High-density event-related potential (ERP) and magnetoencephalography (MEG) data are presented here to illustrate the plurality and plasticity of speech representation in the spatio-temporal dynamics of neural responses. First, there is evidence for both continuous and discrete processing of speech sounds depending on the acoustic parameters and experimental conditions. Speech discrimination and neural sensitivity measures suggest that the perceptual organization of within-category variations is based on both phonetic knowledge and acoustic analysis. Second, attention plays an important role in determining the neural response patterns. Focused attention significantly changes the hemispheric laterality and strength of neural activation for speech processing. Selectively attending to the different aspects (phonetic vs. social-indexical information) of the same speech stimuli shows different regional recruitment, hemispheric involvement and temporal characteristics. Third, the nature of language input and experience can significantly affect learning outcome. Enriched exposure alters the neural representations for speech sounds and plasticity of the system. Theoretical and practical implications for language learning and treatment are discussed.