Caused motion events (e.g., a boy pulls a box into a room) are basic events where an Agent (the boy) performs an Action (pulling) that causes a Figure (box) to move in a spatial Path (into) to a Goal (the room). These semantic elements are mapped onto lexical and syntactic structures differently across languages. This dissertation investigates the encoding of caused motion events in Turkish, and the development of this encoding in speech and gesture. First, a linguistic analysis shows that Turkish does not fully fit into the expected typological patterns, and that the encoding of caused motion is determined by the fine-grained lexical semantics of a verb as well as the syntactic construction the verb is integrated into. A grammaticality judgment study conducted with adult Turkish speakers further establishes the fundamentals of the encoding patterns. An event description study compares adults’ verbal and gestural representations of caused motion to those of children aged 3 to 5. The findings indicate that although language-specificity is evident in children’s speech and gesture, the development of adult patterns takes time and occurs after the age of 5. A final study investigates a longitudinal video corpus of the spontaneous speech of Turkish-speaking children aged 1 to 3, and finds that language-specificity is evident from the start in both children’s speech and gesture. Apart from contributing to the literature on the development of Turkish, this dissertation furthers our understanding of the interaction between language-specificity and the multimodal expression of semantic information in event descriptions.