

# Learning form-meaning mappings for language

Nancy Chang

University of California, Berkeley and International Computer Science Institute  
1947 Center Street, Suite 600  
Berkeley, CA 94704  
nchang@cs.berkeley.edu

The proposed thesis research addresses two of the main obstacles to building agents that communicate using natural language: the need for richer representations of linguistic constructions that incorporate aspects of conceptual knowledge, context and goals; and the need for a principled approach to the automatic acquisition of such structures from examples. More generally, it explores the idea that patterns that arise in language are inextricably linked with and motivated by patterns of meaning and experience. This view, along with empirical evidence suggesting that linguistic knowledge at all levels can be characterized as mappings between form and meaning, serves as the basis for a computational model of the acquisition of simple phrasal and clausal constructions.

The idea that language involves associations between sounds (form) and relatively richer sets of stimuli (meaning) is neither new nor surprising. Indeed, such an assumption is made without controversy in some recent models of the acquisition of individual words, including those for objects, spatial relations, and actions (Regier 1996; Bailey 1997; Siskind 1997; Roy & Pentland 1998). Thus far, however, models of the learning of larger phrasal and clausal structures have been oriented toward the problem of acquiring symbolic syntactic patterns, usually based around verbal argument structure (e.g., Brent 1994). The meaning of these larger structures is typically assumed to be predictable from the meaning of its constituents.

The current work adopts the less strictly compositional framework of Construction Grammar (Goldberg 1994). On this view, although phrasal and clausal constructions require more complex structural description of the relations between their constituents, they can still be described, like words, as mappings between form and meaning. Parameters of form consist primarily of phonological cues, such as features of intonation, words and their inflections, and word order. Parameters of meaning encompass a much larger set of possibilities, including event structure, sensorimotor control, force dynamics, attentional state, perspective-

switching ability, and communicative and social goals. These features, while complex, play an important role in both cognitively motivated linguistic theories and cross-linguistic evidence from the study of early child language (Slobin 1985).

The model described here assumes some initial set of constructions corresponding to already learned individual words or unanalyzed phonological units, where constructions are mappings or associations between hierarchically structured sets of features corresponding to phonological input (form) and a simple event (meaning). The learning problem is to hypothesize, based on the current inventory of constructions and new input, the best new set of constructions to fit the data and generalize to future input. We extend previous work using Bayesian model merging as the basis for generalization (Stolcke 1994; Bailey 1997) to include composition, subset selection, and extension as possible operations on the set of existing constructions. Since these operations are not sufficient for expressing constructions that include binding *relations* between entities, we propose additional operations that allow a constrained form of unification in which relations detectable from the form representation can be associated with relations over corresponding entities in the meaning representation.

Besides addressing an instance of the general problem of learning relations and correlations over previously learned associations, this approach to grammar learning is also intended as a cognitive model that can make tractable the modeling of empirical data, such as that in the CHILDES corpus of parent-child interactions (MacWhinney 1991). The use of relatively rich cognitive structures for meaning places most of the generalization power in the semantic domain, creating a tradeoff between the (mostly semantic) drive to generalize and the map-enforced need to adhere to specific form-based correlates. The resulting learning behavior accords well with the observation that children acquire constructions on a verb-specific basis before generalizing to more abstract constructions (Tomasello 1992), demonstrating how a richer approach to meaning representation can facilitate the acquisition of natural language for both human and artificial agents.

*Full references available upon request.*

## References

- BAILEY, DAVID R. 1997. *When Push Comes to Shove: A Computational Model of the Role of Motor Control in the Acquisition of Action Verbs*. Computer Science Division, EECS Department, University of California at Berkeley dissertation.
- REGIER, TERRY. 1996. *The Human Semantic Potential*. MIT Press.
- ROY, DEB, & ALEX PENTLAND. 1998. Learning audio-visually grounded words from natural input. In *Proc. AAAI workshop on The Grounding of Word Meaning: Data and Models*.
- SISKIND, JEFFREY MARK. 1997. *A computational study of cross-situational techniques for learning word-to-meaning mappings*, chapter 2. MIT Press.
- SLOBIN, DAN ISAAC. 1985. Crosslinguistic evidence for the language-making capacity. In *Theoretical Issues*, ed. by Dan Isaac Slobin, volume 2 of *The Crosslinguistic Study of Language Acquisition*, chapter 15. Mahwah, New Jersey; London: Lawrence Erlbaum Associates.
- STOLCKE, ANDREAS. 1994. *Bayesian Learning of Probabilistic Language Models*. University of California, Berkeley dissertation.
- TOMASELLO, MICHAEL. 1992. *First verbs: A case study of early grammatical development*. Cambridge, UK: Cambridge University Press.