

Annotating semantic roles: phenomena

Katrin Erk
Sebastian Pado
ESLLI 2006

Structure

1. History of Semantic Roles
2. Contemporary Frameworks
3. Difficult Phenomena (from an empirical perspective)
4. Role Semantics vs. Formal Semantics
5. Cross-lingual aspects

1

Annotating semantic roles

[President Kennedy]_{Speaker} said_{STATEMENT}
[to an astronaut]_{Addressee}
[“ Man is still the most extraordinary
computer of all.”]_{Message}

[She]_{Theme} guided_{COTHEME} [Kim]_{Cotheme}
[along the street]_{Path} and began talking.

(English examples from FrameNet annotated
data, German examples from TIGER corpus
unless stated otherwise)

2

Annotating semantic roles

- The meaning of the predicate is relevant for determining the semantic roles available
- I heard him ask [her]_{Donor} to pass_{GIVING} [the salt]_{Theme}
- [The children]_{Theme} passed_{TRAVERSING} [the neighbor's yard]_{Area}

3

Annotating semantic roles: idioms

- [Kerry]_{Evader} gave [them]_{Pursuer} the slip_{EVADING} and was found Sunday night in Kota Bahru.
- “Give the slip” as a whole is a predicate. “the slip” does not fill any semantic role.

4

Annotating semantic roles: support

- Noun (or PP) as semantic head of a clause rather than the verb.
- Noun = predicate, verb = support:
 - Syntactic dependents of the verb are annotated as semantic roles of the noun.
- Examples:
 - [Someone]_{Speaker} [made] a statement_{STATEMENT} [about my need to correct my spelling]_{Topic}
 - [Frances Patterson]_{Patient} [underwent] an operation_{TREATMENT}

5

Annotating semantic roles: metaphors

- Example: “boil” in literal, non-literal use
 - [Water]_{Entity} boils_{ABSORB_HEAT} [at 100°C]_{Temperature} at one atmosphere pressure.
 - Just now work seemed the best antidote to the frustration [that]_{Emotion} was boiling_{EMOTION_HEAT} [inside her]_{Seat_of_emotion}.
- Design choice for metaphors: annotate literal meaning vs. understood meaning. Results in different semantic roles.

6

Annotating semantic roles: metaphors

- Roles in literal and nonliteral readings:
 - [She]_{Agent} threw_{CAUSE_MOTION} [her pencil]_{Theme} [across the room]_{Path}
 - [Danny's corner]_{Agent} refused to throw_{CAUSE_MOTION} [in]_{Goal} [the towel]_{Theme}.
 - [Danny's corner]_{Capitulator} refused to throw_{CAUSE_MOTION} [in the towel]_{GIVING_IN}.
- Roles of the literal reading may form part of the target in the nonliteral reading.

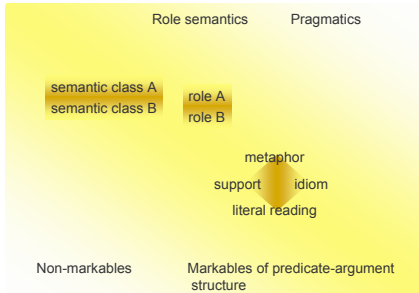
7

A problem of boundaries

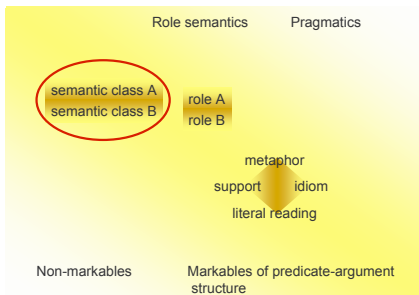
- Next:
 - problems of defining hard boundaries between phenomena in manual Frame-semantic annotation
 - illustrated on examples

8

Boundary problems at a glance



Unclear boundaries between semantic classes of predicates



Unclear boundaries between semantic classes of predicates

- to surprise/überraschen:
 - 👉 to take unawares
 - 👉 to cause emotion of surprise
- But:
 - Die Wurzeln mancher unguter Erscheinungen, die in unseren Gesellschaften auftreten und uns immer wieder überraschen, liegen in dem eben beschriebenen Zustand eines frustrierten postkommunistischen Gemüts. (The roots of some unfortunate events that occur in our societies and surprise us again and again lie in the state of a frustrated post-communist mind that we've just described.)
 - Unexpectedness or emotion? Or some of both?

Unclear boundaries between semantic classes of predicates

- to remark, notice/feststellen, bemerken:
 - ☞ to become aware
 - ☞ to make a statement
- But:
 - Zudem gibt es, so stellte die Prüfungskommission "Gläserne LPG" jüngst fest, viele Unregelmäßigkeiten. (In addition, as the examination board "LPG of glass" recently remarked/noticed, there are many irregularities.)
 - Becoming aware, or statement? Or some of both?

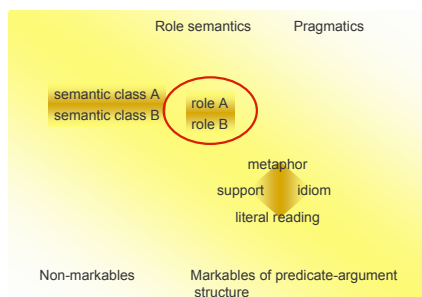
12

Unclear boundaries between semantic classes of predicates

- Slam the door/Tür zuschlagen: contains aspects of both Closure and Impact

13

Unclear boundaries between role labels



14

Unclear boundaries between role labels: granularity

- Frame Hostile_encounter:
 - Issue: something over which the two sides in a hostile encounter are in disagreement
 - Goal: the desired result of the outcome of the hostile encounter
 - struggle for control: goal.
struggle over familiar territory: issue
- **Machtkampf** / **power struggle**: Issue or Goal?

15

Unclear boundaries between role labels: granularity

- Frame Assistance: help, assist, aid, ...
 - **Goal**: the desirable state of affairs that the Benefitted_party is involved in and which is enabled by the Helper.
Jack helped Jill [in the development of the game].
 - **Focal_entity**: This role identifies a Focal_entity involved in achieving the Goal.
Whoever didn't cook has to **help** [with the dishes].
 - **But**: Can you **help** me [do the dishes]?
Neither Goal nor Focal_entity, rather some kind of Activity.
- Granularity problem remains, even with fine-grained FrameNet roles.
 - Goal, Focal_entity and Activity differ syntactically and in the inferences they afford
 - But they all relate to the activity in which the Helper helps.¹⁶

16

Unclear boundaries between role labels: metonymy

- Frame Statement:
 - Speaker: the person who produces the message
 - Medium: the physical entity or channel used by the Speaker to transmit the statement
- Problem with Metonymy:
Die nachhaltigste Korrektur **forderte** [ein Antrag]...
(The most extensive corrections are being demanded by [a motion]...)
- 2 strategies:
 - Ignore metonymy in general (because otherwise it interferes with very many role assignments). Then this should probably be Speaker.
 - Treat metonymy separately in all cases, and assign Medium here

17

Unclear boundaries between role labels: metonymy

- Frame Destroying:
 - Destroyer: conscious entity
 - Cause: an event, or an entity involved in such an event
- Inmitten des aufgeklärten Europas wurde [von Deutschland] die Zivilisation radikal zerstört.
(In the middle of enlightened Europe, civilization was radically destroyed [by Germany].)
- Again, problem of metonymy.

18

Unclear boundaries between role labels: core/noncore

- FrameNet distinguishes **core**, **peripheral**, **extrathematic** roles.
 - Core: conceptually necessary component of a frame; also: distinguishes this frame from others
 - Revenge frame: Avenger, Punishment, Offender, Injury, Injured_party
 - Peripheral: can be instantiated in any appropriate event frame
 - Time, Place, Manner, Means, Degree
- Similar: classical distinction obligatory/optional

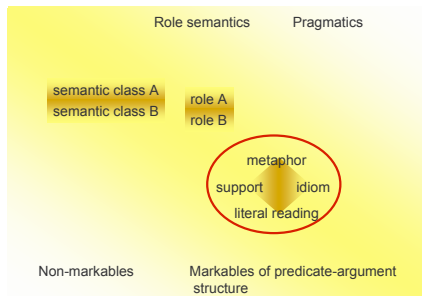
19

Unclear boundaries between role labels: core/noncore

- J. P. Koenig (2003): semantic arguments characterized by
 - obligatoriness (or at least frequency)
 - specificity
- Problematic: PPs
 - live in Berlin
 - stumble over the table ("quasi-valency" in PTB, Lopatkova & Panevova 2005)
 - win by two goals ("quasi-valency" in PTB)
- FrameNet: Addressee not core in Statement

20

Unclear boundaries between limited compositionality phenomena



Unclear boundaries between limited compositionality phenomena

- Metaphor versus idiom, metaphor versus separate but nonmetaphoric reading of the lexeme: How strongly is the literal reading still perceived?
- “Kick the bucket”: The bucket was a bar used by butchers to tie dead pigs to – by their back legs. So these pigs “kicked the bucket”.

22

Unclear boundaries between limited compositionality phenomena

- Metaphor versus separate but nonmetaphoric sense:
- Der “Pluralismus von Erklärungen” aus der CDU/CSU-FDP-Koalition zeige, dass die Einigkeit über die Pflegeversicherung nur “vorgetäuscht” worden sei, “um über die Sommerpause zu kommen”, sagte Klose.
(The “multiplicity of explanations” given by the CDU/CSU-FDP coalition showed that they only “pretended” to agree on nursing care insurance “in order to get through the summer break”, Klose said.)
- Get through (a difficult time):
 - metaphor with Motion source?
 - Or lexicalized and separate sense of “get”?

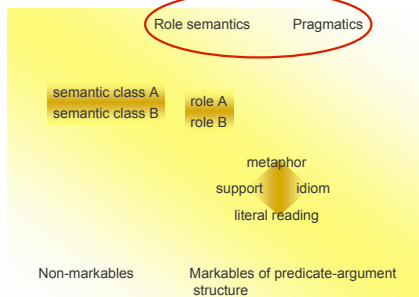
23

Unclear boundaries between limited compositionality phenomena

- Metaphor vs. support vs. separate word sense:
- Zwar liege die Verantwortung allein bei der Bundesregierung... (While responsibility lies solely with the federal government...)
- Support, similar to undergo/perform an operation
- Metaphor, with a Location source
- Separate word sense of lie: being located, even for abstract objects

24

Unclear boundaries between (role) semantics and pragmatics



25

Unclear boundaries between (role) semantics and pragmatics

- “What is X doing Y” (Kay & Fillmore 97)
 - What is this scratch doing on the table?
 - I wonder what the salesman will say this house is doing without a kitchen.
 - What are your children doing playing in my garden?
- Is this a special construction, as Kay&Fillmore argue, or a pragmatic phenomenon?

26

What is X doing Y: semantics or pragmatics?

- Possible analyses of "What's X doing Y":
 - Special construction with its own reading (markable in role semantics)
 - Pragmatics: additional reading derived by conversational implicature
- "Waiter, what is this fly doing in my soup?" – "Madam, I believe that's the backstroke"
 - "How come there is a fly in my soup?"
 - "What is the fly in my soup doing?"

27

What is X doing Y: semantics or pragmatics?

- Arguments in favor of the "WXDY" construction:
 - Context does not change the meaning of the construction:
 - Look what the children are doing in my garden! [Isn't that cute?]
 - But: What are your children doing playing in my garden? [Isn't that cute?]
 - Answer to the literal question often included in the expression:
 - What's your dog doing peeing on my doorstep?
 - Uses progressive even to express events or states that usually do not go with the progressive:
 - What is he doing knowing the answer?

28

Unclear boundaries between (role) semantics and pragmatics

- [We]_{Agent} immediately rushed to the ladies, washed Jessica carefully in the sink and dried_{CAUSE_TO_BE_DRY} [her]_{Dr/tee} [under the hand dryer]_{Place/Instrument}
- under the hand dryer:
 - Place
 - Instrument can be inferred because hand dryers usually blow hot air downward?
- Instrument is defeasible, can be overwritten by continuing the sentence by "... using lots of paper towels"

29

Unclear boundaries between (role) semantics and pragmatics

- [Drei Pötte per annum]_{Theme} verließen_{DEPARTING} während der achtziger Jahre [die Trockendocks]_{Source} ([Three ships a year]_{Theme} left_{DEPARTING} [the dry docks]_{Source} in the eighties.)
- Leaving dry docks = being manufactured?
- [Three ships a year]_{Product} left_{MANUFACTURING} [the dry docks]_{Factory} in the eighties.

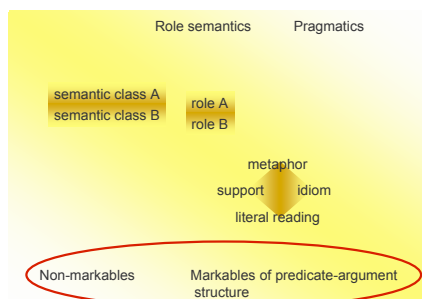
30

Unclear boundaries between (role) semantics and pragmatics

- Frame Statement, role Addressee: the person to whom the Message is communicated
- "vor dem Parlament" (before the parliament), "vor Journalisten" (in front of journalists/in a press conference), "vor den Delegierten" (in front of the delegates):
 - Place?
 - Inferred Addressee?
- Addressee-reading is defeasible, can be changed by context: Er sagte vor Journalisten zu seinem Ministerkollegen... (In a press conference, he said to his co-minister...)

31

Unclear boundaries between markables and non-markables



32

What is a predicate?

- Seen so far:
 - single lemma as frame-evoking: verb, noun, adjective
 - multiword expressions (even, maybe, something like “What’s X doing Y”)
- Also: parts of a word (German compounds)
Machkampf: struggle for power

33

What is a predicate?

“Colon construction”:

- “Das ist eines der weltbesten Teams”, zeigte Klinsmann Respekt.
(“This is one of the world’s best teams”, Klinsmann showed respect.)
- In English much more restricted than in German. Although:
 - And don’t expect many complete games by pitchers – perhaps three out of 288, laughs Mr. Fingers, the former Oakland reliever. (wsj_0214)

34

What are potential semantic role fillers?

- Local assignment:
 - Constituents within the maximal projection of the predicate
 - Arguments of the supporting verb in support constructions
 - Subjects of control verbs
- Nonlocal assignment of semantic roles:
 - Communication: message running longer than one sentence.
 - Rare: possible nonlocal roles for verbs
 - Besitzer von [Zweifamilienhäusern]_{Goods?} [die]_{Buyer} vor 1987 gebaut oder gekauft_{COMMERCE_BUY} haben (TIGER s975)
(Owners of [two-family homes]_{Goods?} [who]_{Buyer} have built or purchased_{COMMERCE_BUY} before 1987...)
Note defeasibility!

35

What are potential semantic role fillers?

- Nonlocal assignment of semantic roles:
 - Not so rare: possible nonlocal roles for frame-evoking nouns
 - Vor Jahren, als [Helmut Kohl]_{Addressee??} erstmals ganz unten war[...], machte [...] [Günter Oettinger]_{Speaker?} bundesweit mit einer [Rücktritts]_{Message} forderung_{Request} von sich reden. (TIGER s1862)
 (Years ago, when [Helmut Kohl]_{Addressee??} was on the rocks for the first time, [Günter Oettinger]_{Speaker?} brought himself into public awareness with a demand_{REQUEST} [for resignation]_{Message}.)

36

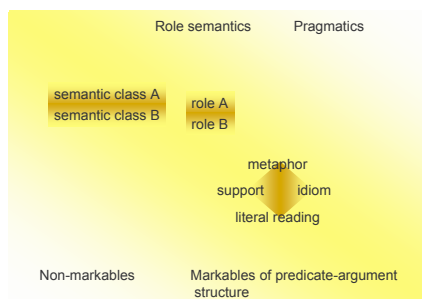
What are potential semantic role fillers?

- Comparatives: content of a semantic role not overtly realized
 - The flooding was worse than expected_{EXPECTATION}.

What is the Phenomenon here?

37

Unclear boundaries: summing up



38

Unclear boundaries: What do we make of this?

- Focus on
 - unclear boundaries between predicate senses (frames)
 - unclear boundaries between semantic roles
- Unclear boundaries in sense and role assignment:
 - Just a result of not defining classes clearly enough?
 - Inherent feature of the phenomena?

39

Unclear boundaries: bug or feature?

- Word Sense Disambiguation:
 - School A: just a problem of bad definitions.

More coarse-grained definitions will do the trick.
e.g. Hovy et al (2006)
 - School B: vagueness inherent to the phenomenon.

Remains persistent with all sense inventories.

40

Unclear boundaries: bug or feature?

- Semantic role assignment:
 - Granularity problem with small set of abstract roles (see earlier class),
 - Granularity problem also with FrameNet roles: the case of Assistance above
- Following "School B", we'll pursue the idea that vagueness may be inherent in the phenomenon
- Next: prototype theory

41

Concepts: The classical theory

- Goes back to Aristotle
- A concept is defined by a set of **necessary and sufficient conditions**
 - **Acquiring concepts:** Conditions are empirically discoverable. We consider instances of a concept and extract distinguishing conditions.
 - **Classifying items:** by checking whether they fulfil conditions.
- **Examples:**
 - BACHELOR: unmarried male human
 - HUMAN: featherless biped

42

Problems with the “necessary and sufficient conditions” theory of concepts

- Do all concepts have a definition?
 - Wittgenstein (1953): GAME as a concept without a definition
 - Same problem with BACHELOR: children? divorced or widowed men? priests?
- Psychological corroboration?
 - People rarely able to provide definitions
 - Definitions differ between people, differ over time
 - Reaction time should depend on complexity of definition
Reaction to MALE should be faster than to BACHELOR. ⁴³
This is not the case.

43

Prototype theory

- Eleanor Rosch and others, 1970s
- Typicality effects in members of a common category
- Experiments:
 - Rating in how far items are good examples of the category FURNITURE, scale 1-7: “chair” and “sofa” on top, “telephone” and “refrigerator” at the bottom (Rosch 1975)
 - Response time: “Is a robin a bird?”
 - Naming examples

44

Prototype theory

- Concepts organized by family resemblances
- Fuzzy boundaries: some borderline items may "more or less" belong to a category
- Concepts represented by a set of weighted features
 - BIRD: has features, can fly, ...
 - Need not be present in all members of a category
 - Features are weighted: Theories name either some notion of salience, or the relative frequency of category members exhibiting a feature

45

Prototypes in lexical semantics

- Linda Coleman and Paul Kay (1981): the verb "lie"
 - Participants were asked to rate situations on whether they represent a lie
 - Participants did not seem to use necessary and sufficient conditions
 - Participants seemed to use cluster of features, none of them necessary, varied in importance:
 - untruth (least important)
 - speaker knows that the statement is not true (most important)
 - intention to deceive (medium important)
 - Participants easily use "more-or-less" ratings, good agreement
 - Prototype for rather abstract concept

46

Prototypes in lexical semantics

- Patrick Hanks (2000): Do word meanings exist?
 - "bank": "financial institution" vs "riverside" reading
 - 1000 random occurrences of the noun "bank" in the BNC:
 - No ambiguity between "finance" and "river" reading
 - But cases where a reading is only partially activated
 - "Financial" reading:
Blood bank, data bank, seed bank?
 - "Riverside" reading:
Sloping land? Water? One slope or two?
- Again: features, none of them necessary, may be "more-or-less" activated. Usually a combination of features is activated.

47

Prototypes and FrameNet

- Fillmore (1975): Alternatives to checklist theories of meaning
Proposes prototypes instead of necessary and sufficient conditions
- FrameNet: a prototype-based theory of meaning

48

Sense and role assignment: Consequences of adopting prototype theory?

- Senses (frames):
 - not assumed disjoint
 - assignment of more than one frame label to signal similar distance to several classes
- Roles:
 - Each role bears properties/inferences, in virtue of being a certain participant in the given frame
 - Properties as prototype features: a given instance may fulfil more or less of the properties of a given role

49

Sense and role assignment: Consequences of adopting prototype theory?

- Frame/role described by features which need not be present in all cases
 - Defeasible features (as in "birds can fly")
 - More-or-less features (as in color terms: more or less green)
- Consequences for doing reasoning over semantic representations: Handle non-certain knowledge?

50

Summary

- Annotation of normal cases, also involving metaphor, idioms, support
- Boundary problems concerning:
 - semantic classes (senses)
 - semantic roles
 - different kinds of Limited Compositionality phenomena
 - markables vs. non-markables
 - semantics vs. pragmatics
- Attempting to explain boundary problems concerning senses and roles:
Prototype theory

51

References

- Collin Baker, Charles Fillmore, John Lowe (1998): The Berkeley Framenet project. In: Proceedings of the COLING-ACL, Montreal.
- Aljoscha Burchardt, Katrin Erk, Anette Frank, Andrea Kowalski, Sebastian Pado and Manfred Pinkal (2006): The SALSA corpus: a German corpus resource for lexical semantics. In: Proceedings of LREC 2006, Genova.
- Linda Coleman and Paul Kay (1981): Prototype semantics: the English verb "lie". *Language* 57/1:26-44
- Michael Ellsworth, Katrin Erk, Paul Kingsbury, and Sebastian Pado (2004): PropBank, SALSA, and FrameNet: how design determines product. In: Proceedings of the LREC 2004 workshop on building lexical resources from semantically annotated corpora, Lisbon.
- Katrin Erk, Andrea Kowalski, Sebastian Pado and Manfred Pinkal (2003): Towards a resource for lexical semantics: a large German corpus with extensive semantic annotation. In: Proceedings of ACL 2003, Sapporo.
- Patrick Hanks (1994): Linguistic norms and pragmatic exploitations, or Why lexicographers need prototype theory, and vice versa. In: F. Kiefer, G. Kiss and J. Pajzs (eds.) *Papers in Computational Lexicography: Complex '94*. Hungarian Academy of Sciences, Budapest.

52

References

- Patrick Hanks (2000): Do word meanings exist? *Computers and the Humanities* 34
- Ed Hovy, Mitchell Marcus, Martha Palmer, Lance Ramshaw and Ralph Weischedel (2006): OntoNotes - the 90% solution. In *Proceedings of HLT-NAACL 2006*, New York.
- Jean-Pierre Koenig, Gail Mauner, and Breton Bienvenue (2003): Arguments for Adjuncts. *Cognition*, 89:67-103
- Marketa Lopatkova and Jmila Panevova (2005): Recent developments in the theory of valency in the light of the Prague Dependency Treebank. In *Insights into Slovak and Czech Corpus Linguistics*, 83-92. Veda Bratislava, Slovakia
- Eleanor Rosch (1975): Cognitive representation of semantic categories. *J Experimental Psychology* 104:192-233
- Josef Ruppenhofer, Michael Ellsworth, Miriam Petruck, and Christopher Johnson (2005): FrameNet: Theory and Practice. <http://framenet.icsi.berkeley.edu>
- Ludwig Wittgenstein, *Philosophical investigations*. Prentice Hall, 1999. (Originally appeared 1953)

53
