Modeling the acquisition of quantifier semantics: a case study in function word learnability Our work studies the acquisition of quantifier meanings as a case study of function word learnability. We suggest that learners construct semantic representations of quantifiers and other function words using a compositional statistical learning mechanism that operates over a small set of domain-general cognitive primitives. We present a simple cross-situational learning model that provably solves key learning problems in this domain, using a developmentally-plausible amount of data. Our model combines tools that have previously been proposed to explain concept learning across cognition, including efficient (Bayesian) statistical inference over compositions of primitive operations with explicit representations like those used in semantics. The operations assumed to be available to learners before they acquire quantifiers include simple logical and set-theoretic functions. The task of the learner is to take positive examples of quantifiers used in context and infer an explicit composition of primitives that explains the observed utterances. Thus, the model ends up learning a representation much like one might find in an elementary semantics textbook. The logic of this approach provides a powerful inductive framework that have been proposed previously to explain learning in domains like magnetism (Ullman, Goodman, & Tenenbaum, 2010), number words (Piantadosi, Tenenbaum, & Goodman, 2012), and nouns and verbs (Siskind, 1996). In the case of quantifiers we present a simple proof that this statistical setup always yields learnability for all quantifier meanings, even when only provided with positive evidence. This strong result contrasts with limited learnability proofs for quantifiers established previously (e.g. Clark 1996, Tiede 1999). More interestingly, the learnability proof extends to other function words in language, establishing broadly that latent (unobserved) semantic representations can always be learned from observing language use in context. Beyond establishing learnability in principle, we establish it in practice, using common and simple statistical inference techniques. Our work provides an

implementation that correctly acquires quantifier meanings from noisy positive data, while requiring relatively little data. The implementation can be used in principle to predict the time-course of quantifier acquisition and mistakes children make. We show how the implementation can be used to acquire both literal and presuppositional aspects of meaning simultaneously, using identical representational and statistical tools. We additionally use the learning model to evaluate the utility of proposed constraints on quantifier meaning like conservativity. We show that learning in an unrestricted space of meanings is not substantially more difficult than learning in highly-constrained frameworks, meaning that constraints on quantifiers should not be posited for reasons of learnability (e.g. Keenan and Stavi 1986). We show that in many cases, learning in a relatively unrestricted space of compositions of set-theoretic primitives is not substantially harder than determining the correct meanings when only the correct meanings are possible hypotheses for the learner.