[S30] Subword linguistic modeling for multiple applications in speech understanding. Stephanie Seneff, Computer Science & Artificial Intelligence Lab, MIT, Cambridge, MA, USA.

Over the past decade, we have been conducting research on the use of subword modeling for many different applications in speech understanding. This line of research is based on our belief that a hierarchical representation that captures the subword structure, ranging from metrical feet and syllables to phonemes and distinctive features, can offer powerful constraints on the description of phonological events. The research is in part aimed at the difficult task of identifying and characterizing unknown words, although the framework also has utility in other speech recognition tasks such as phonological and prosodic modeling, as well as open vocabulary speech recognition. The approach exploits the linguistic substructure of words by describing graphemic, phonemic, phonological, syllabic, and morphemic constraints through a set of context-free rules, and supporting the resulting parse trees with a corpus-trained probability model. Statistical methods are combined with explicit linguistic information to generalize from the observation space to unseen words. A derived finite state transducer representation forms a natural means for integrating the trained model into the search framework of a speech recognizer.

There are several ways in which speech understanding tasks can benefit from such formal modeling of word substructure. These include statistical phonological modeling, hierarchical duration modeling, sound-to-letter and letter-to-sound mapping, and automatic acquisition of unknown words in a speech understanding system. Results from a selected set of experiments in applications involving spoken dialogue systems will be presented.