Category boundary effect for grapheme perception

MAKOTO YASUHARA and THEODORE T. KUKLINSKI
*Research Laboratory of Electronics, Massachusetts Institute of Technology
Cambridge, Massachusetts 02139*

In order to discover whether or not discrimination training is effective in grapheme perception, as it is in phoneme perception, we first sought for a category boundary effect. We found that the discrimination function over the character continuum peaks at the category boundary, as it does over the speech continuum. Our next concern was to check whether or not the same effect could be observed for a noncharacter continuum in which neither grapheme category nor, therefore, training processes were involved. The experimental results indicated that the category boundary effect was not unique to speech sounds but was observed for graphemes as well, and that the effect was the result of linguistic experience.

A series of experiments on the identification and discrimination of speech sounds have been reported, revealing that certain classes of speech sounds such as stop consonants were perceived in a special mode, unlike other auditory stimuli. Liberman, Harris, Hoffman, & Griffith (1957) found that synthetic speech stimuli which varied in acoustically equal steps through the range sufficient to produce the stop consonant /b/, /d/, and /g/ were perceived as members of discrete categories. When listeners were asked to discriminate pairs of these stimuli by an ABX procedure, they were able to discriminate stimuli drawn from different phonetic categories but could not discriminate stimuli drawn from the same phonetic category. The obtained discrimination functions were not monotonic with changes in the physical scale but showed marked discontinuity at points along the continuum that were correlated with changes in identification.

The phenomenon of this "categorical perception" was considered to be quite unusual when compared with the results typically obtained in most psychophysical studies. Stimuli that lie on a single continuum are ordinarily perceived continuously and the discrimination functions are typically monotonic with the physical scale.

These findings led to a similar experiment with vowel stimuli that varied in acoustically equal steps through the range /I/, /ε/, and /ae/. Fry, Abramson, Eimas, and Liberman (1962) reported that these stimuli were perceived continuously, much like nonspeech stimuli. The discrimination functions did not yield

discontinuities along the continuum that were related to changes in identification, but were relatively constant across the whole continuum. Similar differences in perception between consonants and steady-state vowels have been reported by Stevens, Liberman, Studdert-Kennedy, and Öhman (1969).

The results of these experiments have led investigators to propose two different modes for the perception of speech sounds: Categorical or phoneme mode and a continuous or auditory mode. The original explanation of the differences for categorical perception of consonants and continuous perception of vowels was offered by the Haskins investigators (Liberman, Cooper, Harris, & MacNeilage, 1962) in terms of the articulatory differences in the production of these sounds. The motor theory of speech perception proposed that reference to articulation is a mediating stage between the incoming acoustic signal and its perception.

A second explanation of categorical perception has been proposed by Lane (1965), who suggests that categorical perception results from a history of reinforcement with respect to labeling behavior. Lane claimed that the postulation of a "special" perceptual mechanism for speech is unwarranted since categorical perception is not unique to speech sounds but may be obtained with nonspeech continua as well. Lane tried to demonstrate similar categorical perception effects with nonspeech stimuli. Using a variety of visual and auditory stimuli, he tried to show that if two stimuli sampled from the same continuum cause different identification responses, they will be more readily discriminated than two stimuli (separated by equal physical units) that cause the same identification responses. Lane maintained that the observed relations between identification and discrimination that had been used as evidence for a "special" speech perception mode were not peculiar to the speech sound,

The authors would like to thank Professor Murray Eden, Professor Barry Blesser, and Dr. Robert J. Shillman of the Research Laboratory of Electronics, M.I.T., for their discussions of the present study. M. Yasuhara is presently on leave from the University of Electro-Communications, Tokyo 182, Japan.

but, rather, were results of general procedures for discrimination training and testing with discrete response repertoires.

Lane and Schneider (Note 1) trained subjects to identify the extreme members of the nonspeech continuum as /do/ or /to/. Following this training, the responses generalized to the intermediate stimuli such that the stimulus continuum was partitioned into two equivalent response classes. Before the training procedure with the extreme stimuli was initiated, the ABX discrimination functions for the nonspeech continuum showed chance performance. However, after training, the discrimination functions showed peaks at the boundary between the two response categories. Lane contended that categorical perception could be obtained quickly and easily in the laboratory after a few minutes of conditioning. If this is the case, there would be little justification for considering the perception of speech as special, involving perceptual operations and mechanisms basically different from those used in the processing of other auditory stimuli.

The Haskins group undertook to replicate the Lane and Schneider experiment. Their attempts to produce categorical perception by discrimination training were summarized in their reply to Lane (Studdert-Kennedy, Liberman, Harris, & Cooper, 1970). Their results indicated that they were no more successful than he was in establishing categorical perception by discrimination training procedures.

A careful inspection by the authors of the data which Lane presented (1965) reveals that the discrimination functions are in every case less consistent than the data shown by Lane for the one particular subject as representative.

Another available explanation of the differences between categorical and continuous perception of speech sounds was suggested by Fujisaki and Kawashima (Note 2, Note 3). They proposed a model of the perpetual process involved in speech discrimination that considered the separate contributions of phonetic and auditory short-term memory. They reported that vowels of very short duration tended to be perceived more categorically than longer vowels, and suggested that the discrepancy in discrimination between the obtained ABX scores and those predicted from the standard Haskins identification formula¹ were due to a component from auditory memory for timbre information. To compensate for this discrepancy, they proposed a new prediction formula which included both the contribution due to identification and that due to auditory memory. The procedure introduced by them substantially improved the fit of the obtained and predicted discrimination functions and suggested that auditory short-term memory might be involved in some way in the decision processes in speech discrimination tasks.

Although the differences between categorical and continuous perception have stimulated a great deal of discussion, the differences between these two modes of perception are not well understood. Pisoni (1971) examined the differences between categorical and continuous perception and evaluated the above-mentioned three accounts for the phenomena of categorical perception. One of the results suggested that the motor theory account of the difference in perception between consonants and vowels might no longer be tenable. Even vowels can be perceived categorically if stimulus duration is reduced, and the range, spacing, and number of categories may also be important factors contributing to the categorical perception observed with vowels. Secondly, a strong basis was obtained for inferring a role of auditory short-term memory in speech discrimination. Vowel discrimination was found to be better than consonant discrimination because auditory short-term memory for acoustic properties of vowels was more readily available for use in discrimination while auditory short-term memory for the acoustic properties of consonants was less readily available for use in discrimination.

In recent years, a number of investigators have assumed that underlying the perception of speech sounds are feature detectors that are especially tuned to the physical properties of the speech signal (e.g., Stevens, Note 4). The evidences supporting a detector model of feature extraction have come primarily from selective adaptation experiments (e.g., Eimas & Corbit, 1973).

Recent Theory of Character Recognition

One general approach to the character recognition problem is feature detection (Neisser, 1967; Sternberg, 1967). It embodies the notion that class membership is determined by the presence or absence of essential properties or features. Blesser, Shillman, Kuklinski, Cox, Eden, & Ventura (1974) claimed that the use of the word "feature" should be avoided because its meaning is often unclear, sometimes referring to a physical measure and sometimes to an abstraction. Instead, they used two terms to replace the notion of feature: "physical attribute" and "functional attribute." Physical attributes are the parts of the character, usually described in terms of line, angle, intersections, etc., which make up the physical image. Functional attributes are abstractions which provide an underlying description of letters. Under this formulation, Blesser et al. have proposed a general approach for studying the functional rather than the physical attributes of letters.

Figure 1 presents two series of characters which have been generated by variation along one physical dimension, the location of an intersection. For each character in Figure 1, the state of the attribute LEG

Figure 1. Two trajectories involving the attribute LEG.

is indicated by a plus sign for presence and a minus sign for absence. A certain character in a series of the generated characters is ambiguous because the character is just as likely to be labeled as F as it is C. Thus, this character defines the labeling boundary along the C-F trajectory.

Such ambiguous characters form the essential part of the theory. An ambiguous character of necessity has a functional attribute in transition from one state to another. By focusing on ambiguous characters, functional attributes become evident. Shillman (1974) identified a set of ambiguous characters through exhaustive pairwise comparisons of the 26 letters of the English alphabet. Many of the ambiguous characters can be grouped together to form subsets in which each member appears to have the same functional attribute in transition. For example, Figure 2 shows a group of four ambiguous characters which are postulated as involving the same functional attribute LEG in transition; each of these lies on V-Y, C-F, D-P, and O-P trajectories, respectively. By studying ambiguous characters in this manner, Shillman isolated 12 functional attributes which he hypothesized to be relevant to human character recognition. Among them were SHAFT, ARM, LEG, CLOSURE, BAY CROSSING, etc.

Some empirical tests have been reported for determining the psychological validity of any postulated attribute (Blessner, Kuklinski, & Shillman, 1976; Kuklinski, 1975). One of the tests indicated if a particular attribute was involved in a particular letter, and the other tests investigated the commonality of an attribute among different letters using a selective adaptation technique. Some psychological evidence was provided concerning which and how many relevant attributes were involved in a particular set of letters from reaction time measurements (Yasuhara, Note 5).

Research Objectives

The conflicting results of the Lane and the Haskins group experiments with nonspeech sounds indicated a real need for additional attempts at establishing

categorical perception by discrimination training. Pisoni tried to replicate the original Lane and Schneider training procedures with a stimulus variable shown by Mattingly, Liberman, Syrdal, and Halwes (1971) to be perceived categorically in speech and continuously in nonspeech. The results of his experiment were not conclusive with respect to resolving the controversy over the effects of discrimination training on categorical perception. However, little evidence has been found to support the Lane position. The outcome of this experiment does not permit us to exclude the possibility that more adequate training processes may improve the fit between the discrimination function predicted from the standard Haskins identification formula and that obtained from experiments.

Our primary concern here is to know whether the effect of discrimination training exists for the process of human visual perception. Each character has its own category, grapheme, as one of the symbols, and most adults have acquired their own idea of grapheme category for each character through the course of long and adequate discrimination training processes from early childhood. This understanding is quite comparable with that for phonemes in the speech case. Thus, it is our first interest to check whether the category boundary effect occurs for grapheme continua, as in the case of phoneme continua. That is, does the discrimination function along character continua peak at the category boundary as it does along speech continua? If this is the case, the next thing to check will be whether the same effect can

Figure 2. Four characters which are ambiguous due to the transitional state of the functional attribute LEG.

be observed for noncharacter continua in which neither grapheme category nor, therefore, training processes are involved.

METHOD

Subjects

The subjects were six university students who were native speakers of English. They took part in the experiment 10 times each over 10 successive weekdays and were paid \$2 for each trial.

Stimuli

Figure 3 presents a series of character stimuli which were generated along the dimension of a functional attribute CROSSING by variation of a physical value: the location of an intersection, starting at an archetypal V and ending at an archetypal X. Each generated character has its own 1/L value assigned from .000 (V) to .500 (X) with the increment of .033.

Each stimulus was 75-98 mm in width $\times$ 75 mm in height, drawn carefully and photographically reduced to 30-39 $\times$ 30 mm on a white 10 $\times$ 15 cm card.

Apparatus

The apparatus used was a one-field Ralph Gerbrands tachistoscope. The stimulus cards were so mounted that the stimuli were centered in the subject's viewing area, 9 $\times$ 14 cm. The viewing distance was approximately 57 cm.

Procedure and Stimulus Presentation

For the identification part of the experiment, the events of a single trial of the experiment were as follows: After the signal given by the experimenter, the subject pressed the start button, triggering the exposure of one of the 16 shuffled stimulus cards for 300 msec and starting the counter for measuring the reaction time. By the voice response of the subject, either V or X, the counter was stopped, giving the reaction time.

To test discrimination by the ABX procedure, the stimuli were arranged in triads, each of which consisted of an A stimulus, a B stimulus, and a third stimulus, X, which was identical either to A or to B. The subjects were instructed to determine whether X was the same as A or the same as B. It was strongly emphasized to each subject that he was to make his judgment on the basis of any cues he could see. The measure of discriminability was the percentage of the time that the X stimulus was correctly matched to A or B. Each stimulus was paired with the stimuli lying on both of its sides on the continuum shown in Figure 3. Thus, for example, Stimulus 2 was paired with Stimulus 1 to form the A and B stimuli, respectively, of one ABX triad, and was paired also with Stimulus 3 to form the A and B stimuli, respectively, of another ABX trial. This was to cancel the effect caused by the unbalance of the spatial presentation of each stimulus. From the above discussion, it followed that the number of A and B combinations was 30. For each combination of A and B stimuli, there were two triads—one in which X was identical to A and one in which X was identical to B. Therefore, the total number of triads

was 60. These triads were shuffled and arranged into one deck of 180 cards. The events of a single trial of the discrimination experiment were as follows: After the signal given by the experimenter, the subject pressed the start button, exposing the A stimulus of the ABX triad for 300 msec, followed 1 sec later by the exposure of the B stimulus for 300 msec, followed 1 sec later by the exposure of the stimulus X for 300 msec.

RESULTS

The experiment was divided into two parts: Experiment 1, to check whether discrimination enhancement could be observed between stimuli in the restricted region along the V-X trajectory (character condition), followed by Experiment 2, to check whether the same effect could be obtained for noncharacter visual stimuli (noncharacter condition).

Experiment 1: Character Condition

The stimuli were presented to the subjects in two sessions: Session 1, where the stimuli were presented singly to determine how the subjects would label them, as V or X, followed by Session 2, where the stimuli were presented in an ABX arrangement to determine to what extent the subjects could discriminate them on any basis at all.

Identification functions. The individual identification functions obtained for three subjects are shown in Figures 4, 5, and 6. Each identification function indicates that the subjects partitioned the stimulus continuum into two character categories: V and X. The category boundary can be identified as the character which is assigned character labels with equal probability. Another possible method of locating the category boundary would be to measure the time taken to arrive at a response, reaction time, when subjects are asked to identify the character. The averaged reaction time data are shown simultaneously by a dashed line in each figure. We see that the maximum reaction time occurs for stimuli closest to the identification curve crossover. The identification and reaction time measures are consonant and are convincing evidence of the location of the category boundary for each subject.

Inspection of each function reveals that the shift from one response category to another appears to be somewhat gradual. This initial finding is in agreement

Figure 3. Characters along the V-X trajectory.

Figure 4. Identification and discrimination functions along the V-X trajectory for Subject 1.

with that reported in the previous experiments on the identification of vowels (Stevens et al., 1969).

Discrimination functions. The individual ABX discrimination functions for the same three subjects as before are shown below the corresponding individual identification functions in Figures 4-6. Note that the discriminability is higher for the stimuli with the lower 1/L value. As mentioned previously, each generated character has its own 1/L value assigned from .000 (V) to .500 (X) with equal increments of .033. If we were to assign each 1/L value with equal increments in the decibel scale, comparable discrimination functions would be obtained with those for normal speech cases.

A careful examination of each of the obtained discrimination functions indicates that each discrimination function has a peak at the point along the V-X trajectory which corresponds closely to category boundary, and that this is quite consistent for all

three subjects. If there were no special effects, the obtained discrimination functions would show only a monotonic decrease with the 1/L value.

Experiment 2: Noncharacter Condition

Three subjects, different from those of Experiment 1, took part in the experiment only for the purpose of obtaining discrimination functions in an ABX arrangement with the stimulus cards presented in unfamiliar (inverted) orientation. Of course, there was no preceding identification session.

The average discrimination function for non-character stimuli for the three subjects is shown in Figure 7. An inspection of the average discrimination function reveals that there appears to be no discontinuity; only a continuous monotonic decrease is observed. Since the discrimination functions vary from one subject to another in their performances along the continuum, the average discrimination

Figure 5. Identification and discrimination functions along the V-X trajectory for Subject 2.

function might be, in general, less consistent. For this reason, the individual discrimination data for the three subjects are reproduced in Figure 8. These functions reveal that there are no consistent peaks and troughs along the whole range of the continuum among these three subjects. These data permit us to consider that no special effect exists in discrimination performances for the noncharacter condition as expected.

CONCLUSIONS

Though the amount of data used in these experiments is relatively small, the following conclusions are warranted.

The result of Experiment 1, character condition, indicated that the category boundary effect occurs also for grapheme continua, where "category boundary effect" refers to the relative differences in discrimination performance for within-category and

across-category comparisons where the physical differences between two stimuli compared are of equal unit value along the continuum. The category boundary effect was found, in Experiments 1 and 2, to result from a history of reinforcement with respect to categorizing behavior, or discrimination training.

Similar results were reported by Miyawaki, Liberman, Fujimura, Strange, and Jenkins (Note 6) in their cross-language study of the speech perception for the /r-l/ continuum: American listeners, to whom the phonetic distinction is familiar, showed a peak in discrimination function similar to that found in the discrimination of stop consonants. Japanese listeners, who are generally quite unfamiliar with the /r-l/ distinction, discriminated very poorly in comparison with the Americans. Thus, there was an effect of linguistic experience on perception in the speech mode.

In conclusion, as far as the category boundary effect is concerned, the effect was found not to be

Figure 6 Identification and discrimination functions along the V-X trajectory for Subject 3.

unique to speech sounds; it was observed for graphemes as well. The category boundary effect appears to be a result of linguistic experience.

DISCUSSION

We would like to restate the terminology used in the following discussion in order to make the definitions clear.

Noncategorical perception. Perceptual performance showing a discrimination function which is monotonic along a physical continuum, as observed in most ordinary psychophysical studies of noncategory stimuli.

Categorical perception. Perceptual performance showing a discrimination function with the category boundary effect which results from discrimination training.

The criteria for categorical perception were those suggested by Studdert-Kennedy, Liberman, Harris, and Cooper (1970): (1) "peaks" of high discrim-

ability between stimuli in restricted regions along the dimension studied; (2) "troughs" of discrimination performance *near chance* in regions on either side of the peak; and (3) correspondence between discrimination peaks and troughs and the course of identification functions, with peaks occurring at the identification boundary and troughs occurring within each perceptual category.

So far, it would be reasonable to consider two modes of categorical perception: An *abrupt* categorical perception and a *continuous* categorical perception. The former refers to satisfying all the three criteria mentioned above,² the latter satisfying all the three criteria except near chance in Criterion 2. The essential difference between abrupt and continuous categorical perception could be considered to rest on the magnitude of the category boundary effect, especially the level of within-category discrimination performance.

The disagreement between the Lane and the Haskins positions mentioned in the introduction appears to be

Figure 7. Average discrimination function for the noncharacter condition for three subjects.

resolved in the present study. However, the disagreement cannot be resolved satisfactorily until the real origins of the difference between abrupt and continuous perception are understood well. Since discrimination training alone cannot be considered to be sufficient to produce abrupt categorical perception of consonants, it might not be permitted to exclude the possibility of the special perceptual process for consonants. The difference between the perception of consonants and vowels may reflect the differences in the nature of the acoustic cues underlying these speech sounds.

REFERENCE NOTES

1. Lane H., & Schneider, B. *Discrimination control of concurrent response by the intensity, duration and relative onset time of auditory stimuli*. Unpublished report from Behavior Analysis Laboratory, University of Michigan, 1963.
2. Fujisaki, H., & Kawashima, T. *On the modes and mechanisms of speech perception*. Annual Report of the Engineering Research Institute, University of Tokyo, 1969, 28, 67-73.
3. Fujisaki, H., & Kawashima, T. *Some experiments on speech perception and model for the perceptual mechanism*. Annual Report of the Engineering Research Institute, University of Tokyo, 1970, 29, 207-214.
4. Stevens, K. *The potential role of property detector in the perception of consonants*. Paper presented at Symposium on Auditory Analysis and Perception of Speech, Leningrad, USSR, 1973.
5. Yasuhara, M. *On distinctive attributes in character recognition: Some evidence from reaction time measurement*. Annual Report of Research Laboratory of Communication Science, University of Electro-Communications, Tokyo, 1977.
6. Miyawaki, K., Liberman, A., Fujimura, O., Strange, W., & Jenkins, J. *Cross-language study of the perception of the F3 cue for /r/ versus /l/ in speech and nonspeech like patterns*. Status Report on Speech Research, Haskins Laboratories, 1973, SR-33, 67-76.

REFERENCES

- BLESSER, B., SHILLMAN, R., KUKLINSKI, T., COX, C., EDEN, M., & VENTURA, J. A theoretical approach for character recognition based on phenomenological attributes. *International Journal of Man-Machine Studies*, 1974, 4, 701-714.
- BLESSER, B., KUKLINSKI, T., & SHILLMAN, R. Empirical tests for feature selection based on a psychological theory of character recognition. *Pattern Recognition*, 1976, 8, 77-85.

Figure 8. Individual discrimination functions for the non-character condition for three subjects.

- EIMAS, P., & CORBIT, J. Selective adaptation of linguistic feature detectors. *Cognitive Psychology*, 1973, 4, 99-109.
- FRY, D., ABRAMSON, A., EIMAS, P., & LIBERMAN, A. The identification and discrimination of synthetic vowels. *Language and Speech*, 1962, 5, 171-189.
- KUKLINSKI, T. *Plasticity effects in the perception of hand-printed character*. MS thesis, Massachusetts Institute of Technology, 1975.
- LANE, H. The motor theory of speech perception: A critical review. *Psychological Review*, 1965, 72, 275-309.
- LIBERMAN, A., COOPER, F., HARRIS, K., & MACNEILAGE, P. A motor theory of speech perception. *Proceeding of the Speech Communication Seminar*, Stockholm, 1962.
- LIBERMAN, A., HARRIS, K., HOFFMAN, H., & GRIFFITH, B. The discrimination of speech sounds within and across phoneme boundary. *Journal of Experimental Psychology*, 1957, 54, 358-368.
- MATTINGLY, I., LIBERMAN, A., SYRDAL, A., & HALWES, T. Discrimination in speech and non-speech modes. *Cognitive Psychology*, 1971, 2, 131-157.
- NEISSER, U. *Cognitive psychology*. New York: Appleton, 1967.
- PISONI, D. *On the nature of categorical perception of speech sounds*. PhD thesis, University of Michigan, 1971.
- SHILLMAN, R. *Character recognition based on phenomenological attributes: Theory and method*. PhD thesis, Massachusetts Institute of Technology, 1974.
- STERNBERG, S. Two operations in character recognition: Some evidence from reaction time measurement. *Perception & Psychophysics*, 1967, 2, 45-53.
- STEVENS, K., LIBERMAN, A., STUDDERT-KENNEDY, M., & ÖHMAN, S. Cross-language study of vowel perception. *Language and Speech*, 1969, 12, 1-23.
- STUDDERT-KENNEDY, M., LIBERMAN, A., HARRIS, K., & COOPER, F. The motor theory of speech perception: A reply to Lane's critical review. *Psychological Review*, 1970, 77, 234-249.

NOTES

1. To predict discrimination from identification of a two-category continuum, the formula for an ABX discrimination task is $P_c = 1/2[1 + (P_1 - P_2)^2]$, where P_c is probability of a correct discrimination, P_1 is that of assigning Stimulus A to one of the two categories, and P_2 is that of assigning Stimulus B to that same category.
2. The operational definition of an abrupt categorical perception would be a matching between the obtained and predicted discrimination functions.

(Received for publication March 21, 1977;
revision accepted November 15, 1977.)