We examined the distributional nature of input which determines infants’ learning of linguistic rules and exceptions. Fourteen-month-olds were trained with sentences of an unknown natural language going through a movement rule. Input was 100% rule-governed (Exp1), or contained noise instances exhibiting no movement (Exp2-3). During Test, novel (Exps1-2) and noise (Exp3) sentences obeying the trained rule versus another movement rule were presented. Infants discriminated the rules being applied to novel sentences, following training with rule instances of 100% and 80% type frequency. In the 80%-type training, we controlled for the overall token frequency for rule versus noise instances (both 50%). The token-per-type frequency of noise was high (80%), but did not disrupt rule generalization to novel instances. However, when the trained rule and non-trained rule were applied to those noise sentences during Test, infants showed no discrimination, i.e., interpreting the noise as exceptions rather than overgeneralizing them to the rule.