People use words to refer to the world as a means for influencing the beliefs and actions of others. Although many isolated aspects of the structure and use of language have been studied extensively and modeled, a unified model of situated language use remains unexplored. Any attempt to explain unconstrained adult language use appears futile due to the overwhelming complexity of the physical, cognitive, and cultural factors at play. A strategy for making progress towards a holistic account of language use is to study simple forms of language (e.g., conversational speech about objects and events in the here-and-now in limited social contexts) and strive for "vertically integrated" explanatory models. I will present experiments in building conversational robots and modeling children's word learning motivated by this strategy. An emerging theoretical framework provides computational insights into the interaction of perceptual, planning, and memory processes underlying simple speech acts. The approach suggests a novel view of the development of emotive/connotative meaning as a result of embodied experience.