahmen dieser Arbeit durchgeführten Annotation und Analyse anaphorischer Bindung nicht realisierter semantischer Rollen im Wegbeschreibungskorpus vorgestellt.

3. NI-Annotation im Wegbeschreibungskorpus

Dieses Kapitel beinhaltet die Beschreibung der Annotation des Wegbeschreibungskorpus. Am Anfang wird das Tool für Frame-Annotation SALTO vorgestellt (Abs. 3.1). Als nächstes werden Beispiele gegeben, wie die Frames (Abs. 3.2) bzw. Null Instantiation (Abs. 3.3) mit SALTO annotiert werden. Danach werden die Kriterien der Annotation aufgelistet (Abs. 3.4)

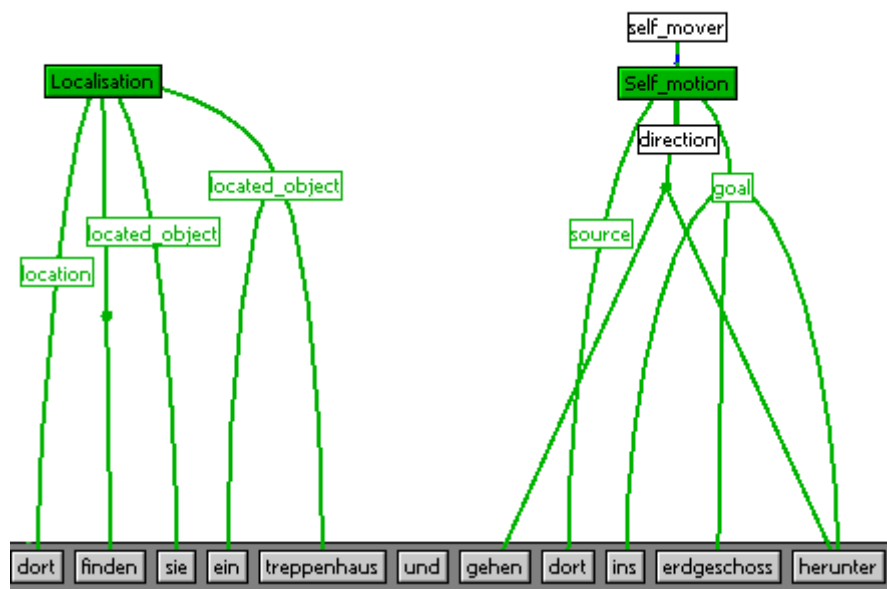
und abschließend Übereinstimmungen zweier Annotatoren bei der Teilannotation der Null Instantiation im SemEval-Korpus berechnet (Abs. 3.5).

3.1. Annotationstool

Für die Annotation des Wegbeschreibungskorpus wurde das grafische Tool für die Annotation der Textkorpora SALTO (Burchardt, 2006a) verwendet, das im Rahmen des Projekts SALSA entwickelt wurde (Burchardt, 2006b). SALTO ermöglicht semantische Annotation der Korpora, die bereits syntaktisch annotiert wurden. Die Ergebnisse der Annotation werden in einer XML-Datei gespeichert.

3.2. Frame-Annotation

Grafik 1 zeigt ein Beispiel für die lokale Frame-Annotation eines Satzes im Wegbeschreibungskorpus mit Hilfe des SALSA-Tools. Hier evozieren die Verben *finden* bzw. *heruntergehen* die Frames LOCALISATION bzw. SELF_MOTION. Die anderen Satzkonstituenten wurden mit den entsprechenden Frame-Elementen annotiert. Die nicht verlinkten Rollen SELF_MOVER und DIRECTION des Frames SELF_MOTION sind lokal nicht realisierte Rollen.



Grafik 1: Frame-Annotation im Wegbeschreibungskorpus.

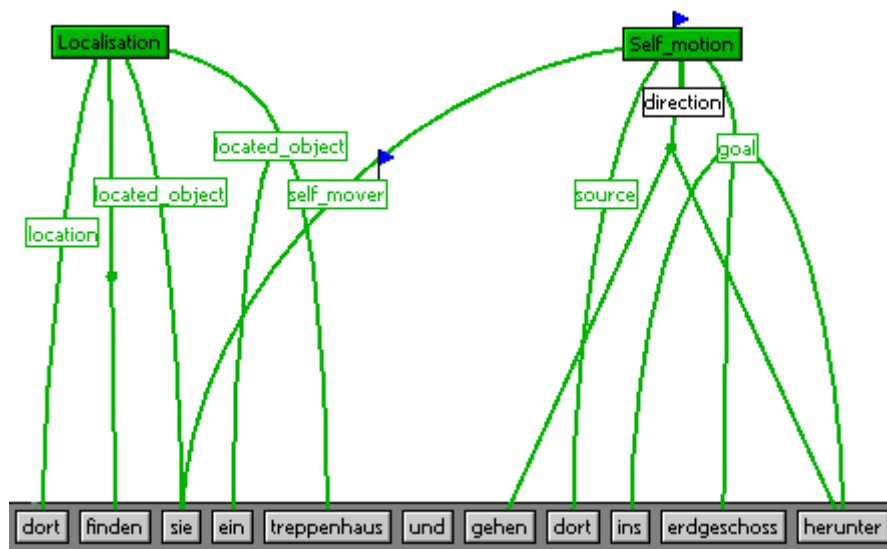
3.3. Ablauf der NI-Annotation

Bei der NI-Annotation wurden folgende Kategorien unterschieden:

- *Definite Null Instantiation*
- *Indefinite Null Instantiation*
- *Lexical licensor*
- *Constructional licensor*³

Die Annotation wurde wie folgt durchgeführt: Bei jeder nicht realisierten Rolle wurde zuerst der Typ von Interpretation bzw. *Licensor* bestimmt, indem das entsprechende Flag gesetzt wurde. Danach, sofern möglich, wurde die klassifizierte lokal nicht realisierte Rolle mit dem Antezedenten des weiteren Kontexts verlinkt.

Grafik 2 zeigt ein Beispiel der NI-Annotation der NI-Rollen SELF_MOVER und DIRECTION aus Grafik 1. Die lokal nicht realisierte Rolle SELF_MOVER wurde an die Rolle LOCATED_OBJECT des Frames LOCALISATION gebunden. Für die Rolle DIRECTION konnte kein Antezedent gefunden werden.

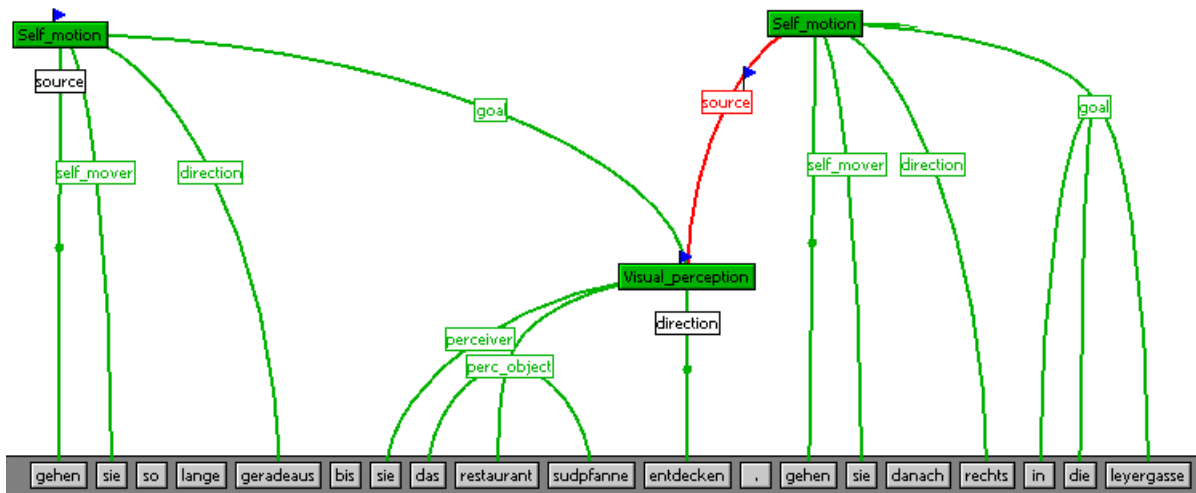


Grafik 2: Beispiel einer NI-Annotation im Wegbeschreibungskorpus.

³ Für die ausführliche Beschreibung der Kategorien siehe Abschnitt 1.3.

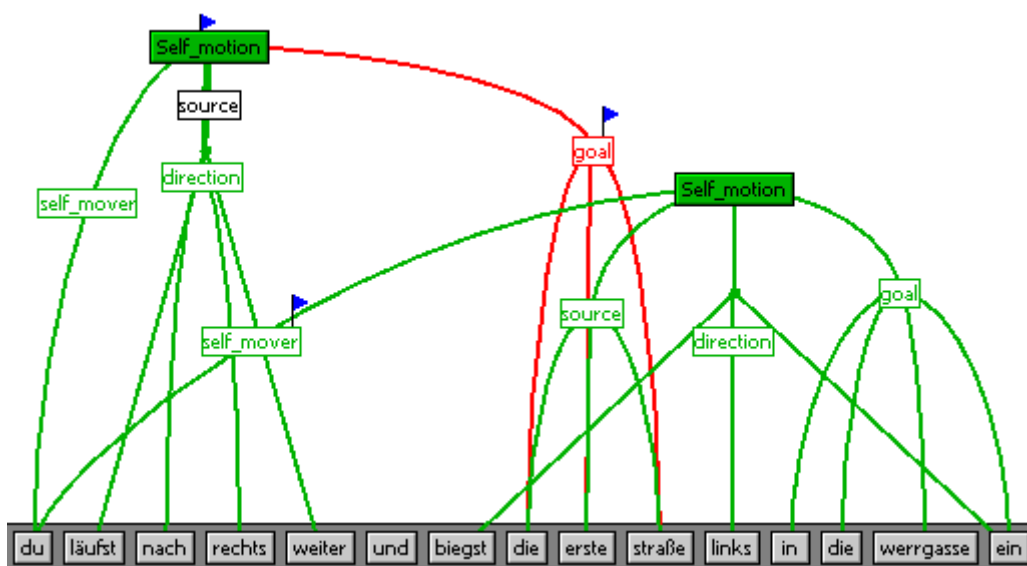
3.4. Annotationskriterien

In der Rolle der Antezedenten konnten sowohl Frame-Elemente, bestehend aus einzelnen lexikalischen Einheiten als auch aus Phrasen auftreten. Außerdem konnte der ganze Frame als Antezedent für eine NI-Rolle dienen. Grafik 3 zeigt ein Beispiel der Verlinkung der NI-Rolle SOURCE (rot markiert) mit dem Frame VISUAL_PERCEPTION, der damit im Frame SELF_MOTION eingebettet wird.



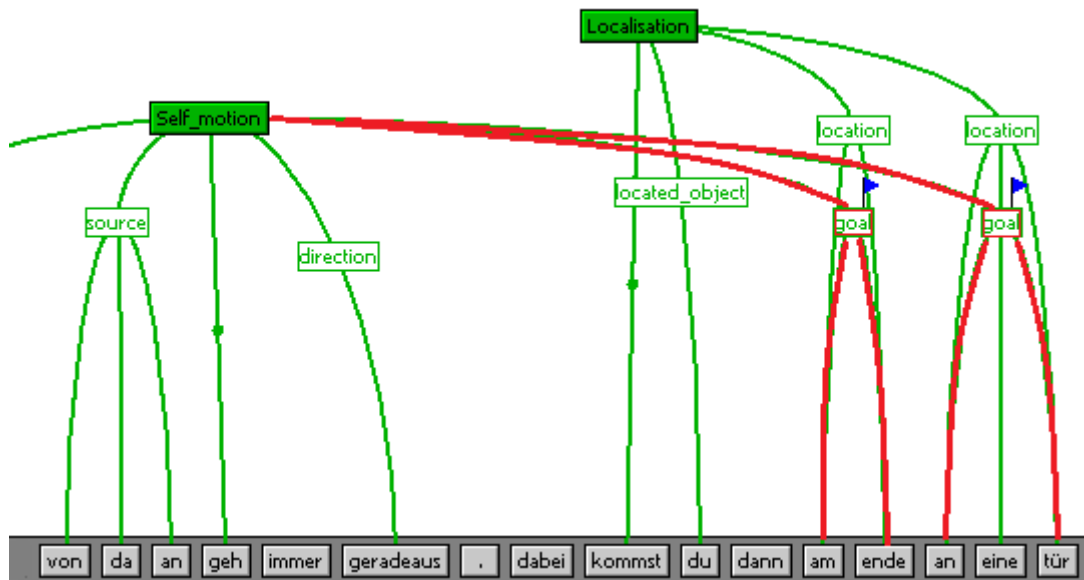
Grafik 3: Beispiel einer Verlinkung der Rolle SOURCE mit dem Frame VISUAL_PERCEPTION

Grundsätzlich wurde versucht, eine NI-Rolle maximal, d.h. mit allen Konstituenten, die zu dem jeweiligen Antezedenten gehören, zu verlinken. In Grafik 4 übernimmt die Rolle GOAL (rot markiert) bei der Verlinkung mit der Rolle SOURCE alle ihre Konstituenten.



Grafik 4: Beispiel einer NI-Annotation der Rolle GOAL

Wenn es für eine Anapher einen Antezedenten bestehend aus mehreren Komponenten gab, wurde bei ihr eine zusätzliche gleiche NI-Rolle eingeführt und mit allen Komponenten des Antezedenten verlinkt. In Grafik 5 kommt das Frame-Element LOCATION des Frames LOCALISATION doppelt vor. Da beide Komponenten als Antezedent für die Rolle GOAL auftreten, wurde auch das zweite GOAL eingeführt.



Grafik 5: Annotationsbeispiel bei doppeltem Antezedenten

In der Regel wurde nach dem Antezedenten im vorausgehenden Kontext gesucht. Konnte in diesem Fall kein Antezedent gefunden werden, wurde der nachfolgende Kontext der NI-Rolle untersucht. Ausnahmen bilden Fälle, beispielsweise Partizipialsätze, bei denen die Verlinkung der Rolle mit dem semantischen Typ *Person* syntaktisch bedingt ist. In solchen Fällen wurde die Rolle mit dem Antezedenten im Hauptsatz verlinkt (29), unabhängig davon, ob er vor oder nach der Anapher steht. Wenn das nicht möglich war (z.B. auf Grund eines expletiven *es*, dem keine Rolle zugewiesen wird (30)), wurde im vorausgehenden Kontext nach dem Antezedenten gesucht.⁴

- (29) Unten angekommen **LOCATED_OBJECT 0**, wenden [**Sie SELF_MOVER**] sich nach rechts.

⁴ Eine Anapher sowie ihr Antezedent werden in diesem Beispiel sowie allen darauf folgenden fettmarkiert.

- (30) Nach ein paar Metern kommt rechts vor dir eine blaue Treppe, die [**du** SELF_MOVER] hinunterläufst. Unten angekommen SELF_MOVER 0, geht es rechts weiter.

Bei der mehrmaligen Erwähnung des Antezedenten wurde eine NI-Rolle an diejenige Erwähnung gebunden, die am nächsten zu ihr liegt. In (31) wird von den zwei Bindungsmöglichkeiten die Rolle LOCATED_OBJECT als Antezedent für die Anapher SELF_MOVER gewählt, weil der Weg zwischen beiden am kürzesten ist.

- (31) [**Sie** SELF_MOVER] gehen den Gang entlang in Richtung Raum 111. Dort finden [**Sie** LOCATED_OBJECT] ein Treppenhaus und gehen SELF_MOVER 0 ins Erdgeschoss herunter.

Es wurden nur Core-Elemente verlinkt. Nicht-Core-Elemente konnten jedoch als mögliche Antezedenten betrachtet werden.

Es ist vorgekommen, dass einige Teile des Textes im Korpus nicht mit Frames annotiert wurden, aber wichtig waren, weil sie als Antezedenten für andere Frame-Elemente auftreten konnten. Deswegen wurden an diesen Stellen die Frames USE_VEHICLE (1 Mal) und READING (1 Mal) nachannotiert.

3.5. Fehleranalyse bei der NI-Teilannotation im SemEval-Korpus

Vor Beginn der Annotation des Wegbeschreibungskorpus wurde für die Quantifizierung der Übereinstimmungen eine Teilannotation (50 Sätze) auf Basis der Testdaten von SemEval durchgeführt. Die gefundene Übereinstimmung sollte damit für die Annotation des Wegbeschreibungskorpus hinreichend charakteristisch sein.

Die Übereinstimmung der im Rahmen dieser Arbeit erstellten Annotation (Annotator 1) mit der Annotation von SemEval (Annotator 2) wurde in folgenden Kategorien gemessen⁵:

⁵ Die Übereinstimmung in jedem NI-Fall wurde nur dann gemessen, wenn beiden Annotatoren eine Aussage über den Typ der Interpretation bzw. des licensors gemacht haben.

- Typ der Interpretation
- Typ des licensors
- verlinkt/nicht verlinkt
- im Fall der Verlinkung, ob der Antezedent gleich ist

Die Resultate des Vergleichs beider Annotationen sind als Übereinstimmungsmatrix in den Tabellen 4, 5 und 6 dargestellt.

Tabelle 4 zeigt die Übereinstimmungen/Nichtübereinstimmungen beider Annotatoren mit Fokus auf den Interpretationstyp. In 47.7 % aller Fälle haben beide Annotatoren eine nicht realisierte Rolle gleich als *DNI* und in 34.1 % als *INI* klassifiziert. Insgesamt wurden damit 81.8 % aller Fälle gleich annotiert, was einem Kappa-Wert (Cohen, 1960) von 0.63 entspricht.

Typ der Interpretation							
	Annotator 2						
		DNI	%	INI	%	Σ	%
Annotator 1	DNI	21	47.7	6	13.7	27	61.4
	INI	2	4.5	15	34.1	17	38.6
	Σ	23	52.2	21	47.8	44	100

Tabelle 4: Übereinstimmungsmatrix beider Annotatoren mit Fokus auf den Interpretationstyp. Neben den absoluten Werten sind die Werte in Prozent angegeben.

Tabelle 5 zeigt die Anzahl der Übereinstimmungen/Nichtübereinstimmungen mit Fokus auf den Typ des licensors. Hier liegt die Übereinstimmung insgesamt bei 79.6 %, wobei von beiden 75.1 % als *lexical licensor* und 4.5 % als *constructional licensor* klassifiziert wurden. Der Kappa-Koeffizient ist in diesem Fall 0.21.

Typ des licensors							
	Annotator 2						
		lexical	%	constr.	%	Σ	%
Annotator 1	lexical	33	75,1	7	15.9	40	91.0
	constr.	2	4.5	2	4.5	4	9.0
	Σ	35	79.6	9	20.4	44	100

Tabelle 5: Übereinstimmungsmatrix beider Annotatoren mit Fokus auf den Typ des licensors. Neben den absoluten Werten sind die Werte in Prozent angegeben.

In Tabelle 6 sind die Annotationsergebnisse beider Annotatoren mit Fokus auf die Verlinkung der NI-Fälle dargestellt, d.h. ob ein klassifizierter NI-Fall auch verlinkt wurde. In 38.6 % haben beide Annotatoren für einen NI-Fall einen Antezedent gefunden, wobei die Übereinstimmung in der Wahl des Antezedenten bei 100 % liegt. In 52.3 % ist eine NI-Rolle in beiden Annotationen nicht verlinkt geblieben. Die insgesamt 90.9 % Übereinstimmung entspricht einem Kappa-Koeffizient von 0.81.

		Annotator 2					
		verlinkt	%	KA	%	Σ	%
Annotator 1	verlinkt	17	38.6	3	6.8	20	45.4
	KA	1	2.3	23	52.3	24	54.6
	Σ	18	40.9	26	59.1	44	100

Tabelle 6: Übereinstimmungsmatrix beider Annotatoren mit Fokus auf die Verlinkung der NI-Fälle. Neben den absoluten Werten sind die Werte in Prozent angegeben.

Nach Landis und Koch (Landis, 1977) wurden die Ergebnisse des Vergleichs auf folgende Weise interpretiert: eine **exzellente** Übereinstimmung wurde bei der Verlinkung der NI-Fälle erreicht ($k = 0.81$). Bei der Bestimmung des Interpretationstyps haben beide Annotatoren eine **gute** ($k = 0.63$) und in Hinsicht auf den Typ des licensors eine **ordentliche** ($k = 0.21$) Übereinstimmung erreicht.

4. Auswertung der NI-Annotation im Wegbeschreibungskorpus

Dieses Kapitel stellt die Ergebnisse der NI-Annotation im Wegbeschreibungskorpus dar. Zuerst werden die Parameter, die bei der Auswertung der Daten berücksichtigt wurden, beschrieben (Abs. 4.1). Dann werden die Ergebnisse präsentiert (Abs. 4.2 - 4.5) und anschließend in Tabellen zusammengefasst (Abs. 4.6). Als Nächstes werden die Annotationsergebnisse zweier Annotatoren im Wegbeschreibungskorpus verglichen (Abs. 4.7). Als Letztes werden die Besonderheiten der NI-Annotation in SemEval und im Wegbeschreibungskorpus verglichen (Abs. 4.8).

4.1. Parameter der Datenanalyse

Der nächste Schritt nach der Annotation des Korpus war die quantitative bzw. qualitative Analyse der annotierten Daten. Für jede Rolle wurden folgende Parameter betrachtet:

4.1.1. Klassifikation

Hier wurde für jede Rolle die Anzahl ihrer Vorkommen, Anzahl der Zuordnungen jedes NI-Falls einer bestimmten Kategorie (*DNI/INI* bzw. *lexical licensor/constructional licensor*), Anzahl der verlinkten bzw. nicht verlinkten Rollen sowie alle Prozentwerte automatisch berechnet. Die Anzahl der Sonderfälle (Kategorien *VP-Koordination* und *Andere* des *constructional licensors*) wurde manuell ermittelt.

4.1.2. Antezedenten

Als nächstes wurde für alle verlinkte NI-Fälle jeder Rolle automatisch errechnet, an welchen Antezedenten die jeweilige Rolle verlinkt wurde und wie groß der Abstand (in dazwischen liegenden Frames) zwischen ihr und dem Antezedenten ist (inkl. aller Prozentwerte). Das heißt, wie viele Frames liegen zwischen dem Frame der Anapher und dem Frame des Antezedenten. Außerdem wurden auffällige Phänomene analysiert und entsprechende Beispiele gegeben.

4.1.3. Semantischer Typ der Antezedenten

Diese Kategorie wurde halbautomatisch analysiert: es wurden alle Antezedenten der jeweiligen Rolle automatisch extrahiert und manuell einer bestimmten Kategorie (einem semantischen Typ) zugeordnet.

4.1.4. Verteilung der NI-Fälle auf die Verben des Frames

Hier wurde automatisch die Verteilung der NI-Fälle jeder Rolle auf die Verben des Frames, zu dem die jeweilige Rolle gehört, sowie dazugehörige Prozentwerte errechnet. Dabei wurde gezählt, bei welchen Verben eine NI einer Rolle vorkam und in wie vielen Fällen davon es möglich war, einen Antezedenten zu finden.

4.1.5. Zusammenfassung der wichtigsten Beobachtungen

Abschließend wurden auffällige Phänomene, die bei der Analyse beobachtet wurden, zusammengefasst.

4.2. Self_Motion

In diesem Abschnitt werden die Ergebnisse der NI-Annotation der Core-Elemente SOURCE, GOAL, SELF_MOVER und DIRECTION des Frames SELF_MOTION wegen der Vielfalt der zu analysierenden Fälle für jedes Frame-Element einzeln präsentiert⁶.

4.2.1. Source

4.2.1.1. Klassifikation

Tabelle 7 zeigt den Überblick der klassifizierten Fälle für die Rolle SOURCE. Wie man sieht, wurde der größte Teil der Rollen SOURCE (insgesamt 73.6 % aller Vorkommen von SOURCE) lokal nicht realisiert. Es wurden 62.5 % aller Vorkommen von SOURCE als *DNI* und 11.1 % als *INI* klassifiziert. Bei dem Typ des licensors sowohl von *DNI* als auch von *INI* handelt es sich in 100 % der Fälle um den *lexical licensor*.

Klassifikation der NI-Fälle von SOURCE													
Alle Rollen		interpretation				licensor				linking			
gesamt	%	DNI	%	INI	%	lexical	%	constr.	%	verlinkt	%	KA	%
944	15.4	590	62.5	105	11.1	695	73.6	0	0.0	303	43.6	392	56.4

Tabelle 7: Klassifikationsergebnisse der NI-Fälle von SOURCE. **Gesamt:** Anzahl der Vorkommen von SOURCE (in %: relativ zur Gesamtzahl aller Rollen im Korpus). **Interpretation** bzw. **licensor:** Anzahl der Vorkommen von *DNI/INI* bzw. *lexical licensor/constructural licensor* (in %: relativ zu der Anzahl aller Vorkommen von SOURCE). **Linking:** Anzahl der verlinkten bzw. nicht verlinkten Rollen (in %: relativ zu allen NI-Fällen von SOURCE).

Bei den nicht realisierten Fällen der Rolle SOURCE treten sowohl Definite- als auch Indefinite Null Instantiation auf. Dabei gehören alle *INI*-Rollen zu den nicht verlinkten Fällen. In (32) ist ein nicht verlinkter *INI*-Fall der Rolle SOURCE dargestellt. Hier wurde statt der Angabe des Ortes, von welchem die Strecke beginnt, eine unbestimmte Angabe *als nächstes* verwendet.

(32) Als nächstes geht man geradeaus SOURCE 0.

Bei *DNI* wurden 48.6 % der Fälle an keinen Antezedenten gebunden. So wurde beispielsweise SOURCE in (33) nicht explizit genannt, jedoch kann geschlossen werden, dass der Prozess

⁶ Framebeschreibung hier und im Folgenden siehe Abschnitt 2.2.2.

des Laufens über die Brücke in dem Moment beginnt, in welchem die Person die Brücke erreicht hat. Da aber dem Ausdruck *über die Brücke* das Frame-Element PATH:CONTAINER desselben Frames zugewiesen ist, kann SOURCE mit ihm nicht verlinkt werden. Somit ist SOURCE in der Rolle PATH:CONTAINER implizit und wird als *DNI* klassifiziert.

(33) Laufen Sie [über die Brücke PATH:CONTAINER] SOURCE 0/*DNI*.

Eine andere Situation, bei welcher ein der Kategorie *DNI* zugeordneter SOURCE nicht verlinkt wurde, ist der Anfang der Routenbeschreibung. Dort wird SOURCE ausgelassen, weil der Startpunkt beiden Gesprächsteilnehmern bekannt ist und der Sprecher ihn deswegen nicht explizit nennt. Der Startpunkt kann hier aus dem situativen Kontext abgeleitet werden. Dieses Phänomen wurde 53 Mal beobachtet, was 18.5 % aller nicht verlinkten *DNI*-Fälle und 13.5 % aller nicht verlinkten Fälle von SOURCE entspricht.

Die *DNI*-Rollen, für die ein Antezedent gefunden wurde, werden im nächsten Abschnitt diskutiert.

4.2.1.2. Antezedenten

Tabelle 8 zeigt den Überblick der Antezedenten für SOURCE. Die fettmarkierten Antezedenten sind diejenigen Frames, mit denen SOURCE in 17.5 % der Fälle verlinkt wurde.

Antezedenten für SOURCE													
		Anapher											
		verlinkt		Abstand									
		gesamt	%	0	%	1	%	2	%	3	%	4	%
Antez.	GOAL	123	40.7	115	93.5	7	5.7	1	0.8	-	-	-	-
	LOC	45	14.9	33	73.4	6	13.3	5	11.1	-	-	1	2.2
	SOURCE	27	8.9	21	77.8	5	18.5	-	-	1	3.7	-	-
	LO (LS)	19	6.3	19	100	-	-	-	-	-	-	-	-
	REF	9	3.0	7	77.8	1	11.1	-	-	1	11.1	-	-
	PAREA	8	2.6	8	100	-	-	-	-	-	-	-	-
	PLINE	6	2.0	6	100	-	-	-	-	-	-	-	-
	PCONT	5	1.6	5	100	-	-	-	-	-	-	-	-
	PPOINT	4	1.3	3	75.0	1	25.0	-	-	-	-	-	-
	GOAL (UV)	2	0.7	2	100	-	-	-	-	-	-	-	-
	VEHICLE	1	0.3	-	-	1	100	-	-	-	-	-	-
	PO	1	0.3	-	-	1	100	-	-	-	-	-	-
	F:LS	28	9.2	25	89.3	3	10.7	-	-	-	-	-	-
	F:RL	17	5.6	10	58.8	3	17.7	4	23.5	-	-	-	-
	F:VP	4	1.3	4	100	-	-	-	-	-	-	-	-
	F:SM	3	1.0	1	33.3	1	33.3	1	33.3	-	-	-	-
	READING	1	0.3	1	100	-	-	-	-	-	-	-	-
	Alle	303	100	260	85.8	29	9.6	11	3.6	2	0.7	1	0.3

Tabelle 8: Antezedenten für SOURCE (Anapher). **Verlinkt:** Anzahl der Verlinkungen von SOURCE mit dem jeweiligen Antezedenten (**gesamt**) (in %: relativ zu allen Verlinkungen von SOURCE). **Abstand:** Abstand (in dazwischenliegenden Frames) zwischen der Anapher und dem Antezedenten (in %: relativ zu allen Verlinkungen von SOURCE mit dem jeweiligen Antezedenten). **Alle:** obengenannte Parameter für Bindungsfälle der Rolle insgesamt.

Wie man aus Tabelle 8 sieht, können gänzlich unterschiedliche Rollen und Frames für das Frame-Element SOURCE als Antezedenten auftreten. Es fällt jedoch auf, dass es am häufigsten mit GOAL (40.7 %) verlinkt wird. Danach folgen LOCATION (14.9 %), SOURCE (9.2 %), Frame LOCALISATION (9.2 %) usw. Im Folgenden werden die obengenannten vier Antezedenten ausführlicher betrachtet.

Eine typische Situation, in der SOURCE an GOAL gebunden wurde, ist in (34) dargestellt. Hier stehen 2 SELF_MOTION Frames nebeneinander. Bei dem Ersten, evoziert durch das Verb *vor-*

beilaufen, ist GOAL realisiert. Bei dem Zweiten, evoziert durch das Verb *einbiegen*, ist SOURCE nicht realisiert. In diesem Fall kann SOURCE mit GOAL verlinkt werden, weil dort, wo ein Pfad endet, der Nächste anfängt.

- (34) Laufen sie vorbei an Gebäude 324 bis [**der Weg auf eine Straße trifft GOAL**]. Biegen **SOURCE 0** Sie links in die Straße ein.

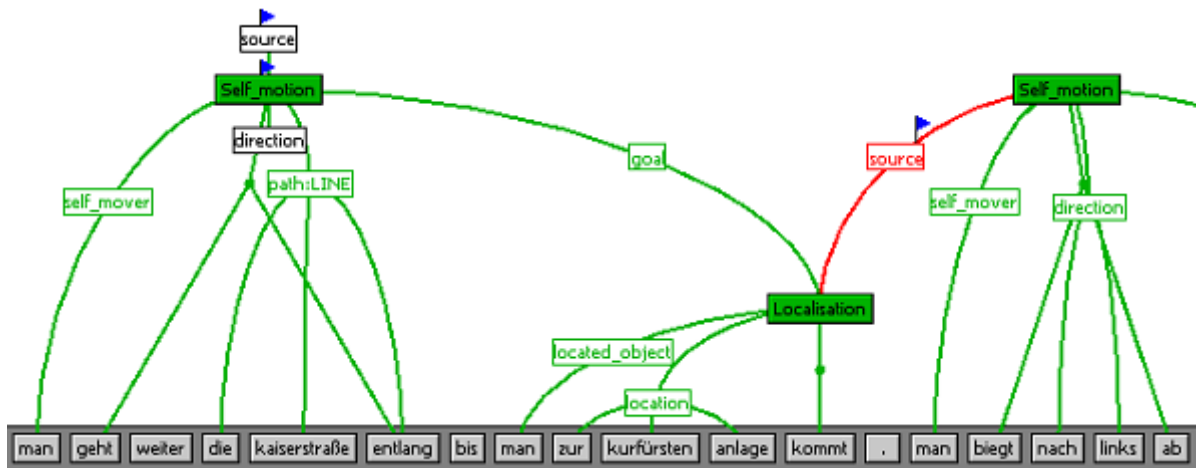
Die Bindung SOURCE an GOAL wie in (34) passiert aber nur in dem Fall, wenn beide Frames verschiedene Pfade beschreiben und der zweite Pfad nach dem Ersten folgt. Eine Ausnahme bilden Fälle wie in (35). Hier verweisen beide Frames auf den gleichen Pfad, wobei die Beschreibung des Zweiten die Bedeutung des Ersten präzisiert. Der Prozess des Umlaufens passiert während des Steuerns und dementsprechend kann SOURCE von *umlaufen* nicht gleichzeitig GOAL von *steuern* sein. In solchen Fällen wird die nicht realisierte SOURCE des zweiten Frames mit der SOURCE des Ersten (sofern vorhanden) verlinkt.

- (35) [**Von hier SOURCE**] steuern Sie auf die Tür zu, indem Sie eine Halle umlaufen **SOURCE 0**.

Ein typisches Beispiel der Bindung von SOURCE an die Rolle LOCATION ist in (36) angegeben. In dieser Situation befindet sich die Person an einem Ort (oder kommt an einem Ort an), von dem sie der Strecke (weiter)folgt. So kann dieser Ort als SOURCE für den nächsten Pfad betrachtet werden.

- (36) Man gelangt [**in einen großen Raum mit grünen Arbeitstischen LOCATION**] und läuft **SOURCE 0** geradeaus auf die Tür „Hörsaal I“ zu.

Auch der Frame LOCALISATION kann einen Startpunkt, entweder der Route oder des Anfanges, von einem Pfad bezeichnen. In Grafik 6 wurde LOCALISATION als Antezedent für SOURCE gewählt, weil er bereits bei der Frame-Annotation als eingebetteter Frame von SELF_MOTION annotiert wurde und damit die Rolle des Frame-Elementes GOAL erfüllte.



Grafik 6: Ein Beispiel der Verlinkung der Rolle SOURCE mit dem Frame LOCALISATION

4.2.1.3. Semantischer Typ der Antezedenten

Die semantischen Typen der Antezedenten der Rolle SOURCE sind in Tabelle 9 dargestellt⁷.

Semantischer Typ des Antezedenten		
	SOURCE	
	gesamt	%
Objekt	43	14.2
Gebäude	45	14.9
Straße	128	42.2
Ort	71	23.4
Andere	16	5.3
Σ	303	100

Tabelle 9: **gesamt**: Anzahl der semantischen Typen der Antezedenten für nichtrealisierte Rolle SOURCE (in %: relativ zu der Anzahl aller Verlinkungen von SOURCE). Σ : semantische Typen der Antezedenten von SOURCE insgesamt.

Aus Tabelle 9 ist ersichtlich, dass ein großer Teil der Antezedenten auf den semantischen Typ *Straße* (42.2 %) entfallen. Die zweithäufigste Kategorie ist *Ort* (23.4 %). Danach folgen die semantischen Typen *Gebäude* (14.9 %) und *Objekt* (14.2 %). Am geringsten ist die Kategorie *Andere* (5.3 %) vertreten. In 68.7 % (11 Mal) wurden dieser Gruppe Lokaladverbien *hier*, *dort* usw. zugeordnet (37). In 31.3 % gehören dieser Kategorie die Wörter und Ausdrücke, die keinen bestimmten Startpunkt beschreiben, z.B. *bei der nächsten Gelegenheit* in (38) und die das Wort *Startpunkt* als Start der Route nennen, ohne weitere Spezifizierung des Ortes (39).

⁷ Beispiele der Kategorien hier und im Folgenden siehe Tabelle 3 Abschnitt 2.2.1.

- (37) Folgen Sie [**hier SOURCE**] halb rechts dem Straßenverlauf der unteren Straße. Laufen **SOURCE 0** Sie geradeaus und passieren Sie vier Querstraßen.
- (38) [**Bei der ersten Gelegenheit SOURCE**] folgen Sie rechts dem kurzen Buckel. Folgen **SOURCE 0** Sie dem Weg bis zur Einmündung in die Zwingerstraße.
- (39) [**Vom Startpunkt aus SOURCE**] geht es zuerst nach links bis das Gebäude aufhört und danach nochmal nach links. Dabei steuert **SOURCE 0** man auf das Gebäude 325 zu.

4.2.1.4. Verteilung der NI-Fälle auf die Verben des Frames

Tabelle 10 stellt die Verteilung der NI-Fälle der Rolle SOURCE auf die Verben⁸ des Frames SELF_MOTION dar.

⁸ Verben mit synonymischer Bedeutung wurden gruppiert (in Klammern sind Beispiele zu der jeweiligen Gruppe gehörende Verben angegeben). Bei der Analyse wird zur Abkürzung nur ein Verb der Gruppe genannt.

Verteilung der NI-Fälle der Rolle SOURCE auf die Verben des Frames SELF_MOTION						
	Verbenvorkommen		SOURCE			
	gesamt	%	NI	%	verlinkt	%
<i>gehen (laufen, weitergehen, weiterlaufen)</i>	446	47.8	376	84.3	184	48.9
<i>abbiegen (einbiegen)</i>	136	14.6	69	50.7	23	33.3
<i>überqueren (durchgehen, hindurchgehen durchqueren, durchlaufen)</i>	90	9.6	90	100	42	46.7
<i>vorbeikommen (vorbeilaufen, vorbeigehen, passieren)</i>	84	9.0	59	70.2	13	22.0
<i>zulaufen (zugehen, zusteuern)</i>	39	4.2	33	84.6	10	30.3
<i>sich wenden (sich drehen)</i>	34	3.6	18	52.9	11	61.1
<i>sich halten</i>	28	3.0	25	89.2	12	48.0
<i>verlassen (rausgehen, rauskommen)</i>	24	2.6	1	4.2	1	100
<i>starten (losgehen, loslaufen)</i>	21	2.2	2	9.5	0	0.0
<i>betreten (reingehen, eintreten, hineingehen)</i>	14	1.5	8	57.1	2	25.0
<i>nehmen</i>	8	0.9	5	62.5	2	40.0
<i>bleiben</i>	5	0.5	5	100	1	20.0
<i>umlaufen</i>	3	0.3	2	66.6	1	50.0
<i>verwenden</i>	2	0.2	2	100	1	50.0
Alle	934 ⁹	100	695	74.4	303	43.6

Tabelle 10: **Verbenvorkommen (gesamt)**: zeigt für jedes Verb des Frames SELF_MOTION die Anzahl seiner Vorkommen im Korpus (in %: relativ zu dem Gesamtvorkommen aller Verben in diesem Frame). **NI**: Anzahl der NI-Fälle von SOURCE bei dem jeweiligen Verb (in %: relativ zum Gesamtvorkommen des jeweiligen Verbs), **verlinkt**: Anzahl der gefundenen Antezedenten für den jeweiligen NI-Fall (in %: relativ zu der Anzahl aller NI-Fälle von SOURCE bei dem jeweiligen Verb). **Alle**: obengenannte Parameter für alle Verben.

Tabelle 10 zeigt, dass es Verben gibt, bei denen SOURCE eine hohe NI-Häufigkeit hat, z.B. frequente Verben wie *überqueren* (100 %), *sich halten* (89.2 %) oder *zugehen* (84.6 %).

Die Semantik des Verbes *überqueren* setzt keinen Startpunkt, sondern die Beschreibung eines Pfades voraus. Der Startpunkt ist dann entweder in der Aussage implizit (siehe Beispiel (33)) oder mit einem Antezedenten verlinkt (46.7 %).

⁹ Die Gesamtzahl der Verben des Frames SELF_MOTION stimmt nicht mit dem Gesamtvorkommen der Rolle SOURCE überein. Der Grund dafür sind Fälle mit doppeltem Vorkommen der Rolle innerhalb eines Frames (siehe Abschnitt 3.4). Diese Begründung gilt auch für andere Rollen und Frames.

Das Verb *sich halten* wird mit dem Ziel der Angabe einer Richtung (*sich links/rechts halten*) verwendet, wobei nicht auf den Ausgangspunkt fokussiert wird. Es geht bei dem Frame, evociert durch dieses Verb, vor allem um die Präzisierung eines Pfades, der bereits beschrieben wurde. Da die beiden Frames denselben Pfad beschreiben, kann NI SOURCE mit dem SOURCE des vorhergehenden Frames (sofern vorhanden) verlinkt werden (40).

- (40) [Nach dem Überqueren SOURCE] gehen Sie immer geradeaus. Halten SOURCE 0 Sie sich dabei links.

Beim Verb *zulaufen* liegt die Betonung auf dem Ziel des Pfades (41), wobei SOURCE ausgelassen wird.

- (41) Laufen SOURCE 0 Sie [auf das Gebäude zu GOAL].

Auffällig sind auch die Verben *verlassen* und *starten*, weil bei denen SOURCE selten ausgelassen wird (4.2 % bzw. 9.5 %). Das obligatorische Objekt des Verbs *verlassen* erfüllt die Funktion der Rolle SOURCE und ist daher meistens (zu 95.8 %) realisiert (42). Die Semantik des Verbs *starten* setzt einen Startpunkt voraus (43).

- (42) Wenden Sie sich nach dem Verlassen [des Hauses SOURCE] nach links.

- (43) Sie starten [im Gebäude 325 SOURCE].

In 9.5 % der Fälle, in denen SOURCE bei dem Verb *starten* nicht realisiert wurde, handelt es sich um als *DNI* klassifizierte Auslassungen am Anfang der Route. Sie können, wie es schon oben erwähnt wurde, aus dem situativen Kontext rekonstruiert werden.

4.2.1.5. Zusammenfassung der wichtigsten Beobachtungen

Bei der Rolle SOURCE gehören 62.5 % ihrer Gesamtvorkommen der Kategorie *DNI* und 11.1 % der Kategorie *INI* an. Alle NI-Fälle sind lexikalische Auslassungen. Insgesamt wurde in 43.6 % aller Auslassungen von SOURCE ein Antezedent gefunden. Alle verlinkten Fälle von SOURCE wurden als *DNI* definiert. Bei nicht gebundenen Fällen geht es in 73.2 % um *DNI* und 26.8 % um *INI*.

Der häufigste Antezedent für SOURCE ist die Rolle GOAL des Frames SELF_MOTION (40.7 %). Mit großem Abstand folgen die Rollen LOCATION (14.9 %), SOURCE (9.2 %) sowie der Frame LOCALISATION (9.2 %).

Als Antezedenten von SOURCE können die Wörter verschiedener semantischer Typen auftreten. Jedoch fällt auf, dass im Fall von SOURCE vor allem die Straßen ausgelassen werden (42.2 %).

Die Auslassung von SOURCE ist durch die Semantik der Verben bedingt. So setzen z.B. *verlassen* und *starten* einen Punkt, der verlassen wird bzw. einen Punkt, von dem gestartet wird, voraus. Daher ist die NI-Rate bei diesen Verben sehr niedrig (4.2 % bzw. 9.5 %). Bei anderen Verben wie z.B. *überqueren* und *bleiben* wurde NI in 100 % der Fälle beobachtet, weil der Zweck ihrer Verwendung die Beschreibung des Pfades oder des Befindens an einem Ort ist. Die Erwähnung eines Startpunktes ist in diesem Fall in der Regel nicht nötig.

4.2.2. Goal

4.2.2.1. Klassifikation

Tabelle 11 zeigt die Klassifikationsergebnisse der NI-Fälle von GOAL. Hier ist zu sehen, dass 38 % aller Vorkommen von GOAL als *DNI* und 26.8 % als *INI* klassifiziert wurden. Alle nicht realisierten Fälle, was insgesamt 64.8 % ergibt, wurden der Kategorie *lexical licensor* zugeordnet. Verlinkt wurden 43.8 % aller NI-Fälle von goal.

Klassifikation der NI-Fälle von GOAL													
Alle Rollen		interpretation				licensor				linking			
gesamt	%	DNI	%	INI	%	lexical	%	constr.	%	verlinkt	%	KA	%
959	15.6	364	38.0	257	26.8	621	64.8	0	0	272	43.8	349	56.2

Tabelle 11: Klassifikationsergebnisse der NI-Fälle von GOAL. **Gesamt**: Anzahl der Vorkommen von GOAL (in %: relativ zur Gesamtzahl aller Rollen im Korpus). **Interpretation** bzw. **licensor**: Anzahl der Vorkommen von *DNI/INI* bzw. *lexical licensor/constructional licensor* (in %: relativ zur Anzahl aller Vorkommen von GOAL). **Linking**: Anzahl der verlinkten bzw. nicht verlinkten Rollen (in %: relativ zu allen NI-Fällen von GOAL).

Gruppe *verlinkt* in der Tabelle 11 bilden nur *DNI*. Die Gruppe der nicht verlinkten Fälle bilden in 73.6 % *INI* und in 26.4 % *DNI*. In (44) ist ein Beispiel für einen *INI*-Fall gezeigt. Hier

wurde im ersten Satz nur die Richtung der Bewegung aber kein Ziel genannt. Auch im zweiten Satz ist es unmöglich, dieses Ziel zu finden.

(44) Man geht nach links GOAL 0. Als nächstes geht man geradeaus.

Ein nicht verlinkter *DNI*-Fall ist in (45) dargestellt. Hier wird gesagt, dass man abbiegen muss, sobald *die erste Straße* kommt. Dadurch ist der Ausdruck *die erste Straße* der Startpunkt des Abbiegens, dem dementsprechend bei der Frame-Annotation die Rolle SOURCE zugewiesen wurde. Es ist ersichtlich, dass die erste Straße außerdem auch das Ziel des Abbiegens ist. Aus diesem Grund wird GOAL als *DNI* klassifiziert.

(45) Biegen Sie [die erste Straße SOURCE] links ab GOAL 0/*DNI*.

Außerdem wurden Fälle gefunden, in welchen GOAL am Ende der Routenbeschreibung komplett ausgelassen wurde. Dies hat den Grund, dass den beiden Gesprächsteilnehmern das Ziel bereits bekannt ist. Insgesamt kam dieses Phänomen 11 Mal vor, was 12 % der nicht verlinkten *DNI*-Fälle bzw. 3.1 % aller nicht verlinkten Fälle von GOAL bildet (vgl.: bei SOURCE in einer ähnlichen Situation am Anfang der Route 18.5 % bzw. 13.5 %).

Nachfolgend werden die gebundenen Fälle von GOAL betrachtet.

4.2.2.2. Antezedenten

Tabelle 12 zeigt, an welche Antezedenten GOAL gebunden wurde. Die fettmarkierten Antezedenten sind die Frames, mit denen GOAL verlinkt wurde.

Antezedenten für GOAL													
		Anapher											
		verlinkt		Abstand									
		gesamt	%	0	%	1	%	2	%	3	%	4	%
Antez.	LOC	67	24.6	53	79.1	8	11.9	4	6.0	1	1.5	1	1.5
	LO (LS)	36	13.2	25	69.4	5	13.9	4	11.1	-	-	2	5.6
	SOURCE	34	12.5	29	85.3	4	11.8	1	2.9	-	-	-	-
	PLINE	30	11.1	27	90.0	2	6.7	1	3.3	-	-	-	-
	GOAL	29	10.7	20	69.1	5	17.2	3	10.3	-	-	1	3.4
	PAREA	18	6.6	16	88.9	2	11.1	-	-	-	-	-	-
	REF	7	2.6	4	57.1	2	28.6	1	14.3	-	-	-	-
	PPOINT	6	2.2	5	83.3	1	16.7	-	-	-	-	-	-
	PCONT	5	1.8	4	80.0	-	-	1	20.0	-	-	-	-
	PO	3	1.1	2	66.7	1	33.3	-	-	-	-	-	-
	F:RL	19	7.0	14	73.7	4	21.0	-	-	1	5.3	-	-
	F:LS	14	5.1	7	50.0	6	42.9	1	7.1	-	-	-	-
	F:VP	4	1.5	1	25.0	2	50.0	1	25.0	-	-	-	-
	Alle	272	100	207	76.1	42	15.4	17	6.3	2	0.7	4	1.5

Tabelle 12: Antezedenten für GOAL (Anapher). **Verlinkt**: Anzahl der Verlinkungen von GOAL mit dem jeweiligen Antezedenten (**gesamt**) (in %: relativ zu allen Verlinkungen von GOAL). **Abstand**: Abstand (in dazwischenliegenden Frames) zwischen der Anapher und dem Antezedenten (in %: relativ zu allen Verlinkungen von GOAL mit dem jeweiligen Antezedenten). **Alle**: obengenannte Parameter für Bindungsfälle der Rolle insgesamt.

Aus der Tabelle 12 ist ersichtlich, dass sowohl einzelne Rollen als auch ganze Frames Antezedenten für GOAL sein können. Weiter unten wird auf die am häufigsten auftretenden Antezedenten LOCATION, LOCATED_OBJECT und SOURCE eingegangen. Außerdem wird ein Beispiel der Verlinkung von GOAL mit dem Frame RELATIVE_LOCATION gegeben.

Die Rolle LOCATION ist der häufigste Antezedent für GOAL (24.6 %). Das Beispiel in (46) stellt eine typische Situation dar, bei welcher nach dem Prozess der Bewegung (*vorbeilaufen*) ein Ort, an dem die Person angekommen ist, genannt wird. Die Rolle GOAL des Frames SELF_MOTION, die dabei nicht realisiert wird, kann dann an diesen Ort (Rolle LOCATION des Frames LOCALISATION) gebunden werden.

- (46) Laufen Sie geradeaus an McDonald's vorbei **GOAL 0**. Sie befinden sich schon [**in der Rohrbacher Straße LOCATION**].

Ein interessanter Satz aufgrund der doppelten Bindung von GOAL und gleichzeitig ein Beispiel der Verlinkung von GOAL mit dem zweithäufigsten Antezedenten dieser Rolle LOCATED_OBJECT (13.2 %) ist in (47) gezeigt. Dort wurde das nicht realisierte GOAL von *durchqueren* sowohl an das LOCATED_OBJECT *das Ziel* als auch mit LOCATION *am anderen Ende des Raumes* gebunden, weil beide gemeinsam einen Zielpunkt bilden.

(47) Man durchquert den Raum **GOAL 0**. [**Das Ziel LOCATED_OBJECT**] ist der Uni-Shop, [**am anderen Ende des Raumes LOCATION**].

Ein relativ häufiges Phänomen (12.5 % aller Verlinkungen von GOAL) bei der Beschreibung eines Weges ist die Auslassung des Ziels (GOAL) eines Pfades, wobei dann im folgenden Frame die Beschreibung des nächsten Pfades mit der Angabe zum Startpunkt (SOURCE) beginnt. So kann ein nicht realisiertes GOAL dem SOURCE des nachstehenden Frames gleichgesetzt werden (48).

(48) Man gelangt auf die große Zwingerstraße und folgt dieser **GOAL 0**. [**Am Ende der Straße SOURCE**] biegt man links ein.

In (49) ist ein Beispiel der Verlinkung von GOAL mit dem Frame RELATIVE_LOCATION angegeben. Dort ist das Ziel des Laufens über die Brücke ein nicht explizit genannter Ort, der *nach der Brücke* kommt. Der Frame RELATIVE_LOCATION ist in diesem Fall im Frame SELF_MOTION, der durch das Verb *folgen* evoziert ist, eingebettet und erfüllt bei ihm die Funktion der Rolle SOURCE.

(49) Laufen Sie über die Brücke **GOAL 0**. [**Nach der Brücke RELATIVE_LOCATION**] folgen Sie dem Straßenverlauf.

4.2.2.3. Semantischer Typ der Antezedenten

Tabelle 13 zeigt semantische Typen der Antezedenten von GOAL.

Semantischer Typ der Antezedenten		
	GOAL	
	gesamt	%
Objekt	31	11.4
Gebäude	37	13.5
Straße	97	35.7
Ort	75	27.6
Andere	32	11.8
Σ	272	100

Tabelle 13: **gesamt**: Anzahl der semantischen Typen der Antezedenten für die nichtrealisierte Rolle GOAL (**in %**: relativ zur Anzahl aller Verlinkungen von GOAL). Σ : semantische Typen der Antezedenten von GOAL insgesamt.

Wie aus Tabelle 13 ersichtlich, kommen am häufigsten die Kategorien *Straße* (35.7 %) und *Ort* (27.6 %) vor. Die restlichen Antezedenten sind relativ gleichmäßig auf die Kategorien *Objekt* (11.4 %), *Gebäude* (13.5 %) und *Andere* (11.8 %) verteilt.

Zur Kategorie *Andere* gehören zum Einen in 78.1 % (25 Mal) Wörter und Ausdrücke, die keinen bestimmten Zielpunkt darstellen, z.B. *wenn man entweder links oder rechts abbiegen muss* in (50), und die, die das Wort *Ziel* als Endpunkt der Route nennen, ohne weitere Spezifizierung des Ortes (51). Zum Anderen wurden dieser Gruppe in 21.9 % (5 Mal) Lokaladverbien wie *dort*, *hier* usw. zugeordnet (52).

(50) Laufen Sie den Gang weiter **GOAL 0**. [**Wenn man entweder links oder rechts abbiegen muss SOURCE**] gehen Sie nach rechts.

(51) Folgen Sie der Straße **GOAL 0**. Kurz nach der Nadlerstraße, die rechts abgeht und der Neugasse, die links abgeht, befindet sich [**das Ziel LOCATED_OBJECT**].

(52) Du läufst von der Poststraße aus Richtung McDonald's **GOAL 0**. [**Dort SOURCE**] überquerst du die Straße und läufst links davon die Plöck herunter.

4.2.2.4. Verteilung der NI-Fälle auf die Verben des Frames

Tabelle 14 zeigt die Verteilung der NI-Fälle der Rolle GOAL auf die Verben des Frames SELF_MOTION.

Verteilung der NI-Fälle der Rolle GOAL auf die Verben des Frames SELF_MOTION						
	Verbenvorkommen		GOAL			
	gesamt	%	NI	%	verlinkt	%
<i>gehen (laufen, weitergehen, weiterlaufen)</i>	446	47.8	296	66.4	174	58.8
<i>abbiegen (einbiegen)</i>	136	14.6	60	44.1	17	28.3
<i>überqueren (durchgehen, hindurchgehen durchqueren, durchlaufen)</i>	90	9.6	90	100	36	40.0
<i>vorbeikommen (vorbeilaufen, vorbeigehen, passieren)</i>	84	9.0	59	70.2	9	15.2
<i>Zulaufen (zugehen, zusteuern)</i>	39	4.2	0	0.0	0	0.0
<i>sich wenden (sich drehen)</i>	34	3.6	31	91.2	2	6.4
<i>sich halten</i>	28	3.0	28	100	19	67.8
<i>verlassen (rausgehen, rauskommen)</i>	24	2.6	22	91.7	2	9.0
<i>starten (losgehen, loslaufen)</i>	21	2.2	17	80.9	4	23.5
<i>betreten (reingehen, eintreten, hineingehen)</i>	14	1.5	1	7.1	0	0.0
<i>nehmen</i>	8	0.9	8	100	4	50.0
<i>bleiben</i>	5	0.5	4	80.0	4	100
<i>umlaufen</i>	3	0.3	3	100	0	0.0
<i>verwenden</i>	2	0.2	2	100	1	50.0
Alle	934	100	621	66.5	272	43.8

Tabelle 14: **Verbenvorkommen (gesamt)**: zeigt für jedes Verb des Frames SELF_MOTION die Anzahl seiner Vorkommen im Korpus (in %: relativ zum Gesamtvorkommen aller Verben in diesem Frame). **NI**: Anzahl der NI-Fälle von GOAL bei dem jeweiligen Verb (in %: relativ zum Gesamtvorkommen des jeweiligen Verbs), **verlinkt**: Anzahl der gefundenen Antezedenten für den jeweiligen NI-Fall (in %: relativ zu der Anzahl aller NI-Fälle von GOAL bei dem jeweiligen Verb). **Alle**: obengenannte Parameter für Bindungsfälle von GOAL insgesamt.

Es fällt auf, dass GOAL bei den Verben *überqueren*, *sich halten*, *nehmen*, *umlaufen* und *verwenden* immer lokal nicht realisiert wird. Ein sehr hohes NI-Vorkommen haben die Verben *verlassen* (91.7 %) und *sich wenden* (91.2 %). Unten werden die von diesen am häufigsten vorkommenden Verben *überqueren*, *sich halten* und *verlassen* betrachtet.

Die Semantik aller drei Verben wurde bereits im Abschnitt 4.2.1.4 bei der Betrachtung der Rolle SOURCE wegen ihrer vielen (überqueren (100 %), sich halten (89.2 %)) bzw. wenigen (verlassen (4.2 %)) NI-Vorkommen bei diesen Verben diskutiert.

Wie schon erwähnt, wird bei *überqueren* auf den Pfad selbst, auf ein Objekt, das überquert wird, fokussiert. In dieser Beschreibung wird nicht nur kein Startpunkt, sondern auch kein Ziel vorausgesetzt. In 40 % der Fälle war es jedoch möglich, einen Antezedenten für das ausgelassene GOAL zu finden (siehe Beispiel (47)).

Das Verb *sich halten* wird vor allem für die Angabe einer Richtung verwendet. Auf die Angabe des Ziels des Pfades wird in diesem Fall verzichtet.

Im Gegensatz zur Rolle SOURCE, die beim Verb *verlassen* nur in 4.2 % lokal nicht realisiert wird, hat GOAL bei diesem Verb ein hohes NI-Vorkommen mit 91.7 %, weil dort der Ort, an welchem man nach dem Verlassen ankommt, in der Regel nicht explizit genannt wird (53).

(53) Verlassen Sie das Gebäude GOAL 0 und laufen Sie in Richtung Mensa.

Eine andere wichtige Beobachtung sind die Verben *zulaufen*, bei welchem GOAL nie ausgelassen wurde und *betreten*, mit einem geringen NI-Vorkommen (7.1 %). Bei beiden Verben lässt sich diese Besonderheit dadurch erklären, dass sie genau für die Angabe des Objektes/Ortes etc., auf welches man zugeht (54) oder welches betreten wird (55), also des Ziels, verwendet werden.

(54) Gehen Sie nach rechts und Sie werden genau [auf den Zip-Pool GOAL] zulaufen.

(55) Man betritt [einen Parkplatz, der als Parkplatz der Mensa ausgeschildert ist GOAL].

4.2.2.5. Zusammenfassung der wichtigsten Beobachtungen

Die Klassifikation der NI-Fälle von GOAL hat folgende Resultate ergeben: 38 % aller Vorkommen dieser Rolle sind *DNI*, 26.8 % sind *INI*. Insgesamt bilden diese NI-Fälle 64.8 % aller Vorkommen von GOAL, die ausschließlich als *lexical licensor* klassifiziert wurden. In 43.8 % aller NI-Fälle der Rolle GOAL konnte ein Antezedent gefunden werden.

Die Antezedenten von GOAL sind durch ganz unterschiedliche Rollen vertreten. Am häufigsten sind das jedoch die Rollen des Frames LOCALISATION, LOCATION und LOCATED_OBJECT. Danach folgen Core- und Non-Core-Elemente des Frames SELF_MOTION, SOURCE, PATH:LINE, GOAL usw. In 13.6 % der Fälle wurde ein ganzer Frame als Antezedent für ein nicht realisiertes GOAL gewählt.

Unter den Antezedenten von GOAL findet man große Variation an semantischen Typen. Mit 35.7 % ist *Straße* der am häufigsten vorkommende semantische Typ. Den zweiten Platz nimmt die Kategorie Ort mit 27.6 % ein.

Goal wurde im Korpus bei den Verben *überqueren*, *sich halten*, *nehmen*, *umlaufen* und *verwenden* zu 100 % lokal nicht realisiert, da die Angabe des Ziels der Bewegung in diesen Fällen entweder nicht vorgesehen oder überflüssig ist. Im Gegensatz dazu ist bei den Verben *zulaufen* und *betreten* die Angabe eines Zielpunktes notwendig. Aus diesem Grund ist das NI-Vorkommen beim Verb *zulaufen* gleich 0. Beim Verb *betreten* beträgt es 7.1 % (1 Mal).

4.2.3. Self_mover

4.2.3.1. Klassifizierung

Tabelle 15 zeigt die Ergebnisse der NI-Klassifikation der Rolle SELF_MOVER. Man sieht, dass alle NI-Fälle zur Kategorie *DNI/constructional licensor* gehören. Davon wurde die Anapher in 96.3 % mit einem Antezedenten verlinkt. In 3.7 % aller NI-Fälle konnte SELF_MOVER trotz seiner Zuordnung zur Kategorie *DNI/Constructional licensor* wegen der kompletten Auslassung des Subjekts über die ganze Beschreibung der Route nicht verlinkt werden (aufgrund Imperativ 2. P. Sg.).

Klassifikation der NI-Fälle von SELF_MOVER													
Alle Rollen		Interpretation				licensor				linking			
gesamt	%	DNI	%	INI	%	lexical	%	Constr.	%	verlinkt	%	KA	%
936	15.2	241	25.7	0	0	0	0	241	25.7	232	96.3	9	3.7

Tabelle 15: Klassifikationsergebnisse der NI-Fälle von SELF_MOVER. **Gesamt:** Anzahl der Vorkommen von SELF_MOVER (in %: relativ zur Gesamtzahl aller Rollen im Korpus). **Interpretation** bzw. **licensor:** Anzahl der Vorkommen von *DNI/INI* bzw. *lexical licensor/constructional licensor* (in %: relativ zur Anzahl der Vorkommen von SELF_MOVER). **Linking:** Anzahl der verlinkten bzw. nicht verlinkten Rollen (in %: relativ zu allen NI-Fällen mit Verlinkung von SELF_MOVER).

Die Analyse der NI-Fälle, welche der Kategorie *constructional licensor* zugeordnet wurden, hat ergeben, dass verschiedene syntaktische Konstruktionen zur Auslassung von SELF_MOVER beitragen können. Tabelle 16 fasst alle Phänomene des *constructional licensors* zusammen.

Phänomene des constructional licensor von SELF_MOVER		
Phänomen	gesamt	%
VP-Koordination	130	53.9
Imperativ	88	36.6
Expletives <i>es</i>	10	4.1
Passiv	5	2.1
Substantivierung	5	2.1
Infinitivkonstruktion	2	0.8
Partizip	1	0.4
Σ	241	100

Tabelle 16: Phänomene des *constructional licensor* von SELF_MOVER. Phänomen: Name des zu der NI beigetragenden Phänomens. **Gesamt:** Anzahl der Vorkommen des jeweiligen Phänomens (in %: relativ zu allen NI-Fällen von SELF_MOVER). Σ: alle NI-Fälle von SELF_MOVER.

Aus Tabelle 16 ist ersichtlich, dass die Weglassung von SELF_MOVER vor allem durch zwei typische Phänomene bedingt ist: VP-Koordination¹⁰ mit Unterdrückung des Subjekts des zweiten Teilsatzes (53.9 %) (56) und Imperativ in der Form der 2.P.Sg (36.6 %) (57).

(56) [Du SELF_MOVER] folgst zuerst weiter dem Straßenverlauf **und** läufst SELF_MOVER 0 dann über die Straße.

(57) Geh SELF_MOVER 0 nach rechts vorbei am Gebäude 324 in Richtung Botanischer Garten.

Andere Phänomene des *constructional licensors* von SELF_MOVER sind: expletives *es* (4.1 %) (58), Passiv (2.1 %) (59), Substantivierung (2.1 %) (60), Infinitivkonstruktion (0.8 %) (61) und Partizip (0.4 %) (62).

(58) Weiter **geht es** SELF_MOVER 0 geradeaus zwischen zwei weiteren Gebäuden.

¹⁰ Im Korpus wurde VP-Koordination jedoch als *DNI/lexical licensor* annotiert

- (59) **Gestartet wird** SELF_MOVER 0 vom Druckerraum im Theoretikum.
- (60) ...beim **Verlassen** SELF_MOVER 0 des Gebäudes 325 gleich nach links.
- (61) ...**um** nach wenigen Metern direkt wieder rechts in die Plöck **abzubiegen** SELF_MOVER 0.
- (62) Weiter geradeaus **gehend** SELF_MOVER 0...

4.2.3.2. Antezedenten und ihr semantischer Typ

Der semantische Typ von SELF_MOTION ist immer eine Person. Deswegen kann dafür nur eine der drei Rollen des Wegbeschreibungskorpus mit Personenbezeichnung SELF_MOVER (63), PERCEIVER (64) oder LOCATED_OBJECT (65) als Antezedent auftreten.

- (63) [**Du SELF_MOVER**] überquerst diese und gehst **SELF_MOVER 0** bei der ersten Möglichkeit nach der Brücke nach rechts.
- (64) Wenn [**du PERCEIVER**] links vor dir Glaskästen mit Mikroskopen siehst, geh **SELF_MOVER 0** nach links.
- (65) [**Du LOCATED_OBJECT**] stehst mit dem Rücken zur Tür. Geh **SELF_MOVER 0** nach rechts Richtung Raum 111.

Tabellen 17 und 18 beschreiben ausführlich die möglichen Antezedenten für SM. Dabei wurde VP-Koordination (Tabelle 17) als häufigstes Phänomen des *constructional licensors* getrennt von den anderen Phänomenen (Tabelle 18) ausgewertet.

Antezedenten für SELF_MOVER (VP-Koordination)									
		Anapher							
		verlinkt		Abstand					
		gesamt	%	0	%	1	%	2	%
Antez.	SM	125	96.1	109	87.2	15	12.0	1	0.8
	LO (LS)	5	3.9	5	100	-	-	-	-
	Alle	130	100	114	87.7	15	11.5	1	0.8

Tabelle 17: Antezedenten für Fälle der VP-Koordination von SELF_MOVER (Anapher). **Verlinkt:** Anzahl der Verlinkungen der Anapher mit dem jeweiligen Antezedenten (**gesamt**) (in %: relativ zu allen Verlinkungen der Anapher). **Abstand:** Abstand (in dazwischenliegenden Frames) zwischen der Anapher und dem Antezedenten (in %: relativ zu allen Verlinkungen der Anapher mit dem jeweiligen Antezedenten). **Alle:** obengenannte Parameter für Bindungsfälle von SELF_MOVER (VP-Koordination) insgesamt.

Antezedenten für SELF_MOVER (Andere)															
		Anapher													
		verlinkt		Abstand											
		gesamt	%	0	%	1	%	2	%	3	%	4	%	Ab5	%
		Antez	SM	42	41.2	18	42.9	10	23.8	10	23.8	3	7.1	1	2.4
LO (LS)	58		56.8	13	22.4	14	24.1	10	17.3	9	15.5	5	8.6	7	12.1
PERC	2		2.0	1	50.0	1	50.0	-	-	-	-	-	-	-	-
	Alle	102	100	32	31.4	25	24.5	20	19.6	12	11.8	6	5.9	7	6.8

Tabelle 18: Antezedenten für SELF_MOVER. Berechnung der gleichen Parameter wie in Tabelle 17 für die anderen Phänomene von *constructional licensor* (ohne VP-Koordination).

4.2.3.3. Verteilung der NI-Fälle auf die Verben des Frames

Tabellen 19 und 20 zeigen die Verteilung der NI-Fälle von SELF_MOVER auf die Verben des Frames SELF_MOTION. Genauso, wie bei Antezedenten wurde VP-Koordination (Tabelle 19) getrennt von den anderen Phänomenen des *constructional licensors* (Tabelle 20) ausgewertet. Wie man sieht, verteilt sich NI auf unterschiedliche Verben des Frames SELF_MOTION. Im Fall von SELF_MOVER kann aber gesagt werden, dass die Auslassung vor allem durch eine bestimmte syntaktische Konstruktion und nicht die Semantik der Verben beeinflusst wird.

Verteilung der NI-Fälle von SELF_MOVER auf die Verben des Frames SELF_MOTION						
	VP-Koordination					
	Verbenvorkommen		SELF_MOVER			
	gesamt	%	NI	%	verlinkt	%
<i>gehen (laufen, weitergehen, weiterlaufen)</i>	446	47.8	72	16.1	72	100
<i>abbiegen (einbiegen)</i>	136	14.6	17	12.5	17	100
<i>überqueren (durchgehen, hindurchgehen durchqueren, durchlaufen)</i>	90	9.6	17	18.8	17	100
<i>vorbeikommen (vorbeilaufen, vorbeigehen, passieren)</i>	84	9.0	6	7.1	6	100
<i>zulaufen (zugehen, zusteuern)</i>	39	4.2	6	15.4	6	100
<i>sich wenden (sich drehen)</i>	34	3.6	4	11.8	4	100
<i>sich halten</i>	28	3.0	4	14.3	4	100
<i>verlassen (rausgehen, rauskommen)</i>	24	2.6	1	4.2	1	100
<i>starten (losgehen, loslaufen)</i>	21	2.2	-	-	-	-
<i>betreten (reingehen, eintreten, hineingehen)</i>	14	1.5	2	14.3	2	100
<i>nehmen</i>	8	0.9	-	-	-	-
<i>bleiben</i>	5	0.5	-	-	-	-
<i>umlaufen</i>	3	0.3	1	33.3	1	100
<i>verwenden</i>	2	0.2	-	-	-	-
Alle	934	100	130	13.9	130	100

Tabelle 19: **Verbenvorkommen (gesamt)**: zeigt für jedes Verb des Frames SELF_MOTION die Anzahl seiner Vorkommen im Korpus (in %: relativ zum Gesamtvorkommen aller Verben in diesem Frame). **NI**: Anzahl der NI-Fälle von SELF_MOVER bei dem jeweiligen Verb (in %: relativ zum Gesamtvorkommen des jeweiligen Verbs), **verlinkt**: Anzahl der gefundenen Antezedenten für den jeweiligen NI-Fall (in %: relativ zu der Anzahl aller NI-Fälle von SELF_MOTION bei dem jeweiligen Verb). **Alle**: obengenannte Parameter für Bindungsfälle von SELF_MOVER (VP-Koordination) insgesamt.

Verteilung der NI-Fälle von SELF_MOVER auf die Verben des Frames SELF_MOTION						
	Andere					
	Verbenvorkommen		SELF_MOVER			
	gesamt	%	NI	%	verlinkt	%
<i>gehen (laufen, weitergehen, weiterlaufen)</i>	446	47.8	60	13.4	52	86.7
<i>abbiegen (einbiegen)</i>	136	14.6	16	11.8	16	100
<i>überqueren (durchgehen, hindurchgehen durchqueren, durchlaufen)</i>	90	9.6	9	10.0	9	100
<i>vorbeikommen (vorbeilaufen, vorbeigehen, passieren)</i>	84	9.0	4	4.8	4	100
<i>zulaufen (zugehen, zusteuern)</i>	39	4.2	3	7.8	3	100
<i>sich wenden (sich drehen)</i>	34	3.6	4	11.8	4	100
<i>sich halten</i>	28	3.0	2	7.1	2	100
<i>verlassen (rausgehen, rauskommen)</i>	24	2.6	4	16.7	4	100
<i>starten (losgehen, loslaufen)</i>	21	2.2	4	19.0	3	75.0
<i>betreten (reingehen, eintreten, hineingehen)</i>	14	1.5	1	7.1	1	100
<i>nehmen</i>	8	0.9	2	25.0	2	100
<i>bleiben</i>	5	0.5	2	40.0	2	100
<i>umlaufen</i>	3	0.3	-	-	-	-
<i>verwenden</i>	2	0.2	-	-	-	-
Alle	934	100	111	12.6	102	92

Tabelle 20: Berechnung der gleichen Parameter wie in der Tabelle 19 für die anderen Phänomene des *constructional licensor* (ohne VP-Koordination).

4.2.3.4. Zusammenfassung der wichtigsten Beobachtungen

Die Weglassung von SELF_MOVER ist in allen Fällen durch besondere Konstruktion des Satzes in dem er vorkommt bedingt. Vor allen sind hier die zwei häufigsten Phänomene VP-Koordination (53.9 %) und Imperativ (36.6 %) zu nennen. In 3.7 % aller NI-Fälle von SELF_MOVER konnte SELF_MOVER nicht verlinkt werden, da die Nennung der Person in der ganzen Routenbeschreibung ausgelassen wurde.

Der semantische Typ von SELF_MOVER ist eine Person. Daher kann ein nicht realisierter SELF_MOVER mit einer Rolle, die eine Person bezeichnet (SELF_MOVER, PERCEIVER oder LOCATED_OBJECT), verlinkt werden. Die Wahl einer Rolle als Antezedent hängt davon ab, welche dieser drei Rollen des vorhergehenden Kontexts (oder, wenn nicht vorhanden, nach-

folgenden) am nächsten an der Anapher liegt. Ausnahmen bilden Substantivierungen, Partizipien und Infinitivsätze, wo der Antezedent zuerst im Hauptsatz, unabhängig von seiner Lage gesucht wurde.

Der SELF_MOVER wurde im Wegbeschreibungskorpus bei unterschiedlichen Verben ausgelassen. In diesem Fall kann aber gesagt werden, dass die Weglassung von SELF_MOVER nicht durch die Semantik der Verben, sondern die konstruktionellen Besonderheiten des Satzes, in dem er vorkommt, bedingt ist.

4.2.4. Direction

4.2.4.1. Klassifikation

Tabelle 21 zeigt den Überblick der Klassifizierungen von NI-Fällen der Rolle DIRECTION. Wie ersichtlich ist, sind 40.7 % aller Vorkommen von DIRECTION *DNI*-Fälle und nur 1.1 % *INI*-Fälle. Dabei wurden alle Auslassungen als *lexical licensor* klassifiziert und nur in 4.1 % aller NI-Fälle mit dem Antezedenten verlinkt.

Klassifikation der NI-Fälle von DIRECTION													
Alle Rollen		interpretation				licensor				linking			
gesamt	%	DNI	%	INI	%	lexical	%	constr.	%	verlinkt	%	KA	%
992	16.2	405	40.7	10	1.1	415	41.8	0	0.0	17	4.1	398	95.9

Tabelle 21: Klassifikationsergebnisse der NI-Fälle von DIRECTION. **Gesamt:** Anzahl der Vorkommen von DIRECTION (in %: relativ zur Gesamtzahl aller Rollen im Korpus). **Interpretation** bzw. **licensor:** Anzahl der Vorkommen von *DNI/INI* bzw. *lexical licensor/constructional licensor* (in %: relativ zu der Anzahl aller Vorkommen von DIRECTION). **Linking:** Anzahl der verlinkten bzw. nicht verlinkten Rollen (in %: relativ zu allen NI-Fällen von DIRECTION).

Es ist auffällig, dass 95.9 % aller NI-Fälle von DIRECTION trotz einer so großen Anzahl der *DNI*-Fälle nicht verlinkt wurden. Diese Besonderheit kann damit erklärt werden, dass die Richtungsangabe in diesen Fällen in der Aussage implizit ist und deswegen nicht separat genannt werden muss. Die Richtungsangabe ist impliziert, wenn entweder ein auffälliges Objekt als Orientierungspunkt genannt wird (66) oder wenn der Weg selbst in eine bestimmte Richtung geht (67). In den 11 *INI*-Fällen war es nicht möglich, aus dem Kontext die Angabe der Richtung abzuleiten (68).

(66) Von dort aus steuern Sie auf den McDonalds zu DIRECTION 0.

(67) Biegen Sie hier in die Vangerowstraße ein und folgen Sie dieser DIRECTION 0.

(68) Man hält sich dabei leicht links biegt aber nicht ab DIRECTION 0.

4.2.4.2. Antezedenten und ihr semantischer Typ

Der semantische Typ der Rolle DIRECTION ist immer eine Richtung. Die Rollen, die eine Richtung darstellen, sind dementsprechend potenzielle Antezedenten für sie. Die Antezedenten für 17 verlinkte NI-Fälle von DIRECTION sind in Tabelle 22 dargestellt.

Antezedenten für DIRECTION							
		Anapher					
		verlinkt		Abstand			
		gesamt	%	0	%	1	%
Antezedent	DIR (SM)	16	94.1	14	87.5	2	12.5
	SRELATION:RIGHT	1	5.9	1	100	-	-
	Alle	17	100	15	88.2	2	11.8

Tabelle 22: Antezedenten für DIRECTION (Anapher). **Verlinkt:** Anzahl der Verlinkungen von DIRECTION mit dem jeweiligen Antezedenten (**gesamt**) (in %: relativ zu allen Verlinkungen von DIRECTION). **Abstand:** Abstand (in dazwischen liegenden Frames) zwischen der Anapher und dem Antezedenten (in %: relativ zu allen Verlinkungen von DIRECTION mit dem jeweiligen Antezedenten). **Alle:** oben genannte Parameter für Bindungsfälle von DIRECTION insgesamt.

Wie aus Tabelle 22 ersichtlich ist, tritt die Rolle DIRECTION in 94.1 % der NI-Fälle als Antezedent für DIRECTION auf. In (69) konkretisiert der zweite Satz die Bewegungsanweisung des ersten Satzes, bezeichnet also den gleichen Pfad. Somit kann auch die fehlende Richtungsangabe im zweiten Satz durch die Richtungsangabe des ersten Satzes vervollständigt werden.

(69) Du läufst [**geradeaus DIRECTION**] am Hotel „Neckar“ vorbei und auf den Bismarckplatz zu. Du bleibst immer auf derselben Straßenseite **DIRECTION 0**.

Mit der Rolle SRELATION:RIGHT des Frames RELATIVE_LOCATION wurde DIRECTION einmal verlinkt (70). Hier kann angenommen werden, dass wenn sich ein Objekt rechts von der Person befindet, sie nach rechts gehen muss, um zu diesem Objekt zu gelangen

- (70) Sie erblicken dann [**rechts SRELATION:RIGHT**] vor sich eine Hütte, die aussieht, wie ein Fliegenpilz. Steuern Sie auf diesen zu **DIRECTION 0**.

4.2.4.3. Verteilung der NI-Fälle auf die Verben des Frames

Tabelle 23 stellt die Verteilung der NI-Fälle der Rolle DIRECTION auf die Verben des Frames SELF_MOTION dar.

Verteilung der NI-Fälle von DIRECTION auf die Verben des Frames SELF_MOTION						
	Verbenvorkommen		DIRECTION			
	gesamt	%	NI	%	verlinkt	%
<i>Gehen (laufen, weitergehen, weiterlaufen)</i>	446	47.8	203	45.5	9	4.4
<i>abbiegen (einbiegen)</i>	136	14.6	13	9.6	0	0.0
<i>überqueren (durchgehen, hindurchgehen durchqueren, durchlaufen)</i>	90	9.6	83	92.2	2	2.4
<i>vorbeikommen (vorbeilaufen, vorbeigehen, passieren)</i>	84	9.0	34	40.5	4	11.8
<i>zulaufen (zugehen, zusteuern)</i>	39	4.2	26	66.7	0	0.0
<i>sich wenden (sich drehen)</i>	34	3.6	1	2.9	0	0.0
<i>sich halten</i>	28	3.0	3	7.9	0	0.0
<i>verlassen (rausgehen, rauskommen)</i>	24	2.6	20	83.3	0	0.0
<i>starten (losgehen, loslaufen)</i>	21	2.2	11	52.4	0	0.0
<i>betreten (reingehen, eintreten, hineingehen)</i>	14	1.5	9	64.2	0	0.0
<i>nehmen</i>	8	0.9	3	37.5	0	0.0
<i>bleiben</i>	5	0.5	5	100	2	40.0
<i>umlaufen</i>	3	0.3	2	66.7	0	0.0
<i>verwenden</i>	2	0.2	2	100	0	0.0
Alle	934	100	415	44.4	17	4.1

Tabelle 23: **Verbenvorkommen (gesamt)**: zeigt für jedes Verb des Frames SELF_MOTION die Anzahl seiner Vorkommen im Korpus (**in %**: relativ zum Gesamtvorkommen aller Verben in diesem Frame). **NI**: Anzahl der NI-Fälle von DIRECTION bei dem jeweiligen Verb (**in %**: relativ zum Gesamtvorkommen des jeweiligen Verbs), **verlinkt**: Anzahl der gefundenen Antezedenten für den jeweiligen NI-Fall (**in %**: relativ zur Anzahl aller NI-Fälle von DIRECTION bei dem jeweiligen Verb). **Alle**: oben genannte Parameter für Bindungsfälle von DIRECTION insgesamt.

Aus Tabelle 23 ist ersichtlich, dass es einerseits Verben gibt, bei denen überwiegend NI vorkommt. So wird DIRECTION z.B. in 100 % beim Verb *bleiben*, in 92.2 % beim Verb *überqueren* und in 83 % aller Vorkommen des Verbs *verlassen* nicht realisiert. Eine solche Besonderheit kann durch die Semantik dieser Verben erklärt werden. Beispielsweise drückt das Verb

bleiben die Lage des Objektes aus. Die Richtungsangabe ist in diesem Fall nicht vorgesehen (*Bleiben Sie auf der Hauptstraße.*). Bei den Verben *überqueren* und *verlassen* ist die Richtungsangabe überflüssig, weil sie aus dem situativen Kontext ableitbar ist (*Überqueren/verlassen Sie den Parkplatz.*).

Andererseits hat DIRECTION bei den Verben wie *sich wenden*, *sich halten* bzw. *abbiegen*, eine niedrige NI-Vorkommensrate (2.9 %, 7.9 % bzw. 9.6 %). Das hängt ebenso mit der Semantik dieser Verben zusammen. Beim Verb *sich wenden* ist es notwendig, die Bewegungsrichtung zu erwähnen (*Wenden Sie sich nach rechts*). Das Verb *sich halten* kommt im Wegbeschreibungskorpus vor allem in den Phrasen wie *halten Sie sich links*, wo die Erklärung der Richtung das Ziel der Aussage ist. Das Verb *abbiegen* setzt unter anderem die Richtungsangabe voraus (*Biegen Sie links/rechts ab.*).

4.2.4.4. Zusammenfassung der wichtigsten Beobachtungen

Die Rolle DIRECTION wurde in 41.8 % aller Vorkommen dieser Rolle nicht realisiert. Bei der Annotation wurden alle diese Fälle als *lexical licensor* klassifiziert. In 40.7 % aller Vorkommen von DIRECTION wurde NI von DIRECTION der Kategorie *DNI* und in 1.1 % der Kategorie *INI* zugeordnet. Bei der Verlinkung konnte in 4.1 % der Fälle ein Antezedent für DIRECTION gefunden werden. In 95.9 % aller NI-Fälle blieb DIRECTION unverlinkt. Der Grund dafür ist entweder die komplette Auslassung der Richtungsangabe (*INI*) oder ihr implizites Ausdrücken durch andere Rollen (*DNI*). Bei der Verlinkung wurde die nicht realisierte Rolle DIRECTION mit den Rollen vom semantischen Typ *Richtung* DIRECTION (94.1 %) und *SRELATION:RIGHT* (5.9 %) verlinkt.

4.3. Localisation

Dieser Abschnitt beschreibt die Ergebnisse der NI-Annotation der Rollen LOCATION und LOCATED_OBJECT des Frames LOCALISATION.

4.3.1. Klassifikation

Tabelle 24 stellt die Klassifikationsergebnisse von der NI-Fälle von LOCATION und LOCATED_OBJECT dar. Hier ist ersichtlich, dass die NI-Häufigkeit dieser Rollen relativ niedrig ist (LOCATION: 4.1 % aller Rollenvorkommen von LOCATION (16 Mal);

LOCATED_OBJECT: 12.7 % aller Rollenvorkommen von LOCATED_OBJECT (43 Mal)). Bei LOCATION handelt es sich in 100 % aller NI-Fälle um lexikalische Auslassungen, die aber nur in 31.3 % (5 Mal) verlinkt werden konnten. Alle Auslassungen der Rolle LOCATED_OBJECT wurden als *constructional licensor* klassifiziert. Dabei wurde das syntaktische Phänomen VP-Koordination als häufigere Ursache der Auslassung dieser Rolle (41.9 % aller NI-Fälle von LOCATED_OBJECT) getrennt von anderen Fällen von *constructional licensor* ausgewertet¹¹.

Klassifikationsergebnisse der NI-Fälle von LOCATION													
Alle Rollen		interpretation				licensor				linking			
gesamt	%	DNI	%	INI	%	lexical	%	constr.	%	verlinkt	%	KA	%
388	6.3	16	4.1	0	0	16	4.1	0	0	5	31.3	11	68.7
Klassifikationsergebnisse der NI-Fälle von LOCATED_OBJECT													
Alle Rollen		interpretation				constructional licensor				linking			
gesamt	%	DNI	%	INI	%	VP- K	%	Andere	%	verlinkt	%	KA	%
339	5.5	43	12.7	0	0	18	5.3	25	7.4	43	100	0	0

Tabelle 24: Klassifikationsergebnisse der NI-Fälle von LOCATION und LOCATED_OBJECT. **Gesamt:** Anzahl der Vorkommen jeder Rolle (in %: relativ zu der gesamten Anzahl von allen Rollen im Korpus). **Interpretation** bzw. **Licensor:** Anzahl der Vorkommen von *DNI/INI* bzw. *lexical/constructional licensor* (in %: relativ zu der Anzahl des Vorkommens der jeweiligen Rolle). **Linking:** Anzahl der verlinkten bzw. nicht verlinkten Rollen (in %: relativ zu allen NI-Fällen der jeweiligen Rolle). **VP-K.** bzw. **Andere** bei LOCATED_OBJECT: Anzahl der Vorkommen der VP-Koordination bzw. anderer Phänomene von *constructional licensor* (in %: relativ zu der Anzahl der Vorkommen der Rolle).

4.3.2. Antezedenten

Tabellen 25 und 26 geben den Überblick der möglichen Antezedenten für die Rollen LOCATION bzw. LOCATED_OBJECT.

¹¹ Im Korpus wurde VP-Koordination jedoch als *DNI/lexical licensor* annotiert.

Antezedenten für LOCATION					
		Anapher			
		verlinkt		Abstand	
		gesamt	%	0	%
Antezedent	PLINE	1	20.0	1	100
	REF	1	20.0	1	100
	GOAL	3	60.0	3	100
	Alle	5	100	5	100

Tabelle 25: Antezedenten für LOCATION (Anapher). **Verlinkt**: Anzahl der Verlinkungen der Anapher mit dem jeweiligen Antezedenten (**gesamt**) (in %: relativ zu allen Verlinkungen dieser Anapher). **Abstand**: Abstand (in dazwischenliegenden Frames) zwischen der Anapher und dem Antezedenten (in %: relativ zu allen Verlinkungen von LOCATION mit dem jeweiligen Antezedenten). **Alle**: obengenannte Parameter für Bindungsfälle von LOCATION insgesamt.

Antezedenten für LOCATED_OBJECT															
		Anapher													
		VP-Koordination								Andere					
		verlinkt		Abstand						verlinkt		Abstand			
		gesamt	%	0	%	1	%	2	%	gesamt	%	0	%	1	%
Antez.	SM	18	100	14	77.8	2	11.1	2	11.1	21	84.0	20	95.2	1	4.8
	LO	-	-	-	-	-	-	-	-	3	12.0	1	33.3	2	66.7
	PERC	-	-	-	-	-	-	-	-	1	4.0	1	100	-	-
	Alle	18	100	14	77.8	2	11.1	2	11.1	25	100	22	88.0	3	12.0

Tabelle 26: Antezedenten für LOCATED_OBJECT. Berechnung der gleichen Parameter wie in der Tabelle 25 für **VP-Koordination** bzw. andere Phänomene (**Andere**) von *constructional licensor* getrennt.

LOCATED_OBJECT

Obwohl das Frame-Element LOCATED_OBJECT im Korpus den Wörtern ganz unterschiedlichen semantischen Typen (Gebäude, Räume etc.) zugewiesen wurde, war jedoch in 100 % der Fälle nur LOCATED_OBJECT des semantischen Typs *Person* ausgelassen (siehe Tabelle 27). Dementsprechend konnte diese Rolle mit den Rollen des gleichen semantischen Typs SELF_MOVER (71), LOCATED_OBJECT (72) oder PERCEIVER (73) verlinkt werden.

(71) [**Wir SELF_MOVER**] biegen wieder nach links ab und erreichen **LOCATED_OBJECT 0** den Haupteingang der Sparkasse.

(72) Folgen Sie dem Weg immer geradeaus, bis [**Sie LOCATED_OBJECT**], unten angekommen **LOCATED_OBJECT 0**, auf einen Platz treffen.

(73) Unten angekommen **LOCATED_OBJECT 0**, sehen [**Sie PERCEIVER**] einen großen Würfel, der ein Symbol für Infothek oder Information darstellt.

Das Weglassen von **LOCATED_OBJECT** kommt in zwei typischen Konstruktionen vor: Zum Einen ist dies die VP-Koordination, bei der das Subjekt im zweiten Teilsatz weggelassen wird (41.9 % aller NI-Fälle von **LOCATED_OBJECT** (18 Mal)) (74). Zum anderen ist das der Partizipsatz (48.8 % aller NI-Fälle von **LOCATED_OBJECT** (21 Mal)), welcher ausschließlich durch die Partizipformen der Verben *ankommen* (z.B. *unten angekommen*) oder *stehen* (z.B. *am Gebäude 325 stehend*) gebildet wird. Bei Partizipsätzen kann die Anapher an die Rolle mit dem semantischen Typ *Person* (sofern vorhanden) des übergeordneten Satzes gebunden werden (75).

(74) [**Sie SELF_MOVER**] laufen einmal um die Ecke und stehen **LOCATED_OBJECT 0** dann direkt vor dem Eingang.

(75) Im größeren Raum angekommen **LOCATED_OBJECT 0**, folgen [**Sie SELF_MOVER**] dem Wegweiser.

Die anderen NI-Fälle von **LOCATED_OBJECT** kommen in 2.3 % wegen der Passivkonstruktion (1 Mal) (76), in 4.7 % wegen der Infinitivkonstruktion (2 Mal) (77) und in 2.3 % wegen Substantivierung des Verbs (1 Mal) (78) vor.

(76) Das Ziel, die Bushaltestelle, **ist erreicht** **LOCATED_OBJECT 0**.

(77) **Um** zur Sparkasse **zu gelangen** **LOCATED_OBJECT 0**...

(78) Folgen Sie dem Weg vorbei an einem Parkhaus **bis zum Erreichen** **LOCATED_OBJECT 0** einer großen Straße.

LOCATION

Wie schon oben erwähnt, konnte die Rolle **LOCATION** nur in 31.3 % ihrer NI-Fälle (5 Mal) verlinkt werden. In (79) ist ein Beispiel der Verlinkung von **LOCATION** mit der Rolle **GOAL**

angegeben. Die Lage des *Ziels* befindet sich dort am Haus mit der Nummer 144, was im Frame SELF_MOTION, ausgelöst durch das Verb *gehen*, die Funktion vom GOAL erfüllt. Deswegen wurde nicht realisierte LOCATION an GOAL *Hausnummer 144* gebunden.

(79) Gehen Sie immer weiter geradeaus am Marktplatz vorbei bis zur [**Hausnummer 144 GOAL**]. Sie haben Ihr Ziel **LOCATION 0** erreicht.

Das Fehlen des Antezedenten in 68.7 % der NI-Fälle (11 Mal) ist dadurch bedingt, dass die Nennung der Lage des Objektes wegen seiner Auffälligkeit und Einzigartigkeit unnötig ist. Dabei ist in 10 Fällen zusätzlich entweder die Richtung (DIRECTION) (80) oder Distanz (DISTANCE) (81) angegeben.

(80) Der Raum ist LOCATION 0 [rechts DIRECTION] und auf der Tür steht 110.

(81) Das Ziel befindet sich LOCATION 0 [wenige Meter weiter DISTANCE].

4.3.3. Semantischer Typ der Antezedenten

In Tabelle 27 sind die semantischen Typen der Antezedenten für LOCATED_OBJECT und LOCATION dargestellt. Der semantische Typ der Antezedenten für die Rolle LOCATED_OBJECT ist immer eine Person. Antezedenten für LOCATION sind Straße, Gebäude und Orte.

Semantischer Typ des Antezedenten				
	LOCATED_OBJECT		LOCATION	
	gesamt	%	gesamt	%
Person	43	100	-	-
Gebäude	-	-	1	20.0
Straße	-	-	1	20.0
Ort			3	60.0
Σ	43	100	5	100

Tabelle 27: **gesamt**: Anzahl der semantischen Typen der Antezedenten für LOCATED_OBJECT und LOCATION des Frames LOCALISATION (in **%**: relativ zu der Anzahl aller Verlinkungen der jeweiligen Rolle). Σ : semantische Typen der Antezedenten für jede Rolle insgesamt.

4.3.4. Verteilung der NI-Fälle auf die Verben des Frames

Tabelle 28 zeigt die Verteilung der nicht realisierten LOCATED_OBJECT und LOCATION auf die Verben des Frames LOCALISATION. Hier ist zu sehen, dass ca. die Hälfte (48.8 % (21 Mal)) der NI-Fälle von LOCATED_OBJECT auf das Verb *kommen* bzw. *ankommen* entfallen. Bei LOCATION wurden 37.5 % (6 Mal) der nicht realisierten Rollen bei dem häufigsten Verb *sich befinden* ausgelassen.

Verteilung der NI-Fälle von LOCATED_OBJECT und LOCATION auf die Verben des Frames LOCALISATION										
	Verbvorkommen		LOCATED_OBJECT ¹²				LOCATION			
	gesamt	%	NI	%	verlinkt	%	NI	%	verlinkt	%
<i>sich befinden</i>	72	22.9	2	2.8	2	100	6	8.3	2	33.3
<i>kommen</i>	66	21.0	21	31.8	21	100	3	4.5	1	33.3
<i>sein</i>	52	16.6	1	1.9	1	100	3	5.8	0	0.0
<i>stehen</i>	32	10.2	6	18.8	6	100	-	-	-	-
<i>erreichen</i>	31	9.9	7	22.6	7	100	-	-	-	-
<i>stoßen</i>	14	4.5	-	-	-	-	-	-	-	-
<i>treffen</i>	10	3.2	1	1.0	1	100	-	-	-	-
<i>liegen</i>	10	3.2	1	10.0	1	100	1	10.0	1	100
<i>finden</i>	10	3.2	-	-	-	-	2	20.0	1	50.0
<i>gelangen</i>	13	4.1	3	23.1	3	100	-	-	-	-
<i>landen</i>	1	0.3	1	100	1	100	-	-	-	-
<i>enden</i>	1	0.3	-	-	-	-	-	-	-	-
<i>anschließen</i>	1	0.3	-	-	-	-	-	-	-	-
<i>beginnen</i>	1	0.3	-	-	-	-	1	100	0	0.0
Alle	314	100	43	13.7	43	100	16	5.1	5	31.3

Tabelle 28: **Verbvorkommen (gesamt)**: zeigt für jedes Verb des Frames LOCALISATION die Anzahl seiner Vorkommen im Korpus (in %: relativ zu dem Gesamtvorkommen aller Verben in diesem Frame). **NI**: Anzahl der NI-Fälle einer Rolle bei dem jeweiligen Verb (in %: relativ zum Gesamtvorkommen des jeweiligen Verbs), **verlinkt**: Anzahl der gefundenen Antezedenten für den jeweiligen NI-Fall (in %: relativ zu der Anzahl aller NI-Fälle der Rolle bei dem jeweiligen Verb). **Alle**: obengenannte Parameter für Bindungsfälle jeder Rolle insgesamt.

¹² Die 25 Fälle der Kategorie *Andere* sind auf *ankommen* (18 Mal), *stehen* (3 Mal), *erreichen* (2 Mal) und *gelangen* (2 Mal) verteilt.

4.3.5. Zusammenfassung der wichtigsten Beobachtungen

LOCATED_OBJECT wurde im Korpus in 12.7 % (43 Mal) der Vorkommen dieser Rolle lokal nicht realisiert und während der Annotation in allen Fällen der Kategorie *DNI/constructional licensor* zugeordnet. Alle NI-Fälle dieser Rolle sind vom semantischen Typ *Person*. Die Auslassung ist überwiegend entweder durch einen Partizipsatz (48.8 %) oder VP-Koordination (41.9 %) bedingt. Als Antezedenten können für **LOCATED_OBJECT** nur die Rollen, die eine Person bezeichnen (**SELF_MOVER**, **LOCATED_OBJECT** oder **PERCEIVER**) auftreten. In 48.8 % der Fälle (21 Mal) wurde **LOCATED_OBJECT** beim Verb *kommen* bzw. *ankommen* ausgelassen.

LOCATION wurde in 4.1 % (16 Mal) der Vorkommen dieser Rolle weggelassen. Alle Fälle wurden in 100 % als *DNI/lexical licensor* klassifiziert, welche dann in 31.3 % (5 Mal) mit den Rollen des semantischen Typs *Straße*, *Gebäude* oder *Ort* verlinkt werden konnten. Komplette Auslassung der Rolle im ganzen Kontext in den restlichen 68.7 % (11 Mal) kann man dadurch erklären, dass es wegen der Einzigartigkeit des Objektes nicht nötig war. **LOCATION** wurde in 37.5 % ihrer NI-Fälle beim Verb *sich befindet* ausgelassen.

4.4. Relative_location

Dieser Abschnitt stellt die NI-Annotationsergebnisse der Core-Elemente **LOCATED_OBJECT** und **REFERENCE** des Frames **RELATIVE_LOCATION** dar.

4.4.1. Klassifikation

Tabelle 29 zeigt die Klassifikationsergebnisse der nicht realisierten Rollen **LOCATED_OBJECT** und **REFERENCE**. Die Rolle **LOCATED_OBJECT** wird im Korpus in 92.2 % ihrer Vorkommen ausgelassen. Alle diese NI-Fälle wurden der Kategorie *lexical licensor* zugeordnet und in allen 100 % verlinkt. Die NI-Häufigkeit der Rolle **REFERENCE** ist mit nur 8.2 % im Korpus relativ niedrig, wobei alle ihre Auslassungen als *DNI* klassifiziert wurden. In 7.8 % handelt es sich um lexikalische und in 0.4 % um konstruktionale Auslassungen. Für die Rolle **REFERENCE** wurde in 94.4 % aller ihrer NI-Fälle ein Antezedent gefunden.

Klassifikation der NI-Fälle von LOCATED_OBJECT und REFERENCE														
	Alle Rollen		interpretation				licensor				linking			
	gesamt	%	DNI	%	INI	%	lexical	%	constr.	%	verlinkt	%	KA	%
LO(LS)	218	3.5	201	92.2	0	0	201	92.2	0	0.0	201	100	0	0.0
REF	219	3.6	18	8.2	0	0	17	7.8	1	0.4	17	94.4	1	5.6

Tabelle 29: Klassifikationsergebnisse der NI-Fälle von LOCATED_OBJECT und REFERENCE. **Gesamt:** Anzahl der Vorkommen jeder Rolle (in %: relativ zur Gesamtzahl aller Rollen im Korpus). **Interpretation** bzw. **Licensor:** Anzahl der Vorkommen von *DNI/INI* bzw. *lexical/constructional licensor* (in %: relativ zu der Anzahl der Vorkommen der jeweiligen Rolle). **Linking:** Anzahl der verlinkten bzw. nicht verlinkten Rollen (in %: relativ zu allen NI-Fällen der jeweiligen Rolle).

4.4.2. Antezedenten

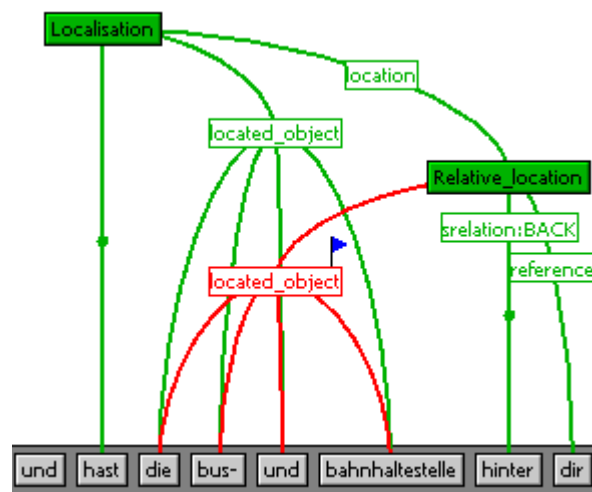
Tabelle 30 zeigt die Antezedenten der Rollen LOCATED_OBJECT und REFERENCE.

Antezedenten für LOCATED_OBJECT und REFERENCE															
		Anapher													
		LOCATED_OBJECT								REFERENCE					
		verlinkt		Abstand						Abstand					
		gesamt	%	0	%	1	%	3	%	gesamt	%	0	%	2	%
Antez.	LO (LS)	117	58.2	114	97.4	2	1.7	1	0.9	2	11.8	2	100	-	-
	SM	54	26.8	53	98.1	1	1.9	-	-	2	11.8	2	100	-	-
	PO	15	7.5	15	100	-	-	-	-	1	5.9	1	100	-	-
	GOAL	5	2.5	5	100	-	-	-	-	5	29.2	4	80	1	20.0
	PCONT	2	1.0	2	100	-	-	-	-	1	5.9	1	100	-	-
	PPOINT	2	1.0	2	100	-	-	-	-	2	11.8	2	100	-	-
	SOURCE	2	1.0	2	100	-	-	-	-	1	5.9	1	100	-	-
	PERC	2	1.0	2	100	-	-	-	-	-	-	-	-	-	-
	REF	1	0.5	1	100	-	-	-	-	1	5.9	1	100	-	-
	LOC	1	0.5	1	100	-	-	-	-	-	-	-	-	-	-
	PLINE	-	-	-	-	-	-	-	-	2	11.8	2	100	-	-
	Alle	201	100	197	98	3	1.5	1	0.5	17	100	16	94.1	1	5.9

Tabelle 30: Antezedenten für LOCATED_OBJECT und REFERENCE (Anapher). **Verlinkt:** Anzahl der Verlinkungen einer Anapher mit dem jeweiligen Antezedenten (**gesamt**) (in %: relativ zu allen Verlinkungen dieser Anapher). **Abstand:** Abstand (in dazwischenliegenden Frames) zwischen der Anapher und dem Antezedenten (in %: relativ zu allen Verlinkungen der jeweiligen Anapher mit dem jeweiligen Antezedenten). **Alle:** obengenannte Parameter für Bindungsfälle jeder Rolle insgesamt.

LOCATED_OBJECT

Wie es schon im Abschnitt 2.2.2 erwähnt wurde, ist der Frame `RELATIVE_LOCATION` ausschließlich ein eingebetteter Frame. Das heißt, er erfüllt die Rolle des Frame-Elementes eines anderen Frames, z.B. die Rolle des Frame-Elementes `LOCATION` in Grafik 7. Die Rolle `LOCATED_OBJECT` wurde im Korpus in ca. 98 % der Fälle nicht realisiert, weil die Konstituenten des Satzes, denen sie zugewiesen werden könnten, schon von den Rollen des einbettenden Frames eingenommen wurden. Deswegen wurde die NI-Rolle `LOCATED_OBJECT` mit den Rollen des einbettenden Frames verlinkt. Grafik 7 zeigt die Verlinkung der nicht realisierten Rolle `LOCATED_OBJECT` des Frames `RELATIVE_LOCATION` mit der Rolle `LOCATED_OBJECT` des einbettenden Frames `LOCALISATION`.



Grafik 7: Verlinkung der NI-Rolle `LOCATED_OBJECT` mit einer Rolle des einbettenden Frames

REFERENCE

Die lokale Realisierung von `REFERENCE` zu 91.8 % kann damit erklärt werden, dass der Frame `RELATIVE_LOCATION` wie im Abschnitt 2.2.2 erwähnt wurde durch lokale Präpositionen (*in, vor, nach* etc.) evoziert wird. Diese setzen die Verwendung eines Bezugswortes voraus, dem das Frame-Element `REFERENCE` zugewiesen wird (*vor der Treppe*).

Was die Verlinkung der NI-Fälle betrifft, so können unterschiedliche Rollen als Antezedenten für den nicht realisierten Frame-Element `REFERENCE` auftreten. In (82) ist das die Rolle `PERC_OBJECT` des Frames `VISUAL_PERCEPTION`. Hier wird die zweite Erwähnung von *eine Art Tankstelle* durch das Adverb *davor* ersetzt, indem die Rolle `REFERENCE` nicht realisiert

wird. Daher kann diese Rolle mit der ersten Erwähnung von *eine Art Tankstelle* (PERC_OBJECT) verlinkt werden.

- (82) Danach geht man nach links, bis man [**eine Art Tankstelle PERC_OBJECT**] sieht.
 Kurz davor **REFERENCE 0** geht man nach rechts.

4.4.3. Semantischer Typ der Antezedenten

Die semantischen Typen der Antezedenten der Rollen OBJECT und REFERENCE sind in Tabelle 31 zusammengefasst. Bei LOCATED_OBJECT fällt auf, dass in 52.3 % der verlinkten Fälle der Antezedent eine Person ist. In diesen Fällen wird die Lage der Person relativ zu einem Objekt genannt. Die kleinsten Gruppen sind *Andere* (4.5 %) und *Straße* (1.5 %). Die Gruppe *Andere* wird im Fall von LOCATED_OBJECT ausschließlich durch die Ausdrücke mit Bezeichnung des Endpunktes der Route als *Ziel*, ohne seine weitere Präzisierung, z.B. *das Ziel*, *Ihr Ziel*, *das Zielpunkt*, gebildet. Die semantischen Typen der Antezedenten von REFERENCE sind auf die Kategorien *Person*, *Gebäude*, *Straße* und *Objekt* verteilt.

Semantischer Typ des Antezedenten				
	LOCATED_OBJECT		REFERENCE	
	gesamt	%	gesamt	%
Person	105	52.3	5	29.4
Gebäude	26	12.9	4	23.5
Straße	3	1.5	2	11.8
Objekt	33	16.4	6	35.3
Ort	25	12.4	-	-
Andere	9	4.5	-	-
Σ	201	100	17	100

Tabelle 31: **gesamt**: Anzahl der semantischen Typen der Antezedenten von LOCATED_OBJECT und REFERENCE des Frames RELATIVE_LOCATION (**in %**: relativ zu der Anzahl aller Verlinkungen der jeweiligen Rolle). Σ : semantische Typen der Antezedenten für jede Rolle insgesamt.

4.4.4. Verteilung der NI-Fälle auf die Verben des Frames

Der Frame RELATIVE_LOCATION wird nicht durch Verben, sondern durch die Präpositionen evoziert.

4.4.5. Zusammenfassung der wichtigsten Beobachtungen

Die Besonderheit des Frames `RELATIVE_LOCATION` besteht darin, dass er ausschließlich durch die Präpositionen evoziert wird und nur als eingebetteter Frame vorkommt.

`LOCATED_OBJECT` wird in 92.2 % seiner Vorkommen nicht realisiert. Es handelt sich in allen Fällen um lexikalische Auslassungen, die in 100 % der Fälle verlinkt wurden. Der Antezedent für das nicht realisierte `LOCATED_OBJECT` ist fast immer (ca. 98 %) in dem einbettenden Frame zu suchen. In mehr als der Hälfte der NI-Fälle von `LOCATED_OBJECT` (58.2 %) ist dies das Frame-Element `LOCATED_OBJECT` des Frames `LOCALISATION`. Bei den semantischen Typen der Antezedenten ist das häufige Auftreten der Kategorie `Person` auffällig (52.3 %).

Das Besondere an der Rolle `REFERENCE` ist eine geringe Anzahl der `IN`-Fälle (8.2 %). Alle Auslassungen, mit Ausnahme von einem Fall, wurden als *DNI* klassifiziert. Ein Antezedent wurde in 94.4 % gefunden. Bei der Verteilung der NI-Fälle auf die semantischen Typen wurden keine besonders auffälligen Phänomene festgestellt.

4.5. Visual_perception

In diesem Abschnitt ist die Analyse der NI-Annotation für die Core-Elemente `PERCEIVER`, `PERC_OBJECT` und `DIRECTION` des Frames `VISUAL_PERCEPTION` dargestellt.

4.5.1. Klassifikation

Tabelle 32 gibt den Überblick der klassifizierten Fälle der Rollen `PERC_OBJECT`, `DIRECTION` und `PERCEIVER` des Frames `VISUAL_PERCEPTION`. Wie man sieht, wurde bei der Rolle `PERC_OBJECT` nur ein NI-Fall festgestellt, was 1.4 % seiner Rollenvorkommen entspricht. Die Rolle `DIRECTION` wurde in 70 % aller Rollenvorkommen (49 Mal) ausgelassen, wobei in 98 % (48 Mal) kein Antezedent gefunden wurde. Alle Auslassungen von `PERC_OBJECT` und `DIRECTION` wurden als *DNI* bzw. *lexical licensor* klassifiziert. Alle 15 NI-Fälle der Rolle `PERCEIVER` wurden den Kategorien *DNI/constructional licensor*¹³ zugeordnet. Bei der Auswertung der Annotation wurde das syntaktische Phänomen VP-Koordination als häufigere

¹³ Im Korpus wurde VP-Koordination jedoch als *DNI/lexical licensor* annotiert

Ursache der Auslassung dieser Rolle (80 % aller NI-Fälle von PERCEIVER) getrennt von den anderen Fällen von *constructional licensor* ausgewertet.

Klassifikation der NI-Fälle von PERC_OBJECT, DIRECTION und PERCEIVER														
	alle Rollen		interpretation				licensor				linking			
	gesamt	%	DNI	%	INI	%	lexical	%	constr.	%	verlinkt	%	KA	%
PO	72	1.2	1	1.4	0	0	1	1.4	0	0	1	100	0	0
DIR(VP)	70	1.1	49	70.0	0	0	49	70.0	0	0	1	2.0	48	98
	alle Rollen		interpretation				constructional licensor				linking			
	gesamt	%	DNI	%	INI	%	VP-K	%	Andere	%	verlinkt	%	KA	%
PERC	67	1.1	15	22.4	0	0	12	17.9	3	4.5	15	100	0	0

Tabelle 32: Klassifikationsergebnisse der NI-Fälle von PERC_OBJECT, DIRECTION und PERCEIVER. **gesamt**: Anzahl der Vorkommen jeder Rolle (in %: relativ zur Gesamtzahl aller Rollen im Korpus). **Interpretation** bzw. **Licensor**: Anzahl des Vorkommens von *DNI/INI* bzw. *lexical/constructional licensor* (in %: relativ zur Anzahl der Vorkommen der jeweiligen Rolle). **Linking**: Anzahl der verlinkten bzw. nicht verlinkten Rollen (in %: relativ zu allen NI-Fällen der jeweiligen Rolle). **VP-K** bzw. **Andere** bei PERCEIVER: Anzahl der Vorkommen von VP-Koordination bzw. anderer Phänomene von *constructional licensor* (in %: relativ zur Anzahl der Vorkommen der Rolle).

4.5.2. Antezedenten

Tabellen 33 bzw. 34 zeigen den Überblick der Antezedenten für die nicht realisierten PERC_OBJECT und DIRECTION bzw. PERCEIVER. Das Phänomen VP-Koordination wurde bei der Rolle PERCEIVER getrennt von anderen Phänomenen ausgewertet.

Antezedenten für PERC_OBJECT und DIRECTION									
		Anapher							
		PERC_OBJECT				DIRECTION			
		verlinkt		Abstand		verlinkt		Abstand	
		gesamt	%	0	%	gesamt	%	0	%
Antezedent	PO	1	100	1	100	-	-	-	-
	SRELATION:RIGHT	-	-	-	-	1	100	1	100
	Alle	1	100	1	100	1	100	1	100

Tabelle 33: Antezedenten für PERC_OBJECT und DIRECTION (Anapher). **Verlinkt**: Anzahl der Verlinkungen einer Anapher mit dem jeweiligen Antezedenten (**gesamt**) (in %: relativ zu allen Verlinkungen dieser Anapher). **Abstand**: Abstand (in dazwischenliegenden Frames) zwischen der Anapher und dem Antezedenten (in %: relativ zu allen Verlinkungen der jeweiligen Anapher mit dem jeweiligen Antezedenten). **Alle**: obengenannte Parameter für Bindungsfälle jeder Rolle insgesamt.

Antezedenten für PERCEIVER											
		Anapher									
		VP-Koordination				Andere					
		verlinkt		Abstand		verlinkt				Abstand	
		gesamt	%	0	%	gesamt	%	0	%	1	%
Antezedent	PERC	1	8.3	1	100	-	-	-	-	-	-
	SM	9	75.0	9	100	3	100	2	66.7	1	33.3
	LO(LS)	2	16.7	2	100	-	-	-	-	-	-
	Alle	12	100	12	100	3	100	2	66.7	1	33.3

Tabelle 34: Antezedenten für PERCEIVER. Berechnung der gleichen Parameter wie in der Tabelle 33 für **VP-Koordination** bzw. andere Phänomene (**Andere**) von *constructional licensor* getrennt.

PERCEIVER

Die Auslassung von PERCEIVER ist in 80 % aller NI-Fälle dieser Rolle durch das syntaktische Phänomen VP-Koordination bedingt, bei der das Subjekt (in diesem Fall eine Person) des zweiten Teilsatzes weggelassen wird. Abhängig davon, durch welchen Frame der erste Teilsatz repräsentiert war, wurde ein nicht realisierter PERCEIVER mit einer der drei Rollen, die eine Person bezeichnet (SELF_MOVER (83), PERCEIVER (84) oder LOCATED_OBJECT (85)), verlinkt.

(83) Dann drehen [**Sie SELF_MOVER**] sich nach rechts um 180 Grad und sehen **PERCEIVER 0** schon Raum 022 den Zip-Pool.

(84) [**Sie PERCEIVER**] schauen nach rechts in den Gang und sehen **PERCEIVER 0** Gebäude 346.

(85) [**Sie LOCATED_OBJECT**] stehen mit dem Rücken zu 325 und sehen **PERCEIVER 0** 305.

Die restlichen 20 % (3 Mal) des *constructional licensor* von PERCEIVER wurden durch die Auslassung des Agens in den Passivkonstruktionen hervorgerufen. In allen drei Fällen wurde die Anapher mit der Rolle SELF_MOVER verlinkt (85).

(85) Betritt [**man SELF_MOVER**] einen großen Parkplatz links, ist ein Gartenrestaurant zu sehen **PERCEIVER 0**.

PERC_OBJECT

Die Rolle PERC_OBJECT wurde zu 98.6 % realisiert. Der einzige NI-Fall (1.4 %) ist in (84) dargestellt. Hier ist *Gebäude 346* gleichzeitig das Objekt des Schauens und des Sehens.

(86) Sie schauen **PERC_OBJECT 0** nach rechts in den Gang und sehen [**Gebäude 346 PERC_OBJECT**].

DIRECTION

Frame-Element DIRECTION war in 70 % aller Rollenvorkommen von DIRECTION nicht realisiert, von denen es nur in 2 % (1 Mal), und zwar mit dem Nicht-Core-Element SRELATION:RIGHT des Frames RELATIVE_LOCATION verlinkt werden konnte (87). Hier haben beide Objektbezeichnungen synonymische Bedeutung, weil sich der Eingang zur Unibibliothek im Gebäude 368 befindet. Deswegen kann die Richtungsangabe *rechts von dir* auch dem Frame VISUAL_PERCEPTION zugewiesen werden.

(87) Wenn du 367 erreicht hast, ist das Gebäude 368 direkt [**rechts von dir SRELATION:RIGHT**] und du siehst **DIRECTION 0** schon den blauen Eingang zur Unibibliothek.

In den restlichen 98 % (48 Mal) blieb DIRECTION ungebunden. Der Grund dafür war in 62.5 % die Beschreibung der Lage des visuellen Objektes durch die Rolle LOCATION, bei der die Richtungsangabe implizit ist (30 Mal) (88) und in 37.5 % die Unnötigkeit der Richtungsangabe wegen der Einzigartigkeit und Auffälligkeit des Objektes (18 Mal) (in (89) *Cafe Rossi* und *McDonald's-Restaurant*).

(88) Dann geht man geradeaus weiter, bis man [**auf der rechten Seite LOCATION**] die Tür 326 sieht **DIRECTION 0**.

- (89) Sie sind nun in der Rohrbacherstraße und sollten das Cafe Rossi und ein McDonald's Restaurant sehen **DIRECTION 0** können.

4.5.3. Semantischer Typ der Antezedenten

Aus Tabelle 35 ist ersichtlich, dass semantischer Typ der Antezedenten für jede Rolle immer gleich ist.

Semantischer Typ des Antezedenten						
	PERCEIVER		PERC_OBJECT		DIRECTION	
	gesamt	%	gesamt	%	gesamt	%
Person	15	100	-	-	-	-
Gebäude	-	-	1	100	-	-
Richtung	-	-	-	-	1	100
Σ	15	100	1	100	1	100

Tabelle 35: **gesamt**: Anzahl der semantischen Typen der Antezedenten für nicht realisierte Core-Elemente des Frames VISUAL_PERCEPTION (in %: relativ zu der Anzahl aller Verlinkungen der jeweiligen Rolle). Σ : semantische Typen der Antezedenten für jede Rolle insgesamt.

4.5.4. Verteilung der NI-Fälle auf die Verben des Frames

Aus Tabelle 36 sieht man, dass die meisten (89.2 % / 58 Mal) nicht realisierte Rollen von VISUAL_PERCEPTION auf das Verb *sehen* entfallen, was sich durch häufigeres Vorkommen (88 % aller Verben des Frames VISUAL_PERCEPTION) dieses Verbs erklären lässt. Die restlichen 10.8 % (7 Mal) sind auf die anderen Verben dieses Frames verteilt.

Verteilung der NI-Fälle von PERCEIVER, PERC_OBJECT und DIRECTION auf die Verben des Frames VISUAL_PERCEPTION														
	Verben- vorkommen		PERCEIVER ¹⁴				PERC_OBJECT				DIRECTION			
	gesamt	%	NI	%	ver- linkt	%	NI	%	ver- linkt	%	NI	%	ver- linkt	%
sehen	59	88	14	23.7	14	100	-	-	-	-	44	74.6	1	2.2
erkennen	5	7.5	-	-	-	-	-	-	-	-	3	60.0	0	0.0
schauen	1	1.5	-	-	-	-	1	100	1	100	-	-	-	-
erblicken	1	1.5	1	100	1	100	-	-	-	-	1	100	0	0.0
entdecken	1	1.5	-	-	-	-	-	-	-	-	1	100	0	0.0
Alle	67	100	15	22.4	15	100	1	100	1	100	49	73.1	1	2.2

Tabelle 36: **Verbenvorkommen (gesamt)**: zeigt für jedes Verb des Frames VISUAL_PERCEPTION die Anzahl seiner Vorkommen im Korpus (in %: relativ zu dem Gesamtvorkommen aller Verben in diesem Frame). **NI**: Anzahl der NI-Fälle einer Rolle bei dem jeweiligen Verb (in %: relativ zum Gesamtvorkommen des jeweiligen Verbs), **verlinkt**: Anzahl der gefundenen Antezedenten für den jeweiligen NI-Fall (in %: relativ zu der Anzahl aller NI-Fälle der Rolle bei dem jeweiligen Verb). **Alle**: obengenannte Parameter für Bindungsfälle jeder Rolle insgesamt.

4.5.5. Zusammenfassung der wichtigsten Beobachtungen

PERCEIVER wurde in insgesamt 22.4 % der Vorkommen dieser Rolle weggelassen und in allen Fällen als *DNI/constructional licensor* klassifiziert. Der semantische Typ von **PERCEIVER** ist immer eine Person. Diese Rolle wurde im Korpus überwiegend (80 %) durch das syntaktische Phänomen der VP-Koordination nicht realisiert. Was die Bindung von **PERCEIVER** betrifft, ist die Tendenz wie bei anderen nicht realisierten Rollen, die eine Person bezeichnen gleich: **PERCEIVER** wird an die am nächsten liegende realisierte Rolle mit dem semantischen Typ **PERSON** (**SELF_MOVER**, **LOCATED_OBJECT** oder **PERCEIVER**) gebunden.

PERC_OBJECT wurde im Korpus zu 98.6 % lokal realisiert. Es wurde nur ein NI-Fall dieser Rolle beobachtet. Alle 49 Auslassungen der Rolle **DIRECTION** wurden als *DNI* markiert, konnten aber in 98 % der Fälle nicht verlinkt werden. Auf die explizite Nennung der Richtung kann beim Frame **VISUAL_PERCEPTION** verzichtet werden, weil entweder die Blickrichtung implizit mit der Rolle **LOCATION** ausgedrückt wird oder ihre Erwähnung aufgrund der Auffälligkeit des zu sehenden Objektes überflüssig ist.

¹⁴ Die drei Fälle der Kategorie *Andere* der Rolle **PERCEIVER** entfallen auf das Verb *sehen*.

Der Frame VISUAL_PERCEPTION wird im Korpus in 88 % der Fälle durch das Verb *sehen* ausgelöst. Dementsprechend entfallen die meisten der NI-Fälle (89.2 %) auf dieses Verb.

4.6. Zusammenfassung der Ergebnisse der quantitativen Analyse

4.6.1. Klassifikation

Tabelle 37 fasst die Ergebnisse der NI-Klassifikation aller Core-Elemente zusammen.

Klassifikation der NI-Fälle aller Rollen														
	Alle Rollen		interpretation				licensor				linking			
	gesamt	%	DNI	%	INI	%	lexical	%	constr.	%	verlinkt	%	KA	%
SOURCE	944	15.4	590	62.5	105	11.1	695	73.6	0	0	303	43.6	392	56.4
GOAL	959	15.6	364	38	257	26.8	621	64.8	0	0	272	43.8	349	56.2
DIR(SM)	992	16.2	405	40.7	10	1.1	415	41.8	0	0	17	4.1	398	95.9
LOC	388	6.3	16	4.1	0	0	16	4.1	0	0	5	31.3	11	68.7
LO (RL)	218	3.5	201	92.2	0	0	201	92.2	0	0	201	100	0	0
REF	219	3.6	18	8.2	0	0	17	7.8	1	0.4	17	94.4	1	5.6
PO	72	1.2	1	1.4	0	0	1	1.4	0	0	1	100	0	0
DIR(VP)	70	1.1	49	70	0	0	49	70	0	0	1	2	48	98
	Alle Rollen		interpretation				constructional licensor				linking			
	gesamt	%	DNI	%	INI	%	VP-K	%	Andere	%	verlinkt	%	KA	%
SM	936	15.2	241	25.7	0	0	130	13.9	111	11.8	232	96.3	9	3.7
LO (LS)	339	5.5	43	12.7	0	0	18	5.3	25	7.4	43	100	0	0
PERC	67	1.1	15	22.4	0	0	12	17.9	3	4.5	15	100	0	0

Tabelle 37: Klassifikationsergebnisse der NI-Fälle aller Rollen. **gesamt**: Anzahl der Vorkommen jeder Rolle (in %: relativ zur Gesamtzahl aller Rollen im Korpus). **Interpretation** bzw. **Licensor**: Anzahl des Vorkommens von *DNI/INI* bzw. *lexical/constructional licensor* (in %: relativ zur Anzahl der Vorkommen der jeweiligen Rolle). **Linking**: Anzahl der verlinkten bzw. nicht verlinkten Rollen (in %: relativ zu allen NI-Fällen der jeweiligen Rolle). **VP-K** bzw. **Andere** bei SM, LO und PERCEIVER : Anzahl der Vorkommen von Koordination bzw. anderer Phänomene von *constructional licensor* (in %: relativ zur Anzahl der Vorkommen der Rolle).

4.6.2. Antezedenten

Tabellen 38-41 zeigen die Antezedenten der Core-Elemente der Frames SELF_MOTION, LOCALISATION, RELATIVE_LOCATION und VISUAL_PERCEPTION.

Antezedenten der Rollen von SELF_MOTION											
		Anapher									
		SOURCE		GOAL		DIRECTION		SELF_MOVER (ges.)		SELF_MOVER (ohne K)	
		gesamt	p	gesamt	p	gesamt	p	gesamt	p	gesamt	p
Antez.	GOAL	123	0.406	29	0.107	-	-	-	-	-	-
	LOC	45	0.149	67	0.246	-	-	-	-	-	-
	SOURCE	27	0.089	34	0.125	-	-	-	-	-	-
	LO (LS)	19	0.063	36	0.132	-	-	63	0.2	58	0.5
	REF	9	0.030	7	0.026	-	-	-	-	-	-
	PAREA	8	0.026	18	0.066	-	-	-	-	-	-
	PLINE	6	0.020	30	0.110	-	-	-	-	-	-
	PCONT	5	0.017	5	0.018	-	-	-	-	-	-
	PPOINT	4	0.013	6	0.022	-	-	-	-	-	-
	GOAL (UV)	2	0.007	-	-	-	-	-	-	-	-
	VEHICLE	1	0.003	-	-	-	-	-	-	-	-
	PO	1	0.003	3	0.246	-	-	2	0.1	2	0.1
	F:LS	28	0.092	14	0.125	-	-	-	-	-	-
	F:RL	17	0.056	19	0.132	-	-	-	-	-	-
	F:VP	4	0.013	4	0.026	-	-	-	-	-	-
	F:SM	3	0.010	-	-	-	-	-	-	-	-
	READING	1	0.003	-	-	-	-	-	-	-	-
	DIR (SM)	-	-	-	-	16	0.9	-	-	-	-
	SRIGHT	-	-	-	-	1	0.1	-	-	-	-
SM	-	-	-	-	-	-	167	0.7	42	0.4	
Σ		303	1.0	272	1.0	17	1.0	232	1	102	1.0

Tabelle 38: Antezedenten der Core-Elemente von SELF_MOTION (Anapher). **gesamt**: Anzahl der Verlinkungen einer Anapher mit dem jeweiligen Antezedenten (**p**: Wahrscheinlichkeit der Verlinkung einer Anapher mit dem jeweiligen Antezedenten). Bei SELF_MOVER: die Gesamtzahl der Verlinkungen mit dem jeweiligen Antezedenten (**ges.**), Anzahl der Verlinkungen mit dem jeweiligen Antezedenten ohne Berücksichtigung der Fälle von VP-Koordination (**ohne K**) Σ: obengenannte Parameter für Bindungsfälle der Rolle insgesamt.

Antezedenten der Rollen von LOCALISATION							
		Anapher					
		LOCATION		LOCATED_OBJECT (ges.)		LOCATED_OBJECT (ohne K)	
		gesamt	p	gesamt	p	gesamt	p
Antez.	PLINE	1	0.2	-	-	-	-
	REF	1	0.2	-	-	-	-
	GOAL	3	0.6	-	-	-	-
	SM	-	-	39	0.9	21	0.84
	LO(LS)	-	-	3	0.08	3	0.12
	PER	-	-	1	0.02	1	0.04
	Σ	5	1.0	43	1.0	25	1.0

Tabelle 39: Antezedenten der Core-Elemente von LOCALISATION (Anapher). **gesamt**: Anzahl der Verlinkungen einer Anapher mit dem jeweiligen Antezedenten (**p**: Wahrscheinlichkeit der Verlinkung einer Anapher mit dem jeweiligen Antezedenten). Bei LOCATED_OBJECT: die Gesamtzahl der Verlinkungen mit dem jeweiligen Antezedenten (**ges.**), Anzahl der Verlinkungen mit dem jeweiligen Antezedenten ohne Berücksichtigung der Fälle von VP-Koordination (**ohne K**) Σ: obengenannte Parameter für Bindungsfälle der Rolle insgesamt.

Antezedenten der Rollen von RELATIVE_LOCATION					
		Anapher			
		LOCATED_OBJECT		REFERENCE	
		gesamt	p	gesamt	p
Antez.	LO (LS)	117	0.582	2	0.118
	SM	54	0.269	2	0.118
	PO	15	0.075	1	0.059
	GOAL	5	0.025	5	0.294
	PCONT	2	0.010	1	0.059
	PPOINT	2	0.010	2	0.118
	SOURCE	2	0.010	1	0.059
	PERC	2	0.010	-	-
	REF	1	0.005	1	0.059
	LOC	1	0.005	-	-
	PLINE	-	-	2	0.118
	Σ	201	1.0	17	1.0

Tabelle 40: Antezedenten der Core-Elemente von RELATIVE_LOCATION (Anapher). **gesamt**: Anzahl der Verlinkungen einer Anapher mit dem jeweiligen Antezedenten (**p**: Wahrscheinlichkeit der Verlinkung einer Anapher mit dem jeweiligen Antezedenten). Σ: obengenannte Parameter für Bindungsfälle der Rolle insgesamt.

Antezedenten der Rollen von VISUAL_PERCEPTION									
		Anapher							
		PERC_OBJECT		DIRECTION		PERCEIVER (ges.)		PERCEIVER (ohne K)	
		gesamt	p	gesamt	p	gesamt	p	gesamt	p
Antez.	PO	1	1	-	-	-	-	-	-
	SRIGHT	-	-	1	1	-	-	-	-
	PERC	-	-	-	-	1	0.07	-	-
	SM	-	-	-	-	12	0.80	3	1
	LO(LS)	-	-	-	-	2	0.13	-	-
	Σ	1	1.0	1	1.0	15	1.0	3	1.0

Tabelle 41: Antezedenten der Core-Elemente von VISUAL_PERCEPTION (Anapher). **gesamt**: Anzahl der Verlinkungen einer Anapher mit dem jeweiligen Antezedenten (**p**: Wahrscheinlichkeit der Verlinkung einer Anapher mit dem jeweiligen Antezedenten). Bei PERCEIVER: die Gesamtzahl der Verlinkungen mit dem jeweiligen Antezedenten (**ges.**), Anzahl der Verlinkungen mit dem jeweiligen Antezedenten ohne Berücksichtigung der Fälle von VP-Koordination (**ohne K**) Σ: obengenannte Parameter für Bindungsfälle der Rolle insgesamt.

4.6.3. Semantischer Typ der Antezedenten

Tabelle 42 zeigt die semantischen Typen der Antezedenten aller Core-Elemente.

Semantischen Typen der Antezedenten aller Core-Elemente														
	Strasse	%	Ort	%	Objekt	%	Gebäude	%	Person	%	Richtung	%	Andere	%
SOUCE	128	55.4	71	40.8	43	38.1	45	39.5	-	-	-	-	16	28.1
GOAL	97	42.0	75	43.1	31	27.4	37	32.4	-	-	-	-	32	56.1
DIR(S)	-	-	-	-	-	-	-	-	-	-	17	94.4	-	-
LOC	1	0.4	3	1.7	-	-	1	0.9	-	-	-	-	-	-
LO(RL)	3	1.3	25	14.4	33	29.2	26	22.8	105	26.2	-	-	9	15.8
REF	2	0.9	-	-	6	5.3	4	3.5	5	1.2	-	-	-	-
PO	-	-	-	-	-	-	1	0.9	-	-	-	-	-	-
DIR(VP)	-	-	-	-	-	-	-	-	-	-	1	5.6	-	-
SM	-	-	-	-	-	-	-	-	232	58.0	-	-	-	-
LO(LS)	-	-	-	-	-	-	-	-	43	10.8	-	-	-	-
PERC	-	-	-	-	-	-	-	-	15	3.8	-	-	-	-
Σ	231	100	174	100	113	100	114	100	400	100	18	100	57	100

Tabelle 42: **gesamt**: Anzahl der semantischen Typen der Antezedenten für die nichtrealisierte Core-Elemente aller Frames (**in %**: relativ zur Gesamtzahl aller Vorkommen des jeweiligen semantischen Typs). Σ: obengenannte Parameter für jeden semantischen Typ insgesamt.

4.6.4. Verteilung der nicht realisierten Core-Elemente auf die Verben der Frames

Tabellen 43-49 zeigen die Verteilung der nicht realisierten Core-Elemente auf die Verben der Frames, zu denen sie gehören.

Verteilung der NI-Fälle der Rolle SOURCE auf die Verben des Frames SELF_MOTION						
	Verbenvorkommen		SOURCE			
	gesamt	%	NI	%	verlinkt	%
<i>gehen (laufen, weitergehen, weiterlaufen)</i>	446	47.8	376	84.3	184	48.9
<i>abbiegen (einbiegen)</i>	136	14.6	69	50.7	23	33.3
<i>überqueren (durchgehen, hindurchgehen durchqueren, durchlaufen)</i>	90	9.6	90	100	42	46.7
<i>vorbeikommen (vorbeilaufen, vorbeigehen, passieren)</i>	84	9	59	70.2	13	22
<i>zulaufen (zugehen, zusteuern)</i>	39	4.2	33	84.6	10	30.3
<i>sich wenden (sich drehen)</i>	34	3.6	18	52.9	11	61.1
<i>sich halten</i>	28	3	25	89.2	12	48
<i>verlassen (rausgehen, rauskommen)</i>	24	2.6	1	4.2	1	100
<i>starten (losgehen, loslaufen)</i>	21	2.2	2	9.5	0	0
<i>betreten (reingehen, eintreten, hineingehen)</i>	14	1.5	8	57.1	2	25
<i>nehmen</i>	8	0.9	5	62.5	2	40
<i>bleiben</i>	5	0.5	5	100	1	20
<i>umlaufen</i>	3	0.3	2	66.6	1	50
<i>verwenden</i>	2	0.2	2	100	1	50
Σ	934	100	695	74.4	303	43.6

Tabelle 43: **Verbenvorkommen (gesamt)**: zeigt für jedes Verb des Frames SELF_MOTION die Anzahl seiner Vorkommen im Korpus (in %: relativ zu dem Gesamtvorkommen aller Verben in diesem Frame). **NI**: Anzahl der NI-Fälle von SOURCE bei dem jeweiligen Verb (in %: relativ zum Gesamtvorkommen des jeweiligen Verbs), **verlinkt**: Anzahl der gefundenen Antezedenten für den jeweiligen NI-Fall (in %: relativ zu der Anzahl aller NI-Fälle von SOURCE bei dem jeweiligen Verb). Σ : obengenannte Parameter für Bindungsfälle von SOURCE insgesamt.

Verteilung der NI-Fälle der Rolle GOAL auf die Verben des Frames SELF_MOTION						
	Verbenvorkommen		GOAL			
	gesamt	%	NI	%	verlinkt	%
<i>gehen (laufen, weitergehen, weiterlaufen)</i>	446	47.8	296	66.4	174	58.8
<i>abbiegen (einbiegen)</i>	136	14.6	60	44.1	17	28.3
<i>überqueren (durchgehen, hindurchgehen durchqueren, durchlaufen)</i>	90	9.6	90	100	36	40.0
<i>vorbeikommen (vorbeilaufen, vorbeigehen, passieren)</i>	84	9.0	59	70.2	9	15.2
<i>Zulaufen (zugehen, zusteuern)</i>	39	4.2	0	0.0	0	0.0
<i>sich wenden (sich drehen)</i>	34	3.6	31	91.2	2	6.4
<i>sich halten</i>	28	3.0	28	100	19	67.8
<i>verlassen (rausgehen, rauskommen)</i>	24	2.6	22	91.7	2	9.0
<i>starten (losgehen, loslaufen)</i>	21	2.2	17	80.9	4	23.5
<i>betreten (reingehen, eintreten, hineingehen)</i>	14	1.5	1	7.1	0	0.0
<i>nehmen</i>	8	0.9	8	100	4	50.0
<i>bleiben</i>	5	0.5	4	80.0	4	100
<i>umlaufen</i>	3	0.3	3	100	0	0.0
<i>verwenden</i>	2	0.2	2	100	1	50.0
Σ	934	100	621	66.5	272	43.8

Tabelle 44: **Verbenvorkommen (gesamt)**: zeigt für jedes Verb des Frames SELF_MOTION die Anzahl seiner Vorkommen im Korpus (**in %**: relativ zum Gesamtvorkommen aller Verben in diesem Frame). **NI**: Anzahl der NI-Fälle von GOAL bei dem jeweiligen Verb (**in %**: relativ zum Gesamtvorkommen des jeweiligen Verbs), **verlinkt**: Anzahl der gefundenen Antezedenten für den jeweiligen NI-Fall (**in %**: relativ zu der Anzahl aller NI-Fälle von GOAL bei dem jeweiligen Verb). Σ : obengenannte Parameter für Bindungsfälle von GOAL insgesamt.

Verteilung der NI-Fälle der Rolle SELF_MOVER auf die Verben des Frames SELF_MOTION						
	VP-Koordination					
	Verbenvorkommen		SELF_MOVER			
	gesamt	%	NI	%	verlinkt	%
<i>gehen (laufen, weitergehen, weiterlaufen)</i>	446	47.8	72	16.1	72	100
<i>abbiegen (einbiegen)</i>	136	14.6	17	12.5	17	100
<i>überqueren (durchgehen, hindurchgehen durchqueren, durchlaufen)</i>	90	9.6	17	18.8	17	100
<i>vorbeikommen (vorbeilaufen, vorbeigehen, pas- sieren)</i>	84	9.0	6	7.1	6	100
<i>zulaufen (zugehen, zusteuern)</i>	39	4.2	6	15.4	6	100
<i>sich wenden (sich drehen)</i>	34	3.6	4	11.8	4	100
<i>sich halten</i>	28	3.0	4	14.3	4	100
<i>verlassen (rausgehen, rauskommen)</i>	24	2.6	1	4.2	1	100
<i>starten (losgehen, loslaufen)</i>	21	2.2	-	-	-	-
<i>betreten (reingehen, eintreten, hineingehen)</i>	14	1.5	2	14.3	2	100
<i>nehmen</i>	8	0.9	-	-	-	-
<i>bleiben</i>	5	0.5	-	-	-	-
<i>umlaufen</i>	3	0.3	1	33.3	1	100
<i>verwenden</i>	2	0.2	-	-	-	-
Σ	934	100	130	13.9	130	100

Tabelle 45: **Verbenvorkommen (gesamt)**: zeigt für jedes Verb des Frames SELF_MOTION die Anzahl seiner Vorkommen im Korpus (**in %**: relativ zum Gesamtvorkommen aller Verben in diesem Frame). **NI**: Anzahl der NI-Fälle von SELF_MOVER (nur VP-Koordination) bei dem jeweiligen Verb (**in %**: relativ zum Gesamtvorkommen des jeweiligen Verbs), **verlinkt**: Anzahl der gefundenen Antezedenten für den jeweiligen NI-Fall (**in %**: relativ zu der Anzahl aller NI-Fälle von SELF_MOVER bei dem jeweiligen Verb). Σ : obengenannte Parameter für Bindungsfälle von SELF_MOVER insgesamt.

Verteilung der NI-Fälle der Rolle SELF_MOVER auf die Verben des Frames						
SELF_MOTION						
	Andere					
	Verbenvorkommen		SELF_MOVER			
	gesamt	%	NI	%	verlinkt	%
<i>gehen (laufen, weitergehen, weiterlaufen)</i>	446	47.8	60	13.4	52	86.7
<i>abbiegen (einbiegen)</i>	136	14.6	16	11.8	16	100
<i>überqueren (durchgehen, hindurchgehen durchqueren, durchlaufen)</i>	90	9.6	9	10.0	9	100
<i>vorbeikommen (vorbeilaufen, vorbeigehen, passieren)</i>	84	9.0	4	4.8	4	100
<i>zulaufen (zugehen, zusteuern)</i>	39	4.2	3	7.8	3	100
<i>sich wenden (sich drehen)</i>	34	3.6	4	11.8	4	100
<i>sich halten</i>	28	3.0	2	7.1	2	100
<i>verlassen (rausgehen, rauskommen)</i>	24	2.6	4	16.7	4	100
<i>starten (losgehen, loslaufen)</i>	21	2.2	4	19.0	3	75.0
<i>betreten (reingehen, eintreten, hineingehen)</i>	14	1.5	1	7.1	1	100
<i>nehmen</i>	8	0.9	2	25.0	2	100
<i>bleiben</i>	5	0.5	2	40.0	2	100
<i>umlaufen</i>	3	0.3	-	-	-	-
<i>verwenden</i>	2	0.2	-	-	-	-
Σ	934	100	111	12.6	102	92

Tabelle 46: Berechnung der gleichen Parameter wie in der Tabelle 45 für andere Phänomene des *constructional licensor* (ohne VP-Koordination).

Verteilung der NI-Fälle der Rolle DIRECTION auf die Verben des Frames SELF_MOTION						
	Verbenvorkommen		DIRECTION			
	gesamt	%	NI	%	verlinkt	%
<i>Gehen (laufen, weitergehen, weiterlaufen)</i>	446	47.8	203	45.5	9	4.4
<i>abbiegen (einbiegen)</i>	136	14.6	13	9.6	0	0.0
<i>überqueren (durchgehen, hindurchgehen durchqueren, durchlaufen)</i>	90	9.6	83	92.2	2	2.4
<i>vorbeikommen (vorbeilaufen, vorbeigehen, passieren)</i>	84	9.0	34	40.5	4	11.8
<i>zulaufen (zugehen, zusteuern)</i>	39	4.2	26	66.7	0	0.0
<i>sich wenden (sich drehen)</i>	34	3.6	1	2.9	0	0.0
<i>sich halten</i>	28	3.0	3	7.9	0	0.0
<i>verlassen (rausgehen, rauskommen)</i>	24	2.6	20	83.3	0	0.0
<i>starten (losgehen, loslaufen)</i>	21	2.2	11	52.4	0	0.0
<i>betreten (reingehen, eintreten, hineingehen)</i>	14	1.5	9	64.2	0	0.0
<i>nehmen</i>	8	0.9	3	37.5	0	0.0
<i>bleiben</i>	5	0.5	5	100	2	40.0
<i>umlaufen</i>	3	0.3	2	66.7	0	0.0
<i>verwenden</i>	2	0.2	2	100	0	0.0
Σ	934	100	415	44.4	17	4.1

Tabelle 47: **Verbenvorkommen (gesamt)**: zeigt für jedes Verb des Frames SELF_MOTION die Anzahl seiner Vorkommen im Korpus (**in %**: relativ zum Gesamtvorkommen aller Verben in diesem Frame). **NI**: Anzahl der NI-Fälle von DIRECTION bei dem jeweiligen Verb (**in %**: relativ zum Gesamtvorkommen des jeweiligen Verbs), **verlinkt**: Anzahl der gefundenen Antezedenten für den jeweiligen NI-Fall (**in %**: relativ zur Anzahl aller NI-Fälle von DIRECTION bei dem jeweiligen Verb). Σ : oben genannte Parameter für Bindungsfälle von DIRECTION insgesamt.

Verteilung der NI-Fälle der Rollen LOCATED_OBJECT und LOCATION auf die Verben des Frames LOCALISATION										
	Verbenvorkommen		LOCATED_OBJECT				LOCATION			
	gesamt	%	NI	%	verlinkt	%	NI	%	verlinkt	%
<i>sich befinden</i>	72	22.9	2	2.8	2	100	6	8.3	2	33.3
<i>kommen</i>	66	21.0	21	31.8	21	100	3	4.5	1	33.3
<i>sein</i>	52	16.6	1	1.9	1	100	3	5.8	0	0.0
<i>stehen</i>	32	10.2	6	18.8	6	100	-	-	-	-
<i>erreichen</i>	31	9.9	7	22.6	7	100	-	-	-	-
<i>stoßen</i>	14	4.5	-	-	-	-	-	-	-	-
<i>treffen</i>	10	3.2	1	1.0	1	100	-	-	-	-
<i>liegen</i>	10	3.2	1	10.0	1	100	1	10.0	1	100
<i>finden</i>	10	3.2	-	-	-	-	2	20.0	1	50.0
<i>gelangen</i>	13	4.1	3	23.1	3	100	-	-	-	-
<i>landen</i>	1	0.3	1	100	1	100	-	-	-	-
<i>enden</i>	1	0.3	-	-	-	-	-	-	-	-
<i>anschließen</i>	1	0.3	-	-	-	-	-	-	-	-
<i>beginnen</i>	1	0.3	-	-	-	-	1	100	0	0.0
Σ	314	100	43	13.7	43	100	16	5.1	5	31.3

Tabelle 48: **Verbvorkommen (gesamt)**: zeigt für jedes Verb des Frames LOCALISATION die Anzahl seiner Vorkommen im Korpus (**in %**: relativ zu dem Gesamtvorkommen aller Verben in diesem Frame). **NI**: Anzahl der NI-Fälle einer Rolle bei dem jeweiligen Verb (**in %**: relativ zum Gesamtvorkommen des jeweiligen Verbs), **verlinkt**: Anzahl der gefundenen Antezedenten für den jeweiligen NI-Fall (**in %**: relativ zu der Anzahl aller NI-Fälle der Rolle bei dem jeweiligen Verb). Σ : obengenannte Parameter für Bindungsfälle jeder Rolle insgesamt.

Verteilung der NI-Fälle der Rollen PERCEIVER, PERC_OBJECT und DIRECTION auf die Verben des Frames VISUAL_PERCEPTION														
	Verben-vorkommen		PERCEIVER				PERC_OBJECT				DIRECTION			
	gesamt	%	NI	%	ver-linkt	%	NI	%	ver-linkt	%	NI	%	ver-linkt	%
<i>sehen</i>	59	88	14 ¹⁵	23.7	14	100	-	-	-	-	44	74.6	1	2.2
<i>erkennen</i>	5	7.5	-	-	-	-	-	-	-	-	3	60.0	0	0.0
<i>schauen</i>	1	1.5	-	-	-	-	1	100	1	100	-	-	-	-
<i>erblicken</i>	1	1.5	1	100	1	100	-	-	-	-	1	100	0	0.0
<i>entdecken</i>	1	1.5	-	-	-	-	-	-	-	-	1	100	0	0.0
Σ	67	100	15	22.4	15	100	1	100	1	100	49	73.1	1	2.2

Tabelle 49: **Verbenvorkommen (gesamt)**: zeigt für jedes Verb des Frames VP die Anzahl seiner Vorkommen im Korpus (**in %**: relativ zu dem Gesamtvorkommen aller Verben in diesem Frame). **NI**: Anzahl der NI-Fälle einer Rolle bei dem jeweiligen Verb (**in %**: relativ zum Gesamtvorkommen des jeweiligen Verbs), **verlinkt**: Anzahl der gefundenen Antezedenten für den jeweiligen NI-Fall (**in %**: relativ zu der Anzahl aller NI-Fälle der Rolle bei dem jeweiligen Verb). Σ: obengenannte Parameter für Bindungsfälle jeder Rolle insgesamt.

4.7. Vergleich der NI-Annotationsergebnisse zweier Annotatoren im Wegbeschreibungskorpus

Die manuelle Annotation ist immer eine subjektive Aufgabe. Deswegen wurde die im Rahmen dieser Arbeit durchgeführte Annotation des Wegbeschreibungskorpus (Annotator 1) für die Qualitätskontrolle mit der Annotation eines anderen Annotators (Annotator 2) verglichen. Dafür stand ein im Rahmen des Projektes „SightSee“ teilweise annotierter Wegbeschreibungskorpus zur Verfügung.

Es wurde am Beispiel von 50 Sätzen (88 NI-Fälle) untersucht, wie sehr sich die Annotationsergebnisse der beiden Annotatoren voneinander unterscheiden. Für jeden NI-Fall wurde die Übereinstimmung in folgenden Kategorien gemessen¹⁶:

- Typ der Interpretation
- Typ des licensors

¹⁵ Die drei Fälle der Kategorie *Andere* der Rolle PERCEIVER entfallen auf das Verb *sehen*.

¹⁶ Die Übereinstimmung in jedem NI-Fall wurde nur dann gemessen, wenn beiden Annotatoren eine Aussage über den Typ der Interpretation bzw. des licensors gemacht haben.

- verlinkt/nicht verlinkt
- im Fall der Verlinkung, ob der Antezedent gleich ist

Die Ergebnisse sind nachfolgend in den Tabellen 50, 51 und 52 dargestellt.

Tabelle 50 zeigt die Annotationsergebnisse beider Annotatoren in Hinsicht auf den Interpretationstyp. Aus der Tabelle ist ersichtlich, dass beide Annotatoren die gleiche Anzahl aller NI-Fälle als *DNI* (51.1 %) und in 26.1 % als *INI* klassifiziert haben. Damit liegt die Übereinstimmung bei 77.2 %, was einem Kappa-Koeffizient von 0.53 entspricht.

Typ der Interpretation							
		Annotator 2					
		DNI	%	INI	%	Σ	%
Annotator 1	DNI	45	51.1	17	19.3	62	70.4
	INI	3	3.4	23	26.1	26	29.6
	Σ	48	54.5	40	45.5	88	100

Tabelle 50: Übereinstimmungsmatrix beider Annotatoren mit Fokus auf den Interpretationstyp. Neben den absoluten Werten sind die Werte in Prozent angegeben.

Tabelle 51 stellt die Annotationsergebnisse zweier Annotatoren in Hinsicht auf den Typ des licensors dar. In diesem Fall ist der Grad der Übereinstimmung am höchsten: 97.8 % aller NI-Fälle haben beide Annotatoren gleich der Kategorie *lexical licensor* und 1.1 % der Kategorie *constructional licensor* zugeordnet. Insgesamt wurden damit 98.9 % aller Fälle gleich klassifiziert. Der Kappa-Koeffizient beträgt 0.66.

Typ des licensors							
		Annotator 2					
		lexical	%	constr.	%	Σ	%
Annotator 1	lexical	86	97.8	1	1.1	87	98.9
	constr.	0	0	1	1.1	1	1.1
	Σ	86	97.7	2	2.2	88	100

Tabelle 51: Übereinstimmungsmatrix beider Annotatoren mit Fokus auf den Typ des licensors. Neben den absoluten Werten sind die Werte in Prozent angegeben.

Tabelle 52 zeigt die Anzahl der Übereinstimmungen/Nichtübereinstimmungen beider Annotatoren in Hinsicht auf die Verlinkung der NI-Fälle, d.h. ob ein klassifizierter NI-Fall auch verlinkt wurde. Man sieht, dass 5.7 % aller NI-Fälle von beiden Annotatoren gleich verlinkt und 71.6 % nicht verlinkt wurden. Hier liegt der Kappa-Koeffizient mit insgesamt 77.3 % Übereinstimmung bei 0.25. Bei der Bestimmung des Antezedenten in 5 von beiden Annotatoren verlinkten Fällen lag die Übereinstimmung bei 80 % (4 Mal). Einmal haben die Annotatoren einen unterschiedlichen Antezedent für eine nicht realisierte Rolle gewählt.

		Annotator 2					
		verlinkt	%	KA	%	Σ	%
Annotator 1	verlinkt	5	5.7	19	21.6	24	27.3
	KA	1	1.1	63	71.6	64	72.7
	Σ	6	6.8	82	93.2	88	100

Tabelle 52: Übereinstimmungsmatrix beider Annotatoren mit Fokus auf die Verlinkung der NI-Fälle. Neben den absoluten Werten sind die Werte in Prozent angegeben.

Wenn man die Kappa-Koeffizienten aller Kategorien nach Landis und Koch (Landis, 1977) interpretiert, erhält man folgende Ergebnisse: In der Bestimmung des Typs des Licensors haben Annotatoren **gute** Übereinstimmung ($k = 0.66$), des Typs der Interpretation **mäßige** Übereinstimmung ($k = 0.53$) und in Hinsicht auf die Verlinkung der NI-Rollen **ordentliche** Übereinstimmung ($k = 0.25$).

4.8. Vergleich der Besonderheiten von NI-Annotation in SemEval und Wegbeschreibungskorpus

Um zu vergleichen, wie das Phänomen der NI in SemEval und im Wegbeschreibungskorpus repräsentiert wird, wurde für jeden Korpus am Beispiel von 50 Sätzen eine quantitative Auswertung der Daten in den Kategorien *Klassifikation*, *Abstand zwischen Anapher und Antezedenten*, *semantischer Typ der Antezedenten* und *Verteilung der NI-Fälle auf Frame-einführende Elemente* durchgeführt.

4.8.1. Klassifikation

Tabelle 53 zeigt die Klassifikationsergebnisse der NI-Fälle in SemEval und im Wegbeschreibungskorpus¹⁷.

Klassifikation der NI-Fälle														
	Alle Rollen		interpretation				licensor				linking			
	gesamt	%	DNI	%	INI	%	lexical	%	constr.	%	verlinkt	%	KA	%
SE	337	100	23	6.8	21	6.2	35	10.3	9	2.7	18	40.9	26	59.1
WB	303	100	76	25.1	23	7.6	99	32.7	0	0	40	40.4	59	59.6

Tabelle 53: Klassifikation der NI-Fälle der Rollen von SemEval und Wegbeschreibungskorpus. **Gesamt**: Anzahl aller Rollenvorkommen. **Interpretation** bzw. **licensor**: Anzahl der Vorkommen von *DNI/INI* bzw. *lexical licensor/constructional licensor* (**in %**: relativ zur Anzahl aller Rollenvorkommen des jeweiligen Korpus). **Linking**: Anzahl der verlinkten bzw. nicht verlinkten Rollen (**in %**: relativ zu allen NI-Fällen des jeweiligen Korpus).

Aus Tabelle 53 ist ersichtlich, dass beide Korpora eine relativ gleiche Anzahl der in ihnen vorkommenden Rollen haben. Jedoch wurde beim Wegbeschreibungskorpus mit 32.2 % im Vergleich zu 13 % bei SemEval eine höhere NI-Häufigkeit beobachtet. Die hohe NI-Häufigkeit des Wegbeschreibungskorpus kann durch die Struktur der Texte erklärt werden. In der Regel besteht eine Weginstruktion aus der Beschreibung der nacheinander folgenden Pfade, so dass sich einige Informationen mit benachbarten Pfaden überlappen können. Damit sich diese Informationen nicht mehrmals wiederholen, können ihre weiteren Erwähnungen ausgelassen werden. Für die Texte in SemEval ist keine feste Struktur charakteristisch. Die Frames können durch unterschiedliche Elemente evoziert werden (siehe Tabelle 56) und es wurden keine Gesetzmäßigkeiten in ihren Vorkommen festgestellt. Dort, wo ein Frame neue Information einführt, müssen ihre notwendigen Komponenten realisiert werden.

Trotz der hohen Anzahl der NI-Fälle ist es bei einer Wegbeschreibung wichtig, dass die ausgelassenen Elemente eines Satzes aus dem Kontext interpretiert werden, damit der Hörer die Weginstruktion verstehen und danach dem Weg folgen kann. Damit kann ein häufigeres Vorkommen der Kategorie *DNI* (21.5 % der Gesamtvorkommen aller Rollen bzw. 76.8 % aller NI-Fälle des Wegbeschreibungskorpus) als *INI* (7.6 % der Gesamtvorkommen aller Rollen bzw. 23.2 % aller NI-Fälle des Wegbeschreibungskorpus) erklärt werden. Bei den Texten in SemEval ist es zwar auch wichtig, dass der Leser sie verstehen kann, es kommt jedoch vor,

¹⁷ Alle Berechnungen beziehen sich hier und im Folgenden bei jedem Korpus nur auf 50 Sätze.

dass die Information komplett ausgelassen wird, weil sie überflüssig ist. Beispiele finden sich in den Rollen DEGREE des Frames AMBIENT_TEMPERATURE bei *cold* oder MEASUREMENT des Frames DIMENSION bei *high*. Bei SemEval ist die Verteilung der NI-Fälle auf die Kategorien *DNI* (6.8 % der Gesamtvorkommen aller Rollen bzw. 52.3 % aller NI-Fälle des SemEval) und *INI* (6.2 % der Gesamtvorkommen aller Rollen bzw. 47.7 % aller NI-Fälle des SemEval) fast gleichmäßig.

4.8.2. Abstand zwischen der Anapher und dem Antezedenten

Tabelle 54 zeigt den Abstand zwischen Anaphern und Antezedenten in SemEval und im Wegbeschreibungskorpus. Hier fällt auf, dass der Abstand beim Wegebeschreibungskorpus in 80 % aller Verlinkungen gleich 0 ist. Jeweils 10 % entfallen auf die Abstände 1 und 2. In den SemEval-Texten variiert der Abstand schon im kleinen Teil von 50 Sätzen vom 0 bis 52. Diese Besonderheit lässt sich erneut durch die Struktur der Texte erklären. Wie schon oben erwähnt werden in einer Weginstruktion die einander folgenden Pfade der Reihe nach beschrieben und Elemente für die Vermeidung der Wiederholung in den nebeneinander stehenden Pfaden weggelassen. Daher kann die fehlende Information in einem Pfad in der Regel durch die Elemente des direkt vorher- oder nachherstehenden rekonstruierbar werden. In SemEval ist es nicht so. Hier kann es sein, dass z.B. in einem Dialog auf die Information zurückgekommen wird, die viel früher eingeführt wurde und dementsprechend der Abstand zwischen der Anapher und dem Antezedenten groß ist.

Abstand zwischen den Anaphern und Antezedenten in beiden Korpora																						
	verlinkt		Abstand																			
	Ges.	%	0	%	1	%	2	%	3	%	4	%	5	%	6	%	7	%	35	%	52	%
SE	18	100	5	27.9	3	16.9	2	11.1	2	11.1	1	5.5	1	5.5	1	5.5	1	5.5	1	5.5	1	5.5
WB	40	100	32	80.0	4	10.0	4	10.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Tabelle 54: Abstand (in dazwischenliegenden Frames) zwischen der Anapher und dem Antezedenten in SemEval und Wegbeschreibungskorpus (in %: relativ zu allen Verlinkungen des jeweiligen Korpus).

4.8.3. Semantischer Typ der Antezedenten

Die semantischen Typen der Antezedenten des Wegbeschreibungskorpus und SemEval stellt Tabelle 55 dar. Hier sieht man, dass beim Wegbeschreibungskorpus alle semantischen Typen mit Ausnahme von *Richtung* vertreten sind. Antezedenten von SemEval sind die semantischen Typen *Person*, *Gebäude*, *Ort* und *Andere*. Interessant ist in diesem Fall die Kategorie

Andere, weil ihre Phänomene im Wegbeschreibungskorpus nicht beobachtet wurden. Diese enthält zum Einen den ganzen Satz (*This is not talk for a police-constable.*), der zweimal als Antezedent auftritt, zum Anderen den Indefinitartikel *a*.

Semantischer Typ des Antezedenten				
	Wegbeschreibungskorpus		SemEval	
	gesamt	%	gesamt	%
Person	8	20.0	7	38.9
Gebäude	5	12.5	7	38.9
Straße	4	10.0	-	-
Objekt	9	22.5	-	-
Ort	14	35.0	1	5.6
Andere	-	-	3	16.6
Σ	40	100	18	100

Tabelle 55: **gesamt**: Anzahl der semantischen Typen der Antezedenten für nicht realisierte Core-Elemente im Wegbeschreibungskorpus und SemEval (in **%**: relativ zu der Anzahl aller Verlinkungen des jeweiligen Korpus). Σ : semantische Typen der Antezedenten für jeden Korpus insgesamt.

4.8.4. Verteilung der NI-Fälle auf Frame-einführende Elemente

Tabelle 56 zeigt Frame-einführende Elemente des Wegbeschreibungskorpus und SemEval, bei denen NI-Fälle vorkommen. Wie bereits bekannt, treten im Wegbeschreibungskorpus nur Verben und Präpositionen als Frame-einführende Elemente auf. Daher kann eine Rolle nur bei einer dieser Wortarten ausgelassen werden. In Tabelle 56 sind es 83.8 % Verben und in 16.2 % Präpositionen. In SemEval können unterschiedliche Wortarten ein Frame evozieren. Deswegen gibt es hier mehr Wortarten, bei denen NI vorkommt. In Tabelle 56 kommen außer Verben und Präpositionen noch Adjektive und Substantive dazu. Die größte Gruppe, bei der Rollen ausgelassen werden, sind *Substantive*.

Verteilung der NI-Fälle auf Frame-einführende Elemente				
	Wegbeschreibungskorpus		SemEval	
	NI	%	gesamt	%
Verben	83	83.8	10	22.7
Präpositionen	16	16.2	2	4.6
Adjective	-	-	7	15.9
Substantive	-	-	25	56.8
Σ	99	100	44	100

Tabelle 56: Frame-einführende Elemente der beiden Korpora (in %: relativ zu der Gesamtzahl aller Frame-einführenden Elemente des jeweiligen Korpus)

4.8.5. Zusammenfassung der wichtigsten Beobachtungen

Der Wegbeschreibungskorpus enthält eine höhere Anzahl NI-Fälle als SemEval (32.2 % bzw. 13.0 %), wobei sie beim Wegbeschreibungskorpus mehr als *DNI*- als *INI* Fälle klassifiziert wurden (25.1 % bzw. 7.6 %). SemEval hat fast die gleiche Anzahl der *DNI*- und *INI*-Rollen (6.8 % bzw. 6.2 %). Alle Auslassungen des Wegbeschreibungskorpus wurden als *lexical licensor* klassifiziert. In SemEval wurden 2.7 % aller Rollenvorkommen der Kategorie *constructional licensor* zugeordnet. Bei der Verlinkung wurden NI-Fälle beider Korpora nur zur Hälfte an einen Antezedenten gebunden. Im Wegbeschreibungskorpus ist der Abstand zwischen dem Frame der Anapher und dem Frame des Antezedenten in 80 % aller verlinkten Fälle gleich 0. Bei SemEval variiert diese Zahl von 0 bis 52.

Was die semantischen Typen der Antezedenten betrifft, sind im Wegbeschreibungskorpus fast alle semantischen Typen, die im ganzen Korpus vorkommen, enthalten. Bei SemEval sind die Ergebnisse wegen der niedrigen Anzahl der NI-Fälle bzw. verlinkter Rollen weniger aussagekräftig. Es wird in diesem Fall eine Analyse mehrerer Daten benötigt.

Da in SemEval unterschiedliche Wortarten (Substantive, Verben, Adjektive etc.) Frames evokieren können, werden auch NI-Fälle auf diese Wortarten verteilt. Im Wegbeschreibungskorpus sind Frame-einführende Elemente und dementsprechend Elemente, bei denen NI-Fälle vorkommen, auf die Verben und Präpositionen beschränkt.

Die Analyse hat gezeigt, dass die Texteigenschaften der beiden Korpora, die in Abschnitten 2.1.1 und 2.2.1 beschrieben wurden, das Vorkommen der NI beeinflussen. Von diesen Eigenschaften hängen die Anzahl der NI-Fälle insgesamt, ihre Verteilung auf die Typen der Interpretation und des licensors sowie der Abstand zwischen der Anapher und dem Antezedenten ab. Auch die semantischen Typen der Antezedenten und die Verteilung der NI-Fälle auf Frame-einführende Elemente werden durch die Besonderheiten der Textgenres beeinflusst. Da die Texte des Wegbeschreibungskorpus und des SemEval über unterschiedliche Eigenschaften verfügen, wird dementsprechend das Phänomen der NI in den beiden Korpora unterschiedlich repräsentiert.

Zusammenfassung

Im Rahmen dieser Arbeit wurde das Phänomen der Null Instantiation im Wegbeschreibungskorpus untersucht. Das Ziel war zu erforschen, wie dieses Phänomen in den von Menschen erstellten Weginstruktionen repräsentiert wird und ob es beim Aufbau eines Systems, das die textuelle Wegbeschreibungen automatisch generiert, beachtet werden muss. Die Ergebnisse der Untersuchung können als Lehrverfahren für die automatische Annotation von Null Instantiation benutzt werden.

Zu Beginn dieser Arbeit wurden die behandelten Begriffe *Frame-Semantik*, *FrameNet*, *Null Instantiation*, *Anapher* und *Antezedent* erklärt. Frame-Semantik ist eine semantische Theorie, die prototypische Situationen und zu ihr gehörende Partizipanten darstellt. Die prototypische Situation wird als Frame und die Partizipanten als Frame-Elemente repräsentiert. Die Frame-Elemente werden den syntaktischen Argumenten eines Prädikats zugewiesen. Bei der Null Instantiation handelt es sich um die Auslassung notwendiger syntaktischer Argumente des Prädikats, die durch den Bezug auf ein Element des weiteren Kontext rekonstruierbar sind. In diesem Sinne können das ausgelassene Argument als eine Nullanapher und das Element des Kontexts auf das sich die Nullanapher bezieht als Antezedent betrachtet werden.

Die NI-Annotation des Wegbeschreibungskorpus, deren Ziel die Verlinkung der nicht realisierten semantischen Rollen mit den Antezedenten des weiteren Kontexts war, wurde nach dem Vorbild des mit NI bereits annotierten SemEval-Korpus durchgeführt. Beide Frame-annotierte Korpora wurden vorgestellt und verglichen. Beim Ersten handelt es sich um Weginstruktionen, beim Zweiten um einen Text aus dem Belletristik Genre. Unterschiedliche Textsorten haben unterschiedliche Eigenschaften. So wurde festgestellt, dass für die Belletristik die Verwendung einer großen Vielfalt unterschiedlicher sprachlicher Mittel, semantischer Typen, Satzstrukturen und Satzlängen charakteristisch ist. Die Texte des Wegbeschreibungskorpus weisen in allen diesen Aspekten dagegen mehr Regularitäten auf. Bei ihnen werden relativ ähnliche Satzstrukturen, semantische Typen etc. verwendet. Beim Vergleich der NI-Fälle des Wegbeschreibungskorpus mit SemEval wurde beobachtet, dass die unterschiedlichen Textsorten NI-Häufigkeit und Vorkommensart beeinflussen. Da die Textsorten der bei-

den Korpora unterschiedliche Eigenschaften aufweisen, wird bei ihnen auch NI unterschiedlich repräsentiert.

Bei der NI-Annotation des Wegbeschreibungskorpus wurde jede lokal nicht realisierte Rolle nach dem Interpretationstyp der Kategorien *DNI* oder *INI* und nach dem Typ des licensors der Kategorien *lexical licensor* oder *constructional licensor* zugeordnet und danach, sofern möglich, mit einem Antezedenten verlinkt. Anschließend wurde eine qualitative und quantitative Analyse der annotierten Daten durchgeführt.

Die Ergebnisse der Analyse haben gezeigt, dass es bei der Repräsentation der NI im Wegbeschreibungskorpus Gesetzmäßigkeiten gibt, die für die Konzeption eines Lernverfahrens hilfreich sein können. Dabei unterscheiden sich immer die Rollen untereinander darin, wie NI repräsentiert ist. Beispielsweise wird die Rolle `LOCATED_OBJECT` des Frames `RELATIVE_LOCATION` zu 92.2 % ihrer Gesamtvorkommen lokal nicht realisiert, dagegen wurde bei der Rolle `PERC_OBJECT` ein NI-Fall nur in 1.4 % und bei der Rolle `LOCATION` 4.1 % ihrer Gesamtvorkommen beobachtet. Jedoch gibt es Phänomene, die für das ganze Korpus charakteristisch sind. So wurden beispielsweise 84.5 % aller NI-Fälle als *DNI* und 87 % als *lexical licensor* klassifiziert. *Constructional licensor* kommt in 99.7 % bei den Rollen mit dem semantischen Typ *Person* vor. Außerdem wurde festgestellt, dass der Antezedent für eine nicht realisierte Rolle in ca. 80 % aller verlinkten Fälle in einem direkt nach- oder vorhergehenden Frame zu suchen ist.

Der Vergleich der NI-Annotation des Wegbeschreibungskorpus zweier Annotatoren, die für die Qualitätskontrolle durchgeführt wurde, hat Folgendes ergeben: die Übereinstimmung in der Bestimmung des Interpretationstyps liegt bei 77.2 %, in der Zuordnung beim Typ des licensors bei 98.9 % und in der Verlinkung der nicht realisierten Rollen bei 77.3 %. Aus diesem Grund kann geschlossen werden, dass NI-Fälle des Wegbeschreibungskorpus in der Art ihrer Vorkommen Regularitäten aufweisen, die sich trotz der Subjektivität von NI-Annotation im größten Teil der Fälle gleich annotieren lassen.

Literaturverzeichnis

Aljoscha Burchardt, Katrin Erk, Anette Frank, Andrea Kowalski, Sebastian Pado, and Manfred Pinkal (2006b). The SALSA corpus: a german corpus resource for lexical semantics. In Proceedings of LREC 2006, Genoa, Italy.

Aljoscha Burchardt, Katrin Erk, Anette Frank, Andrea Kowalski and Sebastian Pado (2006a). SALTO: A versatile multi-level annotation tool. In Proceedings of the 5th International Conference on Language Resources and Evaluation, Genoa, Italy, 22–28 May 2006, pp. 517–520.

Charles J. Fillmore, Christopher R. Johnson, and Miriam R.L. Petruck (2003). Background to FrameNet. *International Journal of Lexicography*, 16(3):235–250.

Collin F. Baker, Charles J. Fillmore, and John B. Lowe (1998). The Berkeley FrameNet project. In Proceedings of COLING-ACL, Montreal, Canada.

Collin F. Baker, Michael Ellsworth, and Katrin Erk (2007). SemEval-2007 Task 19: Frame Semantic Structure Extraction. In Proc. of SemEval.

Irene Heim (1993). Anaphora and Semantic Interpretation: A Reinterpretation of Reinhart's Approach. Ms., *SfS-Report 07-93*. Universität Tübingen.

Jacob Cohen (1960). Coefficient of Agreement for Nominal Scales Educational and Psychological Measurement, April, 1960, 20: 37-46.

Josef Ruppenhofer, Caroline Sporleder, Roser Morante, Collin Baker and Martha Palmer (2009). SemEval-2010 Task 10: Linking Events and Their Participants in Discourse. *The NAACL-HLT 2009 Workshop on Semantic Evaluations: Recent Achievements and Future Directions (SEW-09)*, Boulder, Colorado, USA, June 4.

Josef Ruppenhofer (2010). Annotation guidelines used for Semeval task 10 –“Linking Events and Their Participants in Discourse“.

J. Richard Landis and Gary G. Koch (1977). The measurement of observer agreement for categorical data. *Biometrics* 33:159-174.

Michael Roth, and Anette Frank (2009). A NLG-based application for walking directions. In *Companion Volume to the Proceedings of the 47th Annual Meeting of the Association for Computational Linguistics and the 4th International Joint Conference on Natural Language Processing of the Asian Federation of Natural Language Processing*, Singapore, 2-7.

Stephanie Schuldes, Michael Roth, Anette Frank, and Michael Strube (2009). Creating an Annotated Corpus for Generating Walking Directions. *Proceedings of the ACL 2009 Workshop on Language Generation and Summarisation/*, Singapore, pp. 72-76.