An object-oriented computational model for processing linguistic constructions

Yoichiro Hasebe
Doshisha University, yhasebe@mail.doshisha.ac.jp

A computational model for storing and processing corpus data is introduced. It maps fine-grained linguistic data onto data objects and relates every object to other objects in a cognitively realistic fashion. By implementing this model, a computational network representing linguistic constructions can be built in a way that is highly consistent with the theories and principles of cognitive linguistics (Croft 2002; Goldberg 2006; Langacker 2008).

An advanced multi-layered data processing system (AMPS) takes linguistic data that are tagged in a multi-layered format and maps them onto objects of different classes, such as sentences, morphemes, frames, and conceptual primitives. These classes are in a relationship of “composition” in terms of object-oriented program design (Meyer 2000; Hasebe 2005), and provide an outer layer of the network. Each of these classes, in turn, has its internal complexity also; with superclasses in above and subclasses below, realizing a rather inner layer of the network. Placed amongst such an intertwined network structure, linguistic constructions (or any other smaller linguistic units, for that matter) can be defined not as a set of values assigned to a certain number of parameters, but as a profiled region of the (possibly enormous) whole, corresponding to the fundamental idea of cognitive linguistics.

Software/database design has not been accorded the importance it deserves in the fields of linguistics, although there are a few successful attempts to structure linguistic data in a way that is cognitively plausible (e.g. Fillmore et al. 2002; Bergen and Chang 2004). It is important to note, however, that the application of cognitive linguistics also to the way linguistic data are actually stored and processed is desirable for the purpose of both constructing useful corpus systems and providing a demonstrable proof of theory.

References


