Prolonged listening to a repeated sequence consisting of low and high tones produces spontaneous transitions in the perceptual state between a single coherent stream and two segregated streams. The dissociation between constant physical stimulation and fluctuating perceptual experience in auditory streaming provides a compelling means for studying how auditory percepts are formed in the brain. First, we psychophysically examined the nature of such perceptual transitions in various frequency differences (Δf) between low and high tones. After the initial buildup of streaming, perceptual transitions occurred frequently even at the Δfs that were previously thought to produce a stable percept, and the dominant perceptual state changed depending on Δf. Next, we explored brain activities correlated with the perceptual transitions using functional magnetic resonance imaging (fMRI). An event-related analysis revealed that the auditory cortex and thalamus were activated at the timing of perceptual transitions. The response onset of the activity in the auditory cortex was earlier than that in the thalamus for the transition from dominant to non-dominant percepts, and later for the transition from non-dominant to dominant percepts. These results imply that the interaction of the auditory cortex and thalamus plays a crucial role for perceptual transitions in auditory streaming.