

FrameNet

Éva Mújdricza

Referat

Computerlexikographie PS

Dozentin: Claudia Kunze

Ruprecht-Karls-Universität Heidelberg

26.05.2008

Übersicht

- Theoretische Grundlagen:
 - Kasusgrammatik und semantische Rollen
- FrameNet:
 - Frame
 - Frame-Element
 - Lexikalische Einheit
 - Polysemie
 - Relationen im FrameNet
 - Satz- und Textannotation
 - Anwendungen
- FrameNet für das Deutsche
- Vergleich zwischen WordNet und FrameNet

Theoretische Grundlagen



- Forschung von Charles J. Fillmore ab den 60er Jahren: Kasusgrammatik, Frame-Semantik
- **Kasusgrammatik** (Case Grammar): Tiefenkasus oder Kasusrollen (~ semantische oder thematische Rollen) beschreiben sowohl die semantische als auch die syntaktische Struktur der Sätze.
- **Thematische Rollen:**
 - Agens: Rolle des belebten Urhebers oder Verursachers einer Handlung:
Philip öffnete die Tür.
 - Patiens: Rolle des Objektes, das von einer Handlung direkt betroffen ist:
Philip öffnete die Tür.; Franz schlug Klaus.
 - Instrumental: Rolle des unbelebten Verursachers
 - einer Handlung: *Der Wind öffnete die Tür.*
 - eines Objekts, mit dessen Hilfe eine Handlung vollzogen wird:
Philip öffnete die Tür mit dem Schlüssel.

Theoretische Grundlagen

- **Thematische Rollen:**

- Rezipient / Experiencer / Benefaktiv: Die Rolle des belebten Mitspielers, der von einem Zustand oder einer Handlung weniger als der Patiens betroffen ist: *Philip öffnete seiner Frau die Tür.*
- Lokativ: Rolle für den Ort der Handlung: *Philip nimmt ein Brötchen aus dem Korb.*
- ...

- Es wird ein Zusammenhang zwischen Kasusrollen, Bedeutungsstruktur und syntaktischer Struktur angenommen.
 - z.B.: Der Agens ist in der Regel das Subjekt des Satzes und steht im Nominativ.
- Aber: Die semantischen Rollen können nicht immer von der syntaktischen Struktur abgeleitet werden.
 - z.B.: "Peter_{AGENS} kocht." ↔ "Die Suppe_{PATIENS} kocht."

Theoretische Grundlagen

- **Frame-Semantik:** schemabasierter Ansatz zur Wissensrepräsentation
 - Annahme: Die Wortbedeutungen werden erst verstanden, wenn der Gebrauchsrahmen des Wortes bekannt ist. Der semantische Rahmen eines Wortes definiert die möglichen Partizipanten (~ Mitspieler) und Eigenschaften einer Situation, eines Ereignisses oder eines Objektes, die mit diesem Wort assoziiert werden.
 - semantische Prototypen: **Frames**

FrameNet: Frame

- FrameNet ist ein **framebasierter** Online-Thesaurus, der nach den Prinzipien der Fillmore'schen Frame-Semantik modelliert wurde.
 - ↔ Wortnetze wie WordNet oder GermaNet sind synsetbasiert.
- FrameNet wird am ICSI (International Computer Science Institute) in Berkeley seit 1997 entwickelt.
 - Jetzige Version: FrameNet II
- **Semantischer Frame** (semantic frame): „eine konzeptuelle Struktur zur Beschreibung eines spezifischen Situationstyps, Objekttyps oder Ereignistyps mit seinen jeweilig verknüpften Partizipanten und Eigenschaften.“ (Kunze & Lemnitzer 2007: 149 (Kap. 3.1))
- Bisher wurden ca. 800 Frames ausgearbeitet.
- Index der Frames:
http://framenet.icsi.berkeley.edu/index.php?option=com_wrapper&Itemid=113

FrameNet: Frame-Element

- Die Partizipanten und Eigenschaften werden im Frame als **Frame-Elemente** (frame elements = FE) dargestellt. Sie entsprechen den Kasusrollen in einem erweiterten Sinn.
 - Frame-Elemente werden nach ihren semantischen Rollen benannt.
- Frame-Elemente sind relativ zu ihrem Frame definiert, deswegen kann es vorkommen, dass es mehrere gleichnamige Frame-Elemente gibt, die aber zu verschiedenen Frames gehören (z.B. das Frame-Element AGENT).
- Beispiel: Frame-Elemente des Frames „Apply_heat”:
 - COOK: der Agens, der kocht oder etwas aufwärmt
 - FOOD: was aufgewärmt wird
 - HEATING_INSTRUMENT: womit man etwas aufwärmt
 - CONTAINER: worin sich FOOD während des Aufwärmens befindet
 - TEMPERATURE_SETTING, ...

FrameNet: Frame-Element

- Typen von **Frame-Elementen**:
 - **zentrale Frame-Elemente** (core frame elements): sind spezifisch für den jeweiligen Frame. Im Frame „Apply_heat”:
 - CONTAINER: worin sich FOOD während des Aufwärmens befindet
 - COOK: der Agens, der kocht oder etwas aufwärmt
 - FOOD: was aufgewärmt wird
 - HEATING_INSTRUMENT: womit man etwas aufwärmt
 - TEMPERATURE_SETTING: Grad der Temperaturerhöhung
 - **periphere Frame-Elemente** (periphere frame elements): Diese Frame-Elemente können auch in verschiedenen anderen Frames vorkommen. Beispiele für den Frame „Apply_heat”:
 - MEANS: *by throwing it into a pot of boiling water*
 - DURATION: *for three minutes*
 - MANNER: *gently*
 - MEDIUM: *in oil*

FrameNet: Lexikalische Einheiten

- Basiseinheit: **Lexical Unit (LU)** = **lexikalische Einheit**: ein Form-Bedeutungspaar
 - Im FrameNet gibt es mehr als 10000 LU-s
- Alle Möglichkeiten der syntaktischen und semantischen Valenz der LU werden mit Beispielen aus dem British National Corpus (BNC) oder aus Korpora des Linguistic Data Consortium (LDC) belegt.
 - FrameNet ist vollkommen **korpusbasiert** – jede Lesart wird durch ein Korpusbeispiel belegt.
 - British National Corpus (BNC): <http://info.ox.ac.uk/bnc/>, 100 Millionen Wörter der gesprochenen und geschriebenen englischen Sprache
 - Es gibt mehr als 135000 annotierte Sätze.
- Jede LU ist mit einem passenden semantischen Frame verknüpft.
- Eine LU **aktiviert** den semantischen Frame, dem sie zugeordnet wird.

Beispiel (1)

- z.B. Frame „Apply_heat”:
 - Frame-Elemente (FE): COOK, FOOD, HEATING_INSTRUMENT, CONTAINER, TEMPERATURE_SETTING, ...
 - aktivierende lexikalische Einheiten: *bake, blanch, boil, broil, brown, simmer, steam, ...*

- Beispiel:
 - *Matilde **fried** the catfish in a heavy iron skillet.*

fried aktiviert den Frame „Apply_heat” → Die Konstituenten des Satzes werden anhand der Frame-Elemente des aktivierten Frames klassifiziert.

- [_{COOK}*Matilde*] **fried** [_{FOOD}*the catfish*] [_{HEATING_INSTRUMENT}*in a heavy iron skillet*].

Beispiel (2)

- Die Frames können auch abstrakte Objekte oder Ereignisse abbilden.
- Beispiel: Frame „Change_position_on_a_scale”
 - Frame-Elemente: ITEM, ATTRIBUTE, INITIAL_VALUE, FINAL_VALUE, DIFFERENCE, ...
 - aktivierende lexikalische Einheiten: *decline, decrease, gain, increasingly, plummet, rise, shift, ...*
 - Colgate’s stock **rose** \$3.64 to \$49.94.
 - [ITEM Colgate’s stock] **rose** [DIFFERENCE \$3.64] [FINAL_VALUE to \$49.94].

FrameNet: Lexikalische Einheiten

- Die Bedeutung der LU ist in einer **Definition** (sense description) erfasst.
 - Definition stammt von
 - COD: aus dem Concise Oxford Dictionary
 - FN: von FrameNet-Lexikographen
 - Beispiel:
 - *boil*: FN: cook by immersing in boiling water
 - *roast*: COD: cook or be cooked by prolonged exposure to heat in an oven or over a fire
- Auch Mehrworteinheiten (*given name, Martin Luther King Day*) oder Phraseme (*middle of nowhere*) können als LU dienen.
 - **Lexem** (lexeme) ist ein Wort repräsentiert durch seine Grundform.
 - **Lemma** (lemma): Lemmata bestehen aus mindestens einem Lexem.

Polysemie

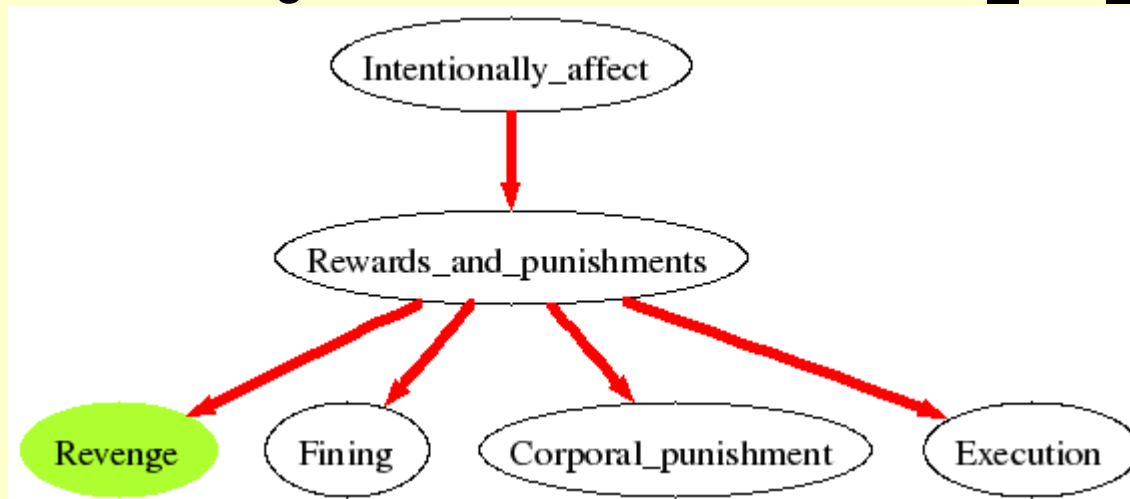
- Das Lemma wird mit mehreren Frames verknüpft: Für jede Lesart gibt es einen Frame.
 - Beispiel: **bake** wird mit den folgenden Frames verknüpft:
 - „Apply_heat”: *Michelle **baked** the potatoes for 45 minutes.*
 - LU-Definition: COD: cook by dry heat without direct exposure to a flame, typically in an oven
 - „Cooking_creation”: *Michelle **baked** her mother a cake for her birthday.*
 - LU-Definition: COD: cook food by dry heat without direct exposure to a flame, typically in an oven
 - „Absorb_heat”: *The potatoes have to **bake** for more than 30 minutes.*
 - LU-Definition: FN: to be cooked by dry heat, especially in an oven and as food

Relationen im FrameNet

- Im FrameNet sind Generalisierungen anhand hierarchischer Beziehungen möglich.
 - **Relationen zwischen Frames:** Jede Relation ist gerichtet.
 - Vererbung (inheritance relation)
 - Enthaltensein (subframe relation)
 - Präsupposition (using relation)
 - Perspektivierung (perspective_on relation)
 - (für andere Relationen siehe Ruppenhofer et al. 2007: 104–111)
 - **Relationen zwischen Frame-Elementen:**
 - Coreness Set
 - Requires
 - Excludes

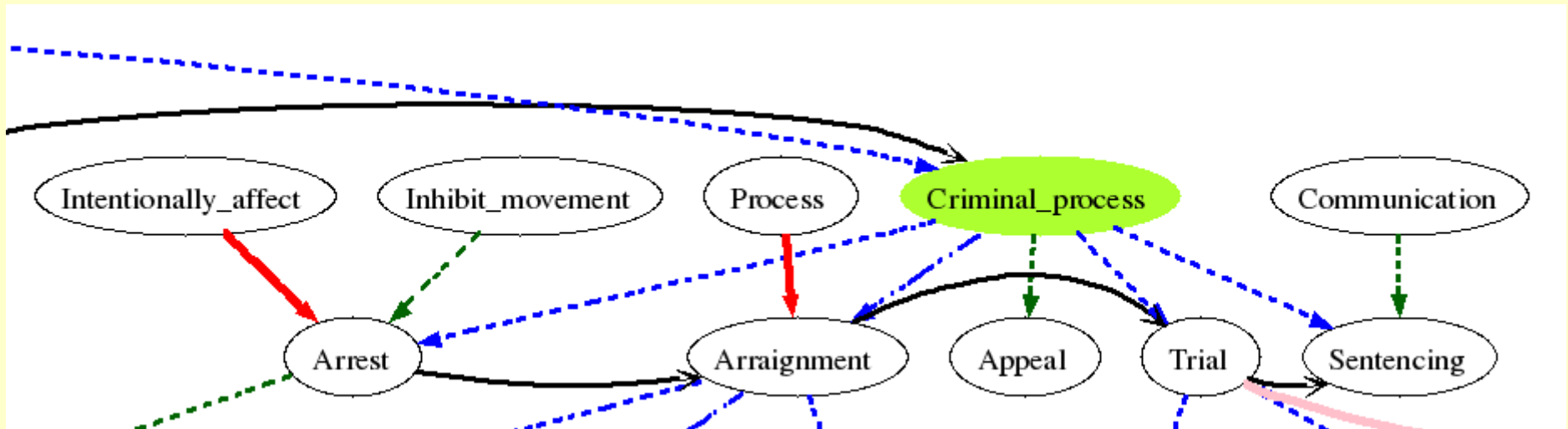
Relationen zwischen Frames

- **Vererbung (inheritance relation):**
 - Eine IS-A-Beziehung zwischen einem Kindframe und einem Elternframe: Der untergeordnete Frame ist eine Art übergeordneter Frame.
 - Die Vererbung ist vollkommen: Auch die Frame-Elemente der beiden Frames sind durch Vererbung miteinander verknüpft.
 - Mehrfachvererbung ist erlaubt: Ein Kindframe (und seine FE-s) kann über mehrere Elternframes verfügen.
- Beispiel: Frame „Revenge“ erbt vom Frame „Rewards_and_punishments“



Relationen zwischen Frames

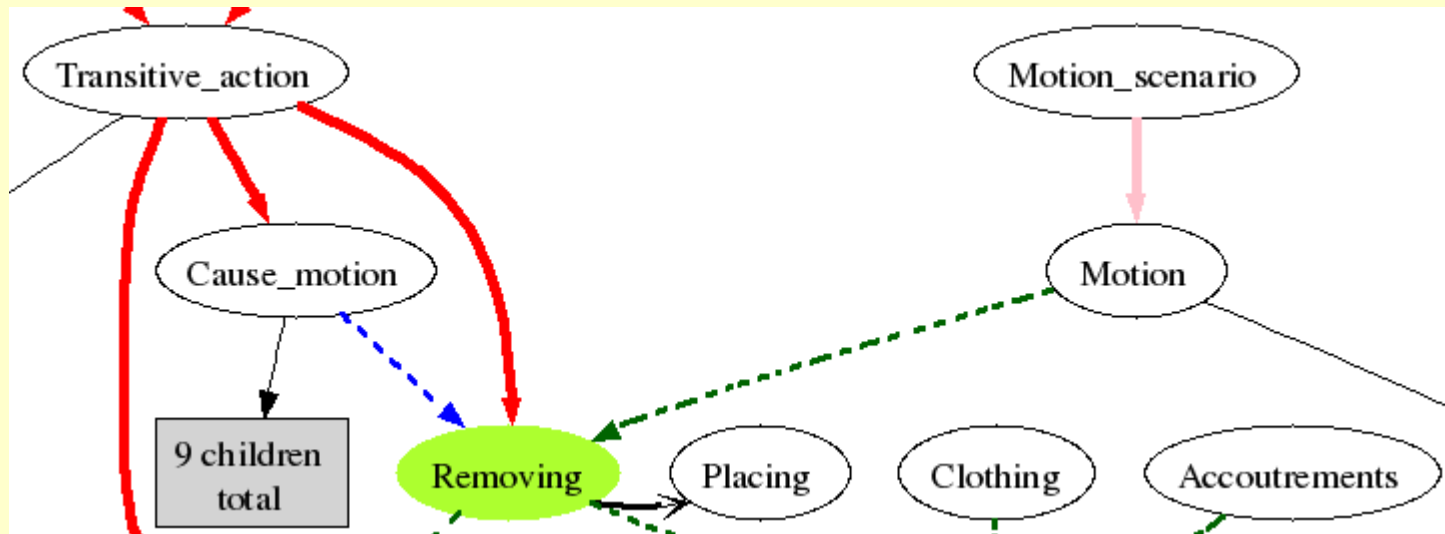
- **Enthaltensein (subframe relation):**
 - Ein komplexes Ereignis kann Subereignisse enthalten.
 - Beispiel: Frame „Criminal_process“ enthält die Subframes
 - „Arrest“
 - „Arraignment“
 - „Trial“
 - „Sentencing“



Relationen zwischen Frames

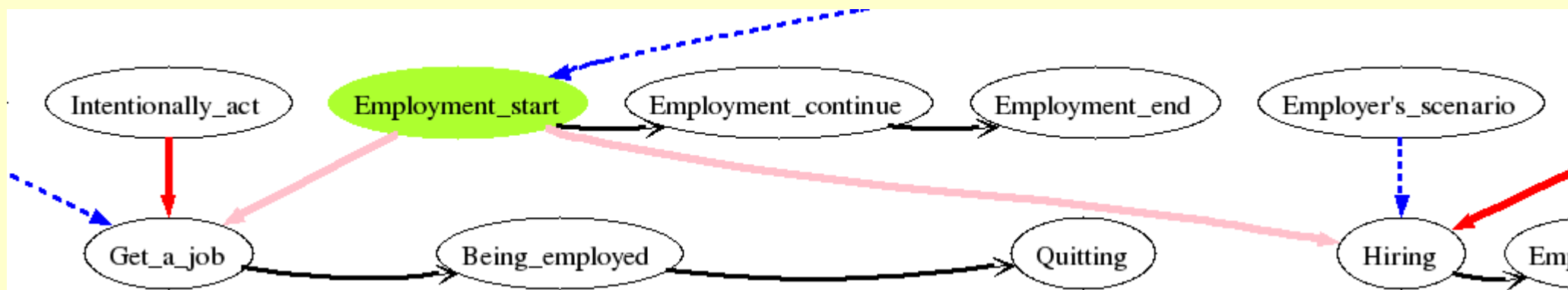
- **Präsupposition (using relation):**

- Untergeordnete Frames setzen übergeordnete Frames voraus. Diese Beziehung ist ähnlich zur Vererbung, aber sie ist nicht so streng. Wenn Frame B Frame A benutzt, dann muss Frame B nicht unbedingt alle entsprechende FE-s aus Frame A enthalten.
- Beispiel: Der Frame „Removing“ präsupponiert / verwendet den Frame „Motion“.



Relationen zwischen Frames

- **Perspektivierung (perspective_on relation):**
 - Ein neutraler Frame hat mehrere „perpektivierte“ untergeordnete Frames, die unterschiedliche Betrachtungsweisen einer Situation darstellen.
- Beispiel: Für den neutralen Frame „Employment_start“ werden die Frame-Elemente EMPLOYER und EMPLOYEE in den Frames „Hiring“ und „Get_a_job“ aus zwei Perspektiven beschrieben:
 - EMPLOYER stellt EMPLOYEE an oder
 - EMPLOYEE findet beim EMPLOYER Arbeit.



Relationen zwischen Frame-Elementen

- **Coreness Set** (CoreSet): wenn für eine LU nicht alle Frame-Elemente realisiert werden müssen
 - Beispiel: Die Frame-Elemente SOURCE, PATH und GOAL sind in vielen Frames, die eine Bewegung beschreiben, vorhanden. Nicht alle drei müssen aber in jedem Satz erscheinen:
 - *Fred went* [SOURCE *from Berkeley*] [PATH *across North America and the Atlantic Ocean*] [GOAL *to Paris*].
 - *Martha hiked* [SOURCE *from Berkeley*] [GOAL *to Oakland*].
 - *Elaine walked* [GOAL *to Monterey*].
 - *I saw Peter sneak* [PATH *past the guard*].
 - *Juan was walking* [SOURCE *out of the office*] *when I arrived*.

Relationen zwischen Frame-Elementen

- **Requires:** wenn die Realisierung eines Frame-Elements obligatorisch die Realisierung eines anderen Frame-Elements mit sich zieht
 - Beispiel: Im Frame „Attaching“ muss das Frame-Element GOAL unbedingt erscheinen, wenn das Frame-Element ITEM realisiert wird:
 - *The robbers tied* [_{ITEM}*Paul*] [_{GOAL}*to his chair*].
 - **The robbers tied* [_{ITEM}*Paul*].
- **Excludes:** wenn Frame-Elemente alternativ erscheinen, d.h. entweder wird das eine oder das andere Frame-Element im Satz realisiert
 - Beispiel: Im Frame „Evading“ die Frame-Elemente CAPTURE und PURSUER:
 - *Sheri'ffs officials said they apprehended a gang member after he evaded* [_{PURSUER}*them*].
 - *He had successfully evaded* [_{CAPTURE}*arrest*].

FrameNet: Annotation der Sätze

- Die Beispielsätze werden anhand der relevanten syntaktischen und semantischen Eigenschaften dreischichtig annotiert. Die drei Schichten:
 - **Frame-Element**: (frame element = FE): Eine Konstituente repräsentiert ein Frame-Element.
 - **Phrasentyp**: (phrase type = PT)
 - **grammatische Funktion**: (grammatical function = GF)
 - (Mehr über die dreischichtige Annotation: Fillmore 2007: 134 ff.)
- Beispiel: *brush* aktiviert den Frame „Filling”. (comp = complement)

<i>(Text)</i>	<i>Joe</i>	<i>brushed</i>	<i>the rolls</i>	<i>with butter.</i>
FE	AGENT		GOAL	THEME
PT	NP		NP	PP
GF	Ext		Obj	Comp

- Außer diesem Valenzmuster gibt es noch 5 andere syntaktische Umgebungen für die Realisierung von *brush* → noch 5 verschiedene Annotationen.

FrameNet: Annotation der Sätze

- Jede Schicht hat null, eins oder mehrere Labels.
- **Null Instantiation:** Für mögliche Frame-Elemente, die im Satz nicht erscheinen: (siehe auch Ruppenhofer et al. 2007: 33–36)
 - Beispiele:
 - *The result should be similar.* [DNI ENTITY_2] (Frame „Similarity”)
 - *I tried to put the toys back in.* [DNI GROUND] (Frame „Locative_relation”)
 - *Molly rarely eats alone.* [INI INGESTIBLES] (Frame „Ingestion”)
 - *Get even* [_{OFFENDER} *with her*] [_{INJURY} *for this*]. [CNI AVENGER] (Frame „Revenge”)
 - Abkürzungen:
 - DNI = definite null instantiation
 - INI = indefinite null instantiation
 - CNI = constructional null instantiation

FrameNet: Annotation der Sätze

- Ein Satz kann auf verschiedener Weise annotiert werden.
 - Z.B. wenn Konstituenten im Satz unterschiedliche Frames aktivieren:
 - *Matilde learned that Rufus had gone to Cairo.*
 - *learn* als FE GOAL aktiviert den Frame „Becoming_aware”
 - *go* als FE GOAL aktiviert den Frame „Self-motion”

FrameNet: Textannotation

- Von der jeweiligen LU (Target) ausgehend werden die aktivierten Frames annotiert. Wenn man auf die LU-s klickt, erscheinen die FE-e zum aktivierten Frame in Bezug auf den Satz.

http://framenet.icsi.berkeley.edu/index.php?option=com_wrapper&Itemid=84

Beispiel: The ART_{Craft} of change-ringing is PECULIAR_{Idiosyncrasy} to the ENGLISH_{People_by_origin} , and , LIKE_{Similarity} most English PECULIARITIES_{Idiosyncrasy} , UNINTELLIGIBLE_{Grasp} to the REST_{Rest} of the WORLD_{Political_locales} .

- The ART^{Target} [ACTIVITY of change-ringing] is peculiar to the English , and , like most English peculiarities , unintelligible to the rest of the world .
- [IDIOSYNCRASY The art of change-ringing] [is]^{Supp} PECULIAR^{Target} [ENTITY to the English] , and , like most English peculiarities , unintelligible to the rest of the world .
- [ENTITY_1 The art of change-ringing] is peculiar to the English , and , LIKE^{Target} [ENTITY_2 most English peculiarities] , unintelligible to the rest of the world .

FrameNet: Aufbau

- **FrameNet II** ist als relationale Datenbank aufgebaut.
 - **lexikalische Datenbank** enthält Informationen in erster Linie über Frames, Frame-Elementen, lexikalische Einheiten, Relationen zwischen diesen Elementen.
 - **Annotationsdatenbank** ist verknüpft mit der lexikalischen Datenbank und enthält das annotierte Korpus (BNC), in dem es Beispielsätze zu den einzelnen lexikalischen Einheiten gibt.

FrameNet: Anwendungen

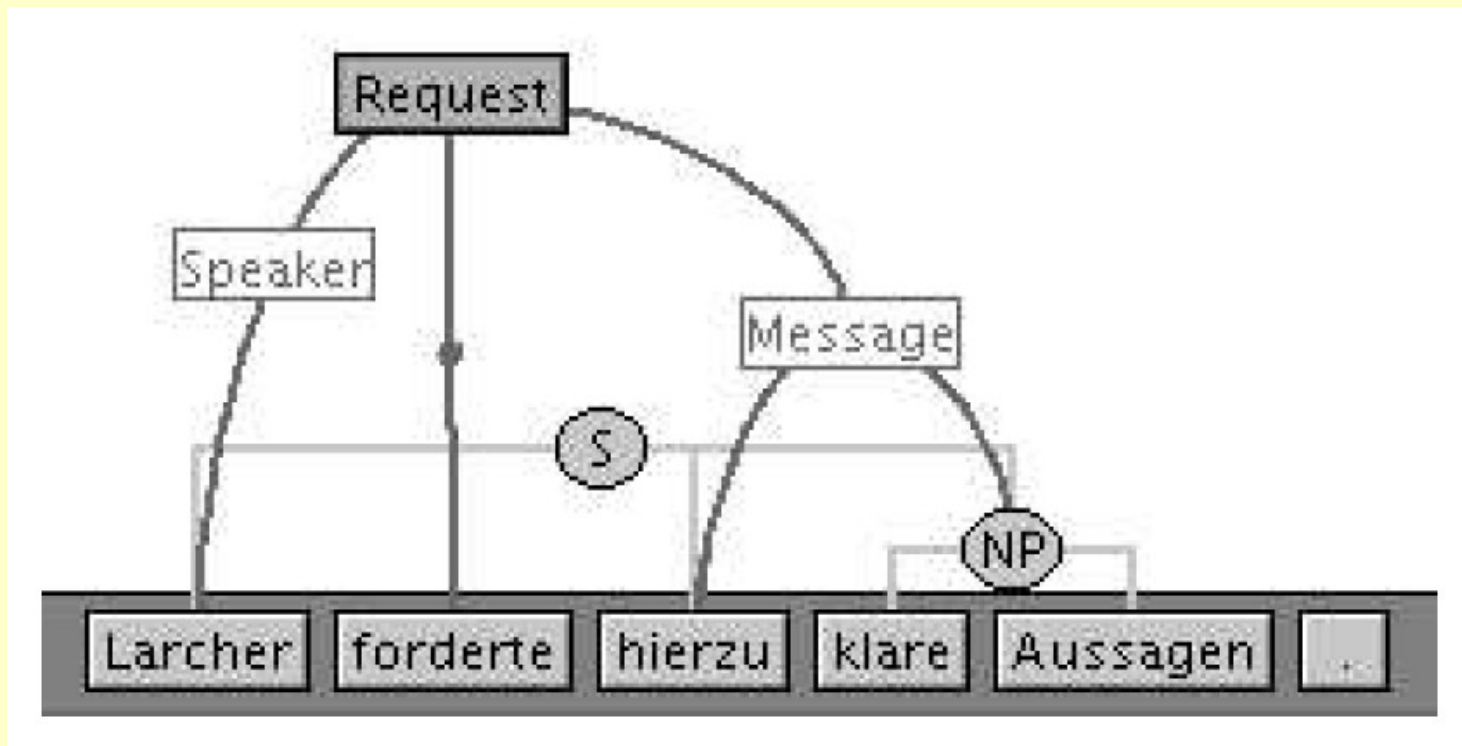
- Verwendung der Daten:
 - Die Datenbank kann im XML-Format exportiert werden.
 - Webbasierter Online-Browser – in für Menschen lesbarem Format
- Ein framebasiertes Lexikon enthält die Beziehungen zwischen bestimmten lexikalischen Bedeutungen und entsprechenden lexiko-syntaktischen Mustern. Jedes Muster wird mit dem aktivierten Frame bzw. mit Teilen des Frames verbunden. (Boas 2005: 141)
- FrameNet dient als Grundlage für den Aufbau paralleler Frame-Fragmente in anderen Sprachen.
 - multilinguale Verknüpfung

FrameNet für das Deutsche

- **German FrameNet:**
 - <http://gframenet.gmc.utexas.edu/>
 - University of Texas at Austin
 - für Verben, Nomen, Adjektive
- **SALSA** = Saarbrücken Lexical Semantics Acquisition Project
 - <http://www.coli.uni-saarland.de/projects/salsa/>
 - Universität Saarland
 - SALSA/TIGER Korpus mit manuell annotierter FrameNet-Framestruktur
 - SALSA bietet ein großes framebasiertes Lexikon für das Deutsche mit reicher semantischer und syntaktischer Annotation an.
 - dynamische semantische Analyse

FrameNet für das Deutsche

- TIGER/SALSA: Graphische Darstellung der Satzannotation mit SALTO ('The SALSA Tool'):
 - http://www.coli.uni-saarland.de/%7Epado/pub/papers/lrec04_erk.pdf (25.05.2008, 20:26)



Vergleich zwischen WordNet und FrameNet

- **Ziel** des WordNet: das mentale Lexikon abbilden ↔ FrameNet: typische Situationen, Objekte, Ereignisse und deren semantische und syntaktische Eigenschaften definieren.
- WordNet ist synsetbasiert. ↔ FrameNet ist framebasiert.
- **Basiseinheit** im FrameNet ist die lexikalische Einheit. ↔ Im WordNet das Synset.
- Frames können Relationen zwischen Lemmata unterschiedlicher **Wortarten** aufweisen. ↔ In Synsets des WordNet gibt es immer eine festgelegte Wortart.
- WordNet verfügt über vielfältige **semantische Relationen** zwischen Synsets und lexikalischen Einheiten. ↔ FrameNet legt Relationen zwischen Frames und Frame-Elementen fest.
- **Polysemie** wird im WordNet in verschiedenen Synsets zu einer lexikalischen Einheit dargestellt. ↔ Im FrameNet werden die Lesarten durch Verknüpfung der lexikalischen Einheiten zu mehreren Frames repräsentiert.
- Sowohl FrameNet als auch WordNet führt für jede Lesart **Beispielsätze** auf.

Quellen

- Baker, Collin F., Charles J. Fillmore & Beau Cronin (2003): The Structure of the FrameNet Database. *International Journal of Lexicography*. 16/3. S. 281–296. Link: <http://ijl.oxfordjournals.org.ubproxy.ub.uni-heidelberg.de/cgi/reprint/16/3/281.pdf> (Stand: 19.05.2008, 13:45)
- Boas, Hans C. (2005): From Theory to Practice: Frame Semantics and the Design of FrameNet. In: Langer, Stefan & Daniel Schnorbusch (Hrsg.): *Semantik im Lexikon*. Tübingen : Narr. S. 129–159. (= Tübinger Beiträge zur Linguistik; 479)
- Bußmann, Hadumod (Hrsg.) (2002): *Lexikon der Sprachwissenschaft*. 3., akt. und erw. Auflage. Stuttgart : Alfred Körner Verlag.
- Fillmore, Charles J. (2007): Valency issues in FrameNet. In: Herbst, Thomas & Katrin Götz-Votteler (Hrsg.): *Valency: Theoretical, Descriptive and Cognitive Issues*. Berlin / New York : Mouton de Gruyter. S. 129–160. (= Trends in Linguistics. Studies and Monographs; 187)
- FrameNet II (Homepage): <http://framenet.icsi.berkeley.edu/>
- Kunze, Claudia & Lothar Lemnitzer (2007): *Einführung in die Computerlexikographie*. Tübingen : Narr.
- Ruppenhofer, Josef, Michael Ellsworth, Miriam R. L. Petruck, Christopher R. Johnson & Jan Scheczyk (2006): *FrameNet II: Extended Theory and Practice*. Link: <http://framenet.icsi.berkeley.edu/book/book.pdf> (Stand: 30.04.2008, 18:05)
- University of California - Berkeley, Charles J. Fillmore (Bild). Link: http://linguistics.berkeley.edu/people/person_detail.php?person=14 (Stand: 30.04.2008, 18:14)