Ann. Bull. RILP, No. 14, 217-224 (1980)

CONTEXT EFFECTS IN PHONETIC AND NON-PHONETIC VOWEL JUDGMENTS*

Sumi Shigeno** and Hiroya Fujisaki***

1. Introduction

The difference between the proactive (forward) and retroactive (backward) influences of a context stimulus upon the perception of a target stimulus has been studied at various levels, and has invoked a number of interpretations. In a study of temporal sensory masking, for example, the effect of a context stimulus (i. e., a masking stimulus) upon a target stimulus (i. e., a masked stimulus) has been regarded as an interference to signal processing in the auditory nervous system. In a study of the anchoring effect, on the other hand, the effect of a context (i. e., an anchor) has been recognized either as response bias (Parducci, 1975) or auditory contrast (Simon and Studdert-Kennedy, 1978). There appears to be one feature which is more or less common to the results of these investigations. Elliott (1962) found that forward masking was greater than backward masking when the inter-stimulus interval exceeds 10 ms, while the situation was reversed for shorter inter-stimulus intervals. Massaro (1973) found that the accuracy of identification of a test tone was decreased by the presence of a preceding or a following tone, and that the decrease was greater in the backward than in the forward masking conditions at 0-, 20- and 40-ms silent intervals. Massaro called this phenomenon "recognition masking," and postulated the existence of a preperceptual auditory image in order to explain the experimental results.

Several perceptual models have been proposed for the forward and backward context effects. The present authors (1977, 1979) investigated the forward context effect, using pure tones, complex tones with a formant, and synthetic vowels. In almost all the experiments, when the shift in the category boundary due to addition of a context was adopted as the index, the context effect was found to vary continuously from assimilation to contrast as the context-target separation is increased on the stimulus continuum or on the time axis. The tendency toward contrast was found to increase with stimulus complexity. A model for the perceptual mechanisms and processes was presented to account for these results, taking into consideration two kinds of short-term memory (i. e., precategorical and categorical). It was hypothesized that the magnitudes of the influences of the two short-term memories upon the shift of category boundary vary depending on the attributes of the stimuli. Repp (1978) indicated that, in isolated synthetic vowel-consonant-vowel (VCV) utterances where C stands for a stop consonant, VC and CV transitions interact assimilatively or contrastively according to the closure durations, and he hypothesized the precategorical and the categorical

* Paper presented at the 10th International Congress on Acoustics, Sydney, July 9-16, 1980.

** Department of Psychology, University of Tokyo

*** Department of Electrical Engineering, University of Tokyo

processes of temporal integration in phonetic perception.

The present study was conducted to investigate the forward and backward context effects in more quantitative terms, and to examine the roles of the two short-term memories. The effects of the inter-stimulus interval upon the phonetic and the non-phonetic categorical judgments of synthetic vowels were investigated, and the shift in the category boundary, measured individually for each subject, was used as an index for the magnitude of the context effect.

2. Method

Subjects

Four female adults with normal hearing participated in the experiment. They had previously taken part in experiments using synthetic sound stimuli.

Stimuli

The stimuli were synthetic vowels with five formants, generated by computer simulation using a terminal-analog speech synthesizer. The stimuli were read out at a sampling rate of 10 kHz with an accuracy of 10 bits, and converted into the analog waveform. They were recorded through a lowpass filter with a cutoff frequency of 4.5 kHz.

Seven points were selected at equal intervals as targets on the two stimulus continua, as shown in Fig. 1. In phonetic judgments the stimulus numbers 4, 6, 8, 10, 12, 14 and 16 were used.

The context stimulus was chosen so as to be 96 Hz in the fundamental frequency (F_0) in the case of non-phonetic judgment, or 330 Hz in the first formant frequency (F_1) in the case of phonetic judgment. Other formant frequencies (F_2, F_3, F_4 and F_5) were held constant at 1250 Hz, 2750 Hz, 3500 Hz and 4500 Hz, respectively. All stimuli were 200 ms in duration and had 20 ms rise-fall times.

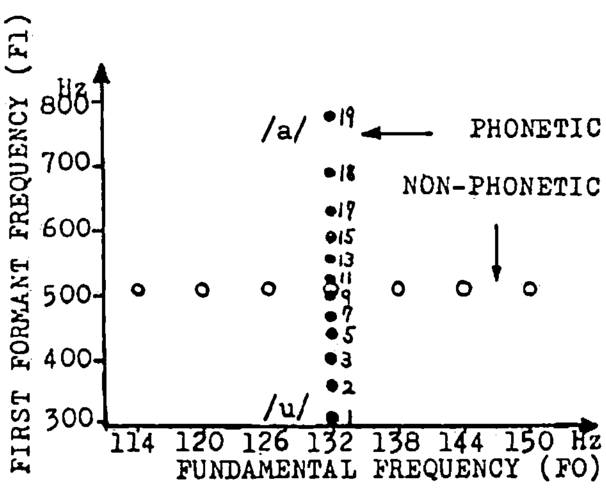

Fig. 1 Stimuli used in Experiment.

Procedure

For both phonetic and non-phonetic judgments, the experiment was conducted under three conditions:

(1) Control condition

The seven target stimuli were randomly presented at intervals of 4 s. The subjects were asked to identify each target stimulus as belonging to either of two categories: /u/ or /a/ in the case of phonetic judgment, and "low" or "high" in the case of non-phonetic judgment.

(2) Forward condition

As shown in Fig. 2, each target stimulus was preceded by the context stimulus. The silent inter-stimulus interval (ISI) between context and target was 100, 200 or 500 ms. The subjects were asked to identify the target stimulus as belonging to either of two categories, which were the same as used in Condition (1).

(3) Backward condition

As shown in Fig. 2, each target stimulus was followed by the context stimulus. The inter-stimulus interval between target and context was 100, 200 or 500 ms. The subjects were asked to identify the target stimulus as belonging to either of two categories, and were also instructed to write down the response after they heard the following stimulus.

The intertrial interval was 4 s. Forty judgments per target stimulus were obtained from each subject for the three values of ISI in the control, forward and backward conditions. In the case of non-phonetic judgment, at the beginning of each experimental session the subjects were requested to identify the target stimulus as "low" or "high."

The recorded stimuli were presented to the subjects in an anechoic room through a loudspeaker at a sound pressure level of about 78 dB (c).

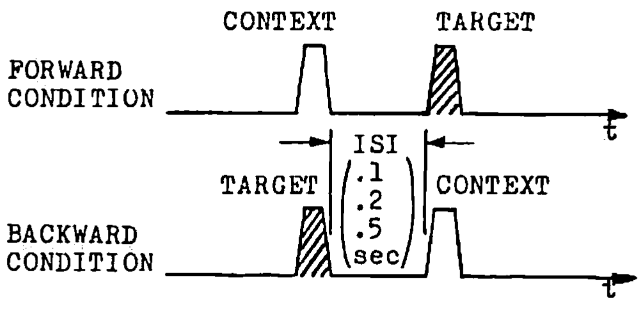

Fig. 2 Presentation of stimuli.

3. Results

In the case of categorical judgment, the distribution of the response probability on the stimulus continuum can be approximated by a cumulative normal distribution. The probability of the /a/ response was calculated in the case of phonetic judgment, while the probability of the "high" response was calculated in the case of non-phonetic judgment. In each case, the mean μ_i and the standard deviation σ_i were estimated from measured probabilities by the maximum likelihood method. The suffix i indicates the condition: (1) control, (2) forward, (3) backward. The mean μ_i corresponds to the category boundary, while the standard deviation σ_i serves as an index for the accuracy of identification. If we denote the increase/decrease in μ from the control condition by $\Delta\mu$, the ratio of $\Delta\mu$ over σ_1 can then be regarded as an index for the context effect. A plus sign for $\Delta\mu/\sigma_1$ indicates a shift of the category boundary toward the context stimulus (meaning contrast), while a minus sign indicates a shift in the opposite direction (meaning assimilation).

The mean identification performance of all the subjects is plotted as a function of the inter-stimulus interval in Fig. 3. The shift toward contrast is conspicuous both in the forward and backward conditions in phonetic judgment, but contrast is hardly observed in non-phonetic judgment.

Though contrast reduced with the increase of the inter-stimulus interval, the context effect appeared to be more contrastive at every inter-stimulus interval in the backward condition than in the forward condition.

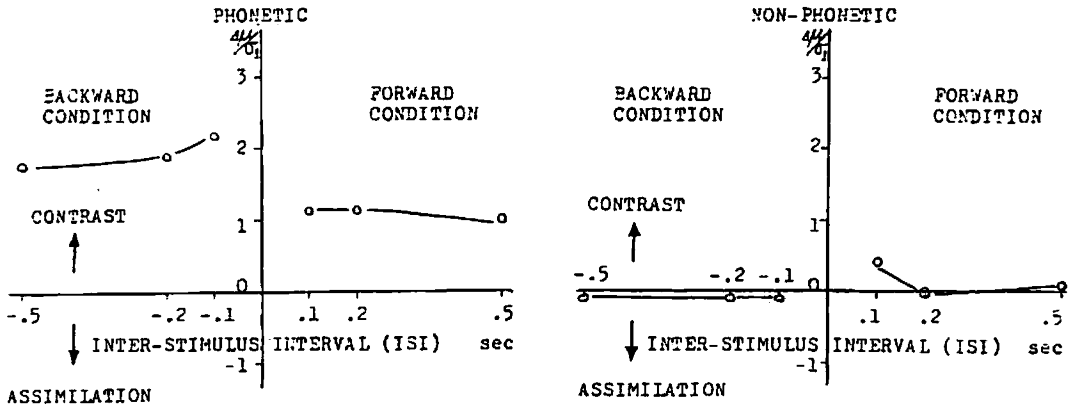

Fig. 3 Context effects shown as the function of ISI (averaged results).

4. Discussion

The forward context effect has been a subject of many experimental studies, and various interpretations have already been proposed. We previously (1977, 1979) presented an interpretation based on a model of the perceptual mechanisms and processes, as shown in Fig. 4. Figure 4 shows the auditory perceptual processes with two short-term memories: one (STM_1) being the precategorical short-term memory and the other (STM_2) being the categorical short-term memory. The first block (AUDITORY MAPPING) represents the auditory process for mapping the stimulus continuum to the sensation continuum. The output from AUDITORY MAPPING is stored in STM_1. The output from AUDITORY MAPPING is at the same time fed to the next stage (CATEGORICAL JUDGMENT). The output from CATEGORICAL JUDGMENT is stored in STM_2. The category boundary for the identification of the stimulus is determined by the influence of a long-term factor and short-term factors. The former is the information stored in the long-term memory LTM, which has been acquired during many years of language learning in the case of phonetic judgment, but by a very brief period of learning in the case of non-phonetic judgment. The latter are the influences of two kinds of short-term memory STM_1 and STM_2. The process in which the category boundary is determined by LTM, STM_1 and STM_2 is shown as INTEGRATION. It is hypothesized that assimilation is caused when the influence of STM_1 upon the determination of the boundary is greater than that of STM_2, while contrast is caused when the influence of STM_2 is greater than that of STM_1. It is also hypothesized that the influence of STM_2 is greater in phonetic judgment than in non-phonetic judgment. The results of the present experiment tend to support this interpretation.

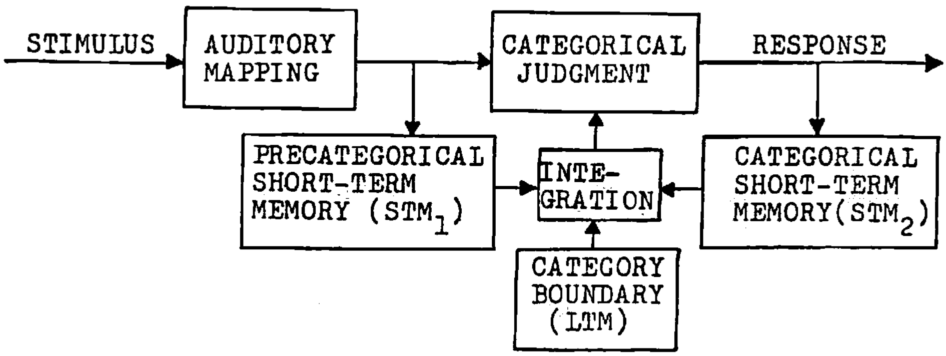

Fig. 4 A model for the process of categorical judgment with two STM's as possible causes of forward context effects.

On the other hand, the results of our present study as well as of previous studies on the backward context effect all indicate that the backward effect is greater than the forward effect at short values of ISI. In order to account for the backward context effect, the above model has to be modified and redrawn as in Fig. 5. Here we assume that a target stimulus is followed by a context stimulus. In the case of phonetic judgment, the target stimulus comes first and is mapped on the sensation continuum at AUDITORY MAPPING. The output from AUDITORY MAPPING is stored in STM_1 and at the same time is fed to CATEGORICAL JUDGMENT, and the result is stored in STM_2. The context stimulus presented after a brief interval is processed in a similar way, and the results are also stored both in STM_1 and in STM_2. The category boundary, however, is influenced mainly by the information stored in STM_2 and is shifted toward the direction of contrast. The phonetic judgment once made upon the target is then corrected on the basis of the new boundary. In the case of non-phonetic judgment, on the other hand, the target stimulus is mapped on the sensation continuum at AUDITORY MAP-

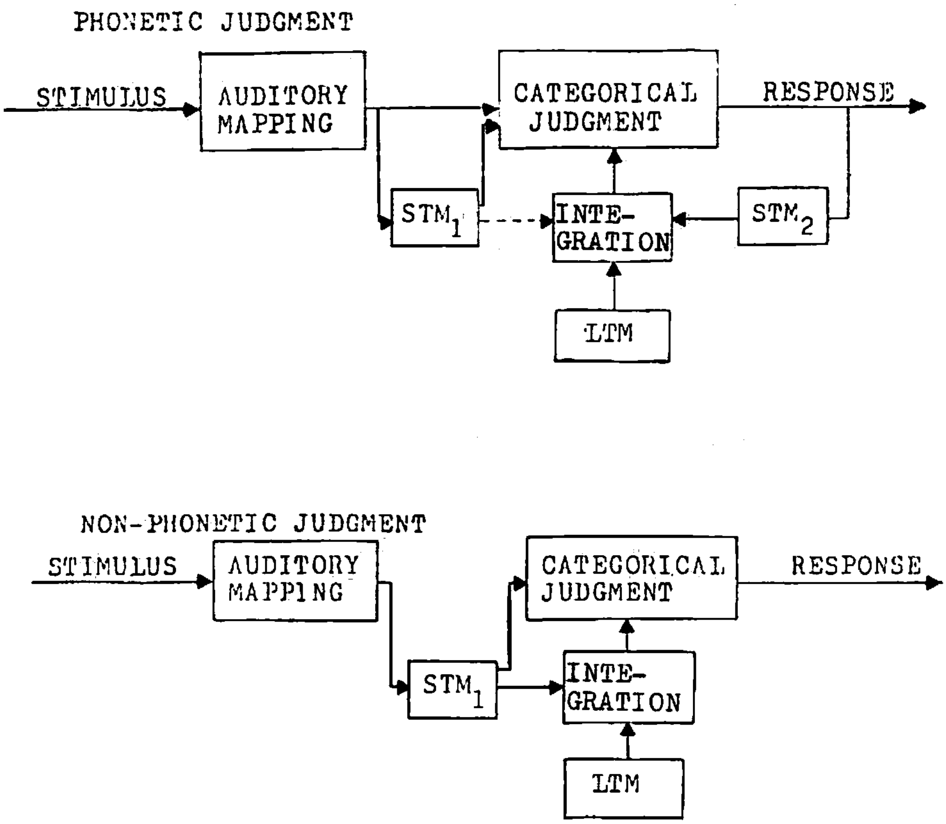

Fig. 5 A model for the process of categorical judgment as possible causes of backward context effects in phonetic and non-phonetic judgments.

PING. The output from AUDITORY MAPPING is stored in STM_1 but is not categorized until the context stimulus is processed at AUDITORY MAPPING after a brief interval. The context stimulus after auditory mapping is stored also in STM_1, but not in STM_2, since it is not categorized. Thus the category boundary is mainly influenced by the information stored in STM_1 and is shifted toward the direction of assimilation. The target stimulus is then categorized on the basis of the boundary which has undergone the influence of the context stimulus.

Taken together, these models can account for the fact that the backward effect is much greater than the forward effect in phonetic judgments, while no significant difference of magnitude is found between forward and backward effects in non-phonetic judgments. In phonetic judgments, categorical judgment of the target stimulus is under the strong influence of the result of categorical judgment of the context stimulus stored in STM_2, and the influence varies to a large extent depending on whether the context stimulus comes before or after the target. While a forward context stimulus exerts its influence only after the lapse of an ISI, a backward context stimulus immediately exerts its influence of the categorization of a preceding target stored in STM_1. In non-phonetic judgments, on the other hand, the categorical judgment does not take place at all on the context stimulus, so that it can exert only weak influence on the categorical judgment of the target stimulus through STM_1, regardless of whether or not the context stimulus comes before or after the target.

5. Conclusion

Context effects along phonetic and non-phonetic continua of synthetic vowels have been investigated with special emphasis on the temporal relationship (order and separation) between the context and the target stimulus. The results indicated (1) that the context effect is generally greater in phonetic judgments than in non-phonetic judgments, and (2) that the backward effect is more contrastive than the forward effect in phonetic judgments, but less contrastive in non-phonetic judgments. A model is presented to account for these results, taking into consideration two kinds of short-term memory: precategorical and categorical.

References

Elliott, L. L. (1962): "Backward and forward masking of probe tones of different frequencies," Journal of the Acoustical Society of America, 34, 1116-1117.

Massaro, D. W. (1973): "A comparison of forward versus backward recognition masking," Journal of Experimental Psychology, 100, 434-436.

Parducci, A. (1975): "Contextual effects: A range-frequency analysis," In E. C. Carterette & M. P. Friedman (Eds.), Handbook of Perception (Vol. II), New York: Academic Press.

Repp, B. H. (1978): "Perceptual integration and differentiation of spectral cues for intervocalic stop consonants," Perception & Psychophysics, 24, 471-485.

Shigeno, S. and H. Fujisaki (1977): "Anchoring effects in the perception of speech and non-speech stimuli," Reports of the Spring Meeting of the Acoustical Society of Japan (Apr. 1977) 2-4-10.
Shigeno, S. and H. Fujisaki (1979): "Effect of a preceding anchor upon the categorical judgment of speech and non-speech stimuli," Japanese Psychological Research, 21, 165-173.
Simon, H. J. and M. Studdert-Kennedy (1978): "Selective anchoring and adaptation of phonetic and non-phonetic continua," Journal of the Acoustical Society of America, 64, 1338-1357.