Speech understanding is influenced by not only the presence, but also the specific nature of maskers. Noise maskers primarily result in energetic masking, whereas speech maskers create additional interference due to linguistic and acoustic similarities to the target. The present study examined the influence of different types of maskers and target word position on the immediate recall of words in sentences by normal-hearing younger adults. In Experiment 1, nonsense sentences with 3 keywords (e.g., A "house" should "dash" to the "bowl"). were presented against a background of speech-spectrum noise or two-talker nonsense speech. With the speech masker, accuracy increased with word position. With the speech-spectrum noise masker, performance was highest for the first word and did not vary linearly with word position. In Experiment 2, when the speech-masker was noise-vocoded to preserve envelope information while disrupting fine structure cues and minimizing semantic content, performance was similar to that found with the speech-spectrum masker. The results suggest that the ability to track a target sentence in conditions of informational masking improves as the target utterance unfolds over time.