

Sentences expressing relations between geographic entitites

*Enriching an ontology of German place names
using local grammars*

Sebastian Nagel

wastl@cis.uni-muenchen.de

Lexis and Grammar Conference

Palermo, 9 sept. 2005

Relations between geographic entities

- a geographic ontology: places (in some spatial representation, e.g. a pair of coordinates), their names, and the relations among them
- relations needed for problem solving
- also in language, e.g.:
Paris rejected the “logic of ultimatums.”
- types of relations
 - hierarchical: containment, administrative subdivision
 - more specific: ‘X is capital of Y’, ‘X flows into Y’, etc.
 - spatial: *Tarrytown is located north of New York City*

Linguistic realization (in German)

- description is started from one meaning, e.g. ‘the fact that one river X flows into a body of water Y’ (= a relation between X and Y)
- various part-of-speeches expressing this fact:
 - verbs
 - X (mündet+fließt+strömt+ergießt sich+entwässert+...) in Y*
 - X (flows+feeds+empties+discharges+drains+...) into Y*
 - nouns
 - X ist ein (Nebenfluss+Zufluss+Quellfluss+...) von Y*
 - X is a (tributary+affluent+confluent+influent+feeder+...) of Y*

Linguistic realization (in German)

● for other relations also:

● prepositions

München liegt in Bayern (containment)

Munich is located/situated in Bavaria

● adjectives/adverbs

Windach liegt $\frac{1}{2}$ Autostunde von München entfernt, ...
(distance)

Windach is $\frac{1}{2}$ hour away by car from Munich

Linguistic realization (in German)

- by various types of phrases:
 - sentence
 - noun phrase
 - die (Mündung+Zusammenfluss+...) des X in Y*
 - the (mouth+confluence+...) of X into Y*
 - compound noun
 - die Isarmündung in die Donau*
 - mouth of Isar into Danube*
 - Donauzufluss*
 - tributary of danube*

Relations, predicates and valency

- most relations between geographic entities are binary relations
- binary relations should be expressed by (at least) bivalent predicates
- The notion of quasi-predicates: a quasi-predicate is an entity having argument slots (Mel'čuk 2004a: 10, cf. Gross 1998: 223) For instance:
 - train(Paris; Munich) = *train [going] from Paris to Munich*
 - father(John; Mary) = *John is Mary's father*
- most (if not all) nouns expressing relations among geographic entities are quasi-predicates

Relations, predicates and valency

the case of *Mündung* vs. *Zufluss*:

- transformation verb → noun (*zufließen* → *Zufluss*, *münden* → *Mündung*)
 - *Die Tauber fließt dem Main zu.* (rare and old fashioned)
Der bedeutendste Zufluss des Mains auf dem Gebiet Baden-Württembergs ist die Tauber.
 - *In Wertheim mündet die Tauber in den Main.*
Wertheim liegt an der Mündung der Tauber in den Main.
- semantics: *Mündung* ('river mouth') is neither 'the fact that a river X flows into another body of water Y'
- nor 'the river X which flows into another body of water Y' (which is *Zufluss* = 'affluent', i.e. a bivalent noun)
- it is 'a place where a river Y flows into another body of water Z'

Relations, predicates and valency

- is the place itself an actant?

- rarely observed sentences:

Die Boca ist die Mündung des Riachuelo in den Rio de la Plata.

Der Meeresarm von Burry ist die Mündung des Loughor-Flusses.

cf. also *Die Camargue ist das Mündungsgebiet der Rhône ins Mittelmeer.*

- if *Mündung* is bivalent there is no good elementary sentence, but only:

This/here is the mouth of ...

- evidence from syntax: shift in realization of actants:
'X flows into Y'

X ist ein Zufluss des Y ↔ die Mündung des X in den Y

compound nouns: *Y#zufluss ↔ X#mündung*

- it depends on the ontology (view of the world): do entities of a class 'river mouth' exist?

Trying to write a grammar

- phrase grammars for noun phrases referring to geographic entities (using large dictionaries of toponymes and class names)
- local grammars for typical inserts and modifiers
- lexicon grammars for predicates
- problems
 - inserts (never ending)
 - coordination
 - specific German: highly inflective, free word order, sentence bracket resulting in non-local constraints
- grammar size: e.g. 160 lexicon grammar entries of nouns expressing binary relations, in terms of an Unitex grammar (=RTN): 1000 subgraphs, 13.000 states, 28.000 transitions, 3.500 symbols

Applications: NER

- lexicon grammars (= predicate-argument-structures) as one component of an NER-system for recognition of German place names
- used to describe external context

Blutbad islamischen Fundamentalisten angelastet
Fünf Franzosen in **Algier** erschossen
Paris fordert seine in **Algerien** lebenden Staatsbürger zur Rückkehr nach **Frankreich** auf
Paris (Reuter/AFP/AP) - In **Algerien** haben mutmaßliche muslimische Fundamentalisten am
Mittwoch erneut ein Blutbad unter Ausländern angerichtet. Bei dem Feuerüberfall in der
Hauptstadt **Algier** wurden fünf französische Staatsbürger getötet und einer verletzt, wie das
Außenministerium in **Paris** mitteilte. Bei den Getöteten handle es sich um drei Polizeibeamte und
zwei Botschaftsangehörige.

Applications: to be done

- ontology: extracting a database of relations from corpora (Senellart 1998a: cf.)
- sentences resulting in same database entries are (partial) paraphrases:

Bei Passau mündet der Inn dann in die Donau.

= Der Inn ist ein rechter Nebenfluss der Donau und 510 km lang.

= Der Inn ist der drittgrößte und der siebt längste Donaunebenfluss.

= ...

Thank you!

Slides can be found on

<http://www.cis.uni-muenchen.de/~wastl/pub/LGpalermoSlide.pdf>

Literature:

- Gross, Gaston 1998: Pour une typologie des prédicats nominaux. In: Forsgren; Jonasson; Kronning (eds.): *Prédication, assertion, information* 221–230.
- Mel'čuk, Igor 2004a: Actants in semantics and syntax I: Actants in semantics. *Linguistics* 42/1, 1–66.
- — 2004b: Actants in semantics and syntax II: Actants in syntax. *Linguistics* 42/2, 247– 291.
- Senellart, Jean 1998: Locating noun phrases with finite state transducers. *Proceedings of the 17th International Conference on Computational Linguistics (COLING-98)* 1212–1219. Montréal.
<http://citeseer.nj.nec.com/senellart98locating.html>