Face is very popular social stimulus for us, and perhaps most important stimulus for infants. In deed, it is repeatedly demonstrated that infants, and even neonates, prefer face/face-like images to other images. However, it is still not clear when infants began to stop looking at face just automatically, and to start noticing the face and its function. To tackle this question, I am doing two lines of studies. One is the time series analysis of scan path during watching face image and its development. The comparison of the development between the scan path to face and house suggest that processing of facial information is specialized even in infants. The other one focuses on the multimodal representation of face in infants. By using simple but ambiguous face-like image and sound, it is demonstrated that not 8- but 10-, and 12-month-olds acquired the representation that a face has a mouth and the mouth is the generator of the sound.