A connectionist approach to modeling the acquisition of spoken words is described in which an internal level of representation is formed under the task demands of spoken word comprehension, production, and imitation. This internal level of representation is hypothesized to become sensitive to the phonological structure that exists in the spoken language environment. An integral part of this modeling approach is the hypothesis that the acoustic consequences of articulatory outputs are learned through babbling and early attempts at spoken utterances (i.e., a forward model of the mapping from articulation to acoustics). Early work on the development of a forward model is reported. The forward model is trained on a database of articulatory and acoustic measurements collected from English speakers. In a more complete implementation of our approach, the forward model will serve as a pathway by which feedback from the spoken language environment can shape internal representations of spoken words. [Supported in part by a pending grant from NIH (PI David Plaut).]