1 Panini’s karākas

The notion of semantic roles was already present in Paninian grammar. Four levels of different syntactic and semantic representations were provided by Panini according to Kiparsky (2002). Semantic roles and abstract tense together constituted the top-level of a four-level hierarchy presented in Figure 1 below. This semantic level had a direct connection with the respective syntactic representation of arguments to the verb in a given sentence. The semantic roles were given a morphosyntactic representation, which was mapped onto an abstract morphological representation, and the abstract morphological representation was finally given a phonetic representation.

| semantic information (karākas and abstract tense) |
| morphosyntactic representation |
| abstract morphological representation |
| phonological output form |

Figure 1: Four-level hierarchy

Kiparsky (2002, pp.3-4) lists these karākas and describes them as follows:

- **AGENT** (*karty*) - the participant which is independent
- **SOURCE** (*apādāna*) - the participant which is fixed
- **GOAL/PATIENT** (*karman*) - the participant which is the primary target of the action
- **LOCATION** (*adhikaraṇa*)

Cardona (1974), in addition lists:

- **BENEFICIARY** (*sampradāna*) - one whom (an agent) intends as goal (of his action)\(^1\)
- **MEANS** (*karaṇa*) - The means par excellence (of accomplishing an action)

(Kiparsky, 2002, p.3ff) exemplifies the semantic roles with the sentence below:

---

\(^1\)Not explicitly said by Cardona, *karman* and *sampradāna* seem to differ with respect to animateness. *karman* is the animate.
and describes thereafter which semantic role is represented by the words in the sentence:

- vánā ‘forest’: source (apádāna)
- gráma ‘village’: goal (karman)
- adyā ‘today’: (temporal) location (adhikaraṇa)
- áśvapatá ‘descendant of Áśvapati: agent (kárty)
- odanā ‘rice’: goal (karman)

These karākas together are the ones that we recognise from listings of basic semantic roles today as well².

2 Semantic roles in language technology

Semantic roles are used to classify the arguments to verbs. In Scandinavian languages it is, for example, not always evident/obvious which is the subject and which is the object. Other semantic features such as animateness can be used, for example, to identify the agent and help us predict that this is the subject of the sentence.

In language technology it is important to know which one is the subject, and which case the other arguments have, e.g. to make appropriate translations. Subject and object can be represented in different ways in different languages, demand a certain case or be put in a certain order to make it correct. This can, for example, be important in a translation system since the mapping between semantic roles and grammatical functions varies from language to language.

There are some machine readable dictionaries and thesauri that include semantic roles. Among them probably FrameNet (FrameNetStaff, 1997-2006) is the most well-known and widely used. FrameNet is a lexical resource mainly for verbs, but also nouns and adjectives. It is build on the research of Frame semantics by Charles Fillmore. Frame semantics is built on semantic roles. FrameNet is used in several NLP applications such as word sense disambiguation, machine translation, information extraction and question answering systems.

References


²Unfortunately, Cardona doesn’t provide examples in this easy-read manner as Kiparsky does. Hence I can’t give examples for beneficiary (sampradāṇa) and means (karaṇa).

2