

Quantifiers in Language: Use and Meaning

Language is the primary means of human communication, and the main vehicle for scientific as well as common sense reasoning. It has been investigated extensively from various theoretical perspectives, that of linguistics, logic, and computer science. At the heart of this multi-faceted scientific enterprise is a simple yet powerful conception of the meaning of linguistic expressions. This notion of meaning has been very fruitful resulting in a wealth of practical applications, e.g., in computer science, including dialogue systems, automated reasoning, information retrieval and search. The undeniable advantages of this theoretical endeavor invite a question: to what extent can it account for the human linguistic behavior? The past decade has seen the increasing interaction between cognitive science and linguistics. The new field of experimental semantics and pragmatics has been facing many challenges. One of the main difficulties is of practical nature: experimental semantics is in serious need of natural linguistic data. The recent advances in the field have focused on acquiring psycholinguistic data via lab experimentation. Such data are crucial to understand language processing, however, there posit a problem from the perspective of semantic interpretation. Namely, the subjects in such experiments do not need to follow their everyday linguistic behavior; they can develop various cognitive strategies, for instance trading of speed and linguistic accuracy. Therefore, to get a full picture of how people interpret language we need data coming from the real use of language. This project will build a linguistic corpus, i.e., a huge database of natural language sentences annotated with linguistic information that will offer a plethora of information for language theory. In particular we will focus on detecting natural language expressions talking about quantities, so called quantifiers, like 'some' or 'most'. The corpus and the toolkit for detecting and analyzing quantifiers will be freely available online. Using the corpus we will be able to better understand why people use some quantifier expressions more frequently than others.