The Enhancing Effect of Subsequent Context on Perception of the Sentence-initial Word.

Takayuki Kawashima*, Makio Kashino †, Takao Sato*

*Department of Psychology, Division of Humanities and Sociologies, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113, Japan † NTT Basic Research Laboratories, 3-1 Morinosato-Wakamiya, Atsugi, Kanagawa, 243-0198, Japan

Abstract: Phonemic restoration effects under periodic noise replacement were compared for two types of Japanese sentences. In the first condition, the initial word of a sentence was semantically predictable from the subsequent context. In the second condition, the initial word was not directly predictable from the subsequent context. The intelligibility of the sentence-initial words was higher in the former condition than the latter condition (58% vs. 10%), indicating that word perception can be enhanced by subsequent semantic context in Japanese.

INTRODUCTION

When a portion of speech is deleted and replaced by a louder sound, under appropriate conditions, listeners perceive the deleted sound as present and are unable to distinguish the perceptually synthesized segment from those actually present (1, 2). This phonemic restoration effect demonstrates the robustness of speech perception. It has been shown that phonemic restorations rely on several information sources, such as coarticulation (3) and semantic context preceding the target (4). However, with the exception of one study (5), we found no formal examination of the role of temporally subsequent context in phonemic restoration. We conducted an experiment to demonstrate the facilitating effect of subsequent context in the perception of preceding words in Japanese sentences.

METHOD

Intelligibility of the initial word was compared for two types of sentences under periodic noise replacement. In the first condition, the initial word of a sentence was semantically predictable from the subsequent context (High-Predictability condition HP). In the second condition, the initial word was not directly predictable from the subsequent context (Low-Predictability condition LP).

For the creation of stimuli, we first selected 300 words (2-4 moras each), then made two types of sentences for each initial word. We presented these sentences without initial words to 13 subjects and asked them to write down a word or two they thought appropriate to fill the blanks (after (6)). When more than 7 subjects reported the word the experimenter intended, the sentence was regarded as a HP sentence. When all subjects reported an intended word, the sentence was excluded. From these sentences, 120 were selected for the main experiment. No subjects reported the intended word for any LP sentence. The length of the sentences was roughly the same for both types of stimuli.

These sentences were produced by a male experimenter and recorded to DAT recorders with a low-pass filter set at 6 kHz. The sentence-initial target words and the two types of carrier sentences were recorded separately and digitally concatenated later. The concatenated sentences were periodically interrupted each 130 ms and the gap filled with 130-ms noise bursts (50 percent duty cycle). The noise bursts were generated digitally in every replacement. An initial 130-ms interval was always interrupted and filled with noise bursts. The noises were low-pass filtered at 3750 Hz. and their level was above that of the speech signal.

The stimuli were presented diotically through headphones (SENNHEISER HD-265). Sixty HP sentences
and 60 LP sentences were presented in random order to 14 subjects. None of the subjects had participated in the previous stimuli creation tasks. The subjects were instructed to repeat aloud these sentences, and their response were recorded. They were allowed to guess at any part of a sentence if they were not sure.

RESULTS

When a reported target word was literally the same as the intended target word, the response was counted as correct. For HP sentences the group mean percent correct was 58%, and for LP sentences the group mean percent correct was 10% (Fig. 1). This difference is statistically significant ($t(13) = 15.5, p < 0.01$). This result indicates that identical target words were better restored when their were semantically related to the subsequent sentences. The intelligibility of the subsequent sentences were also calculated. In this case, each sentence was segmented into a few semantic units, and when a reported segment was literally the same as the presented segment, the response was counted correct. For HP sentences the group mean percent correct was 71% and for LP sentences the group mean percent correct was 66% (Fig. 1). Though this difference was also statistically significant ($t(13) = 2.26, p < 0.05$), it was much smaller than for the target words.

![Figure 1](image)

**FIGURE 1.** The intelligibility of sentence-initial words and carrier sentences for high and low predictability sentences subjected to multiple phonemic restorations.

DISCUSSION

The results show that the perceptual restoration of the initial word in Japanese sentences is better in HP sentences than in LP sentences, suggesting that temporally subsequent sentence context enhances the perception of the preceding word.

Currently we are testing whether the difference observed between HP and LP reflects sensory processes or some non-sensory response biases.

REFERENCES