

Theory-specific sub-ontologies and theory-neutral general ontologies

Graham Wilcock

Department of General Linguistics
University of Helsinki
graham.wilcock@helsinki.fi

This poster describes work on combining a theory-specific sub-ontology with a more general theory-neutral ontology. The domain is linguistic description.

Background

Academic linguists tend to hold strong views on linguistic theories. Each major theory has its own dedicated annual conference, where its advocates present theory-specific linguistic descriptions, and show how their theory explains the data better than rival theories. Field linguists by contrast tend to be busy recording examples of endangered languages before they disappear forever, and less interested in which theory may offer a more elegant analysis. They want a general framework which helps them to capture data about an unknown language with a minimum of theoretical preconceptions.

HPSG: A specific linguistic theory

Head-driven Phrase Structure Grammar (HPSG) (Pollard & Sag 1994) is one specific linguistic theory. Unlike most other theories, it is based on a rigorous foundation of typed feature structures with a clear type hierarchy. HPSG is usually implemented by unification-based typed feature structure systems such as LKB (Copestake 2002).

HPSG OWL: A theory-specific ontology

Using Protégé we have developed a theory-specific HPSG OWL ontology. Linguistic theories, like ontologies, evolve through different versions: we adopted the version of HPSG described by Sag et al. (2003), with an emphasis on multiple inheritance and cross-classification. This work is ongoing. The current focus is on using reasoners to support cross-classification and SWRL to support lexical rules and grammar rules.

GOLD: A theory-neutral general ontology

General Ontology for Linguistic Description (GOLD) (Farrar & Langendoen 2003) (<http://www.linguistics-ontology.org/>) is an OWL ontology that offers a more general framework intended for field linguists. GOLD uses theory-neutral class names, such as GrammaticalUnit, SemanticUnit, FeatureSpecification. Although GOLD initially attempted to be entirely theory-neutral (as regards competing linguistic theories), it has moved towards supporting different "communities of practice" within the overall framework (Farrar & Lewis 2005).

COPE: Community of practice extensions

A community of practice in GOLD may focus on developing a consensus within a specific area, for example in phonology or in describing Bantu languages. In such a case, the aim is to encourage the development of "best-practice" resources. On the other hand, communities of practice may focus on competing theories, where each sub-community has its own distinctive terminology and divergent conceptualization. In this case, the aim is different: to capture explicitly the relationship between the sub-community view and the overall framework, in the form of a Community of Practice Extension (COPE) (Farrar & Lewis 2005). A COPE is a sub-ontology that inherits from, and extends, appropriate parts of the overall GOLD ontology.

HPSG COPE: A theory-specific sub-ontology inside GOLD

HPSG OWL has an HPSG-specific class hierarchy using HPSG terminology. In order to combine this with GOLD, we define the HPSG-specific classes as subclasses of the theory-neutral GOLD classes. For example, in the HPSG type hierarchy there is the type "sign" (Sag et al. 2003, p. 475) with these attributes:

PHON	type: list(form)	(a sequence of word forms)
SYN	type: gram-cat	(a grammatical category)
SEM	type: sem-struc	(a semantic structure)

and in GOLD there is the class `LinguisticSign` which includes the properties:

hasForm	Range: <code>PhonologicalUnit</code>
hasGrammar	Range: <code>GrammaticalUnit</code>
hasMeaning	Range: <code>SemanticUnit</code> .

Using namespaces, we define HPSG types as subclasses of GOLD classes, and define HPSG attributes as subproperties of GOLD properties. For this example:

hpsg:sign	subclass of	gold:LinguisticSign
hpsg:form	subclass of	gold:PhonologicalUnit
hpsg:gram-cat	subclass of	gold:GrammaticalUnit
hpsg:sem-struc	subclass of	gold:SemanticUnit
hpsg:PHON	subproperty of	gold:hasForm
hpsg:SYN	subproperty of	gold:hasGrammar
hpsg:SEM	subproperty of	gold:hasMeaning

The work on both HPSG OWL and HPSG COPE is ongoing.

References

- Copestake, Ann (2002) *Implementing Typed Feature Structure Grammars*. CSLI Publications, Stanford.
- Farrar, Scott & Terence Langendoen (2003) A linguistic ontology for the Semantic Web. *GLOT International* 7(3), 97-100.
- Farrar, Scott & William Lewis (2005) The GOLD Community of Practice: An infrastructure for linguistic data on the web. URL: <http://www.u.arizona.edu/~farrar/>.
- Pollard, Carl & Ivan Sag (1994) *Head-driven Phrase Structure Grammar*. CSLI Publications, Stanford.
- Sag, Ivan, Thomas Wasow & Emily Bender (2003) *Syntactic Theory: A Formal Introduction* (2nd edition). CSLI Publications, Stanford.