Ann. Bull. RILP, No. 14, 285-298(1980)

PERCEPTUAL ABILITIES OF APHASIC PATIENTS TO IDENTIFY FUNDAMENTAL FREQUENCY PATTERNS AND STROKE DIRECTIONS IN VERBAL AND NONVERBAL STIMULI

Itaru F. Tatsumi, Sumiko Sasanuma
and Hiroya Fujisaki

Comprehension of language takes place at different levels of information processing, some of which are known to be modality-independent and some modality-dependent. In aphasia, impairment of comprehension in both spoken and written language constitutes one of the cardinal features of symptomatology. Little is known, however, about the underlying mechanisms of the comprehension difficulties.

The present paper is an attempt to investigate the perceptual aspects of information processing in aphasics by means of administering a set of identification tests of fundamental frequency patterns and stroke directions in verbal and nonverbal stimuli.

METHOD

1. Subjects

The subjects of the experiment were five aphasics and five normals. All were speakers of the Tokyo dialect.

Two of the five aphasics had Broca's aphasia, and the remaining three exhibited Wernicke's aphasia. Table 1 summarizes the age, type and severity of aphasia, presence or absence of hemianopia* in the right visual field, and the post-onset duration (in months) for each patient. No patient had a mean hearing loss of more than 30 dB.

Table 1. Age, type and severity of aphasia, presence or absence of hemianopia, and post-onset duration (in months) for aphasic subjects.

PATIENT	AGE	TYPE	SEVERITY	HEMIANOPIA	POST-ONSET DURATION
B_1	48	Broca	mild	-----	39
B_2	56	Broca	mild or moderate	-----	13
W_1	52	Wernicke	moderate	right, mild	20
W_2	68	Wernicke	mild or moderate	right, mild	48
W_3	54	Wernicke	moderate	-----	9

* The presence of hemianopia in these patients was judged to have no adverse effect on their performance level for the visual tasks employed in the present experiments.

The age of the control subjects ranged from 18 to 34. All of them had normal hearing and visual acuity.

2. Stimuli

Two types of stimuli (one verbal and one nonverbal) in the auditory modality, and three (two verbal** and one nonverbal) in the visual modality were prepared.

(1) Verbal stimuli in the auditory modality

Nineteen two-mora words with gradual changes in their fundamental frequency (F_0) patterns from [awā] (bubble) through [āwa] (the name of a Japanese district) were synthesized.

The synthesis was performed by computer simulation of a terminal-analog speech synthesizer, in which the amplitude and frequency of buzz source, and the first three formant frequencies were controlled as shown in Fig. 1. The duration of each word was fixed at 300 msec, including a rise time and decay time of 20 msec each. The first and second formant frequencies (F_1 and F_2) were kept constant at 800 and 1150 Hz, respectively, during initial and final intervals of 80 msec, corresponding to the initial and final vowels of each. Between these two vowels F_1 and F_2 descend to 400 and 850 Hz and then ascend to 800 and 1150 Hz linearly with a transient interval of 70 msec, respectively. Frequencies of higher formants (F_3 to F_6) were placed at 2700, 3500, 4500 and 5500 Hz. The bandwidths of the first through the sixth formants were 60, 100, 120, 175, 280 and 500 Hz, respectively.

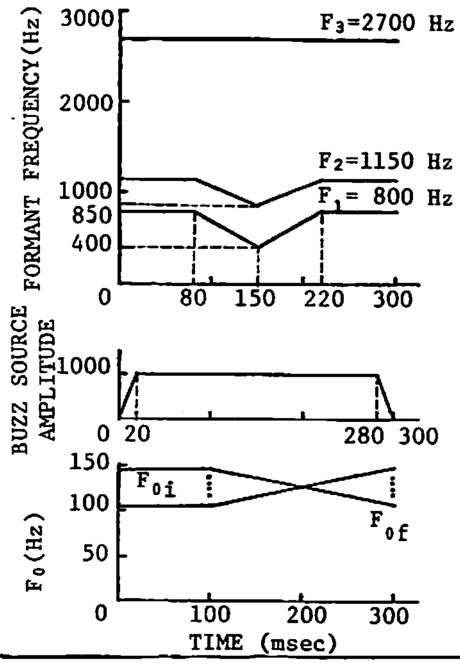


Fig. 1 Parameters used for synthesizing a series of verbal stimuli in the auditory modality ([awā] through [āwa]).

** In Japanese orthography two types of letter systems are used, i. e., the kanji or ideographic system and the kana or phonetic system. The kana system is further divided into two subsystems, katakana and hiragana. In the present experiments katakana was used because of the similarity of geometric features in katakana to those in kanji.

Fujisaki[1] has proposed a functional model of the processes that mediate between the linguistic information on word accents and its physical realization. According to his model, "Commands for voicing and accent are smoothed separately by the low-pass characteristics of their respective control mechanisms, each being approximated by a critically damped second order linear system, and their outputs are combined to control the fundamental frequency of glottal oscillations through a nonlinear mechanism." In the present experiments, however, more simplified F_0 patterns were used in order to preserve a greater similarity to the geometric features of visual stimuli. Namely, the initial value of the fundamental frequency (F_{0i}), which is constant over an interval of 100 msec, rises or falls linearly towards the specified final value (F_{0f}). The F_{0i} and F_{0f} of the 19 stimuli are listed in Table 2. The stimuli denoted by numbers 7, 9, 11 and 13 were excluded from the experimental stimulus set for the aphasics.

These stimuli, then, were recorded on an audio-tape through a D/A converter with a sampling rate of 12 kHz and an accuracy of 12 bits for the off-line experiments. Nineteen stimuli (or fifteen stimuli, for the aphasics) were recorded ten times each, and five dummy stimuli were added before and after these stimuli. A total of 200 (or 160) stimuli, therefore, were included in the test tape.

Table 2. Initial and final values of fundamental frequency (F_{0i} and F_{0f}, respectively) for synthesizing a series of verbal stimuli in the auditory modality ([awā] through [āwa]).

STIMULUS #	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
F_{0i} (Hz)	109	111	113	115	117	120	121	122	123	125	126	127	129	130	133	136	139	142	146
F_{0f} (Hz)	146	142	139	136	133	130	129	127	126	125	123	122	121	120	117	115	113	111	109

(2) Nonverbal stimuli in the auditory modality

A saw-tooth wave was employed as the nonverbal stimulus, since it has a spectral characteristic similar to that of the verbal stimulus, i.e., the spectral envelope of the nonverbal stimulus has a decay characteristic of -6 dB/oct, which corresponds to that of the buzz source (-12 dB/oct) superimposed by radiation characteristics (+6 dB/oct).

The synthesis and compilation of the stimuli were performed on a digital computer, as in the case of verbal stimuli.

(3) <u>Kanji</u> stimuli in the visual modality

A pattern of the <u>kanji</u> 木([ki], tree) shown in the left column of Fig. 2-(a) varies stepwise from 木 to 水([mizu], water), by means of increasing the angle of the right half of the horizontal stroke in 木, with the stroke length being kept constant. The angle was increased by a step of one degree from zero (horizontal) to 14 degrees, yielding a total of 15 patterns. Each pattern was constructed by 360 dots on a CRT-display. Since the time duration required to display the 360 dots was 30 msec and each pattern was displayed ten times, the exposure duration was 300 msec, which was the same as that of the stimuli in the auditory modality.

The synthesis and compilation of the stimulus patterns were carried out on the computer, and the patterns obtained were presented on the CRT-display by means of a computer control.

(4) Kana stimuli in the visual modality

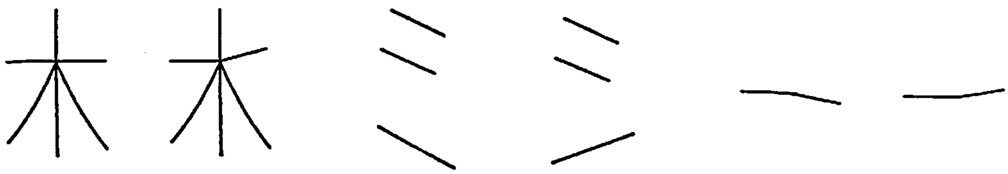
A series of 19 patterns were obtained by increasing the angle of the lowest stroke of ミ ([mi]), shown in the left column of Fig. 2-(b). The angle was varied by a step of 2. 6 degrees from -26. 0 to +20. 8 degrees (シ, [ʃi]).

The same procedures were used for the synthesis and compilation of the kana stimuli as those used for the kanji stimuli.

(5) Nonverbal stimuli in the visual modality

Nineteen patterns, varying stepwise from the left pattern to the right pattern shown in Fig. 2-(c), were generated in similar manner. The angle varied from -9 to +9 degrees by a step of one degree.

It should be noted that the geometric features involved in the three stimuli above ((3)-(5)) are slightly different from one another. In kanji and nonverbal stimuli, identification may be made with reference to the left horizontal stroke, while in kana stimuli there is no such stroke to make reference to. Identification of kana stimuli, therefore, would be more difficult than that of kanji and nonverbal stimuli. There is another difference between the features contained in kanji and nonverbal stimuli. Identification of the kanji stimuli may be made in terms of whether or not the direction of the right stroke is horizontal. Identification of the nonverbal stimuli, on the other hand, may be made with respect to a rise or fall (in the right direction) of the stroke. It is anticipated, therefore, that the identification of kanji stimuli will not be as stable as that of the nonverbal stimuli.



(a) (b) (c)

Fig. 2 Three types of stimuli used for the visual identification tests of kanji (a), kana (b) and nonverbal stimuli (c). Stimuli in the left and right columns have extreme stimulus values.

3. Procedures

(1) Experiments in the auditory modality

The subjects listened to the experimental stimuli through headphones in a soundproof room. The stimuli were presented at 60 dB above threshold and with an interval of 2. 5 or 3 sec between two stimuli.

In the case of the verbal stimuli, the task of the subjects was to judge which of the two test words [awā] and [āwa] was presented. Aphasics were asked to point one of two cards, one with [awā] written by kanji and kana and the other with [āwa], while the normals gave their responses by marking appropriate columns on a mark sense card.

For the nonverbal stimuli, the subjects were required to judge whether the F_0 had a rising or a falling pattern. Aphasics responded to the stimuli by pointing to either of two cards, one with a rising arrow and the other with a falling arrow, respectively, and normals by marking a mark sense card. For each stimulus item the aphasics made 30 to 40 judgments, while normals made 30 to 80 judgments.

(2) Experiments in the visual modality

Stimuli were presented on a frame of 7 cm x 6.5 cm on the CRT-display. The maximum horizontal and vertical dimensions of the stimuli were 3.2 and 4.6 cm, respectively. Aphasics were required to identify the stimuli and to respond to them by pointing to (1) one of two cards, one with a picture of a tree and the other with a picture of a glass of water, respectively, for the kanji task, (2) hiragana scripts corresponding to the katakana stimuli for the kana task, or (3) one of two cards, one with a rising arrow and the other with a falling arrow for the nonverbal task. Normals were instructed to mark an appropriate column on a mark sense card as their responses. Stimuli were separated by an interval of 2 or 2.5 sec.

Since the head position of the subjects was not fixed in order to avoid fatigue, the distance between the eyes and the CRT-display varied within a range of 60 and 90 cm.

4. Analysis of experimental results

If the judgment that a given stimulus belongs to one or the other of the two categories is perfectly precise, the probability of the judgment will jump from zero to one at the category boundary, when plotted against the stimulus value. However, the observed probability was not such a step function, but a continuous and monotonous function as shown in Fig. 3-(a), indicating that the judgment is disturbed by some psychological noises.

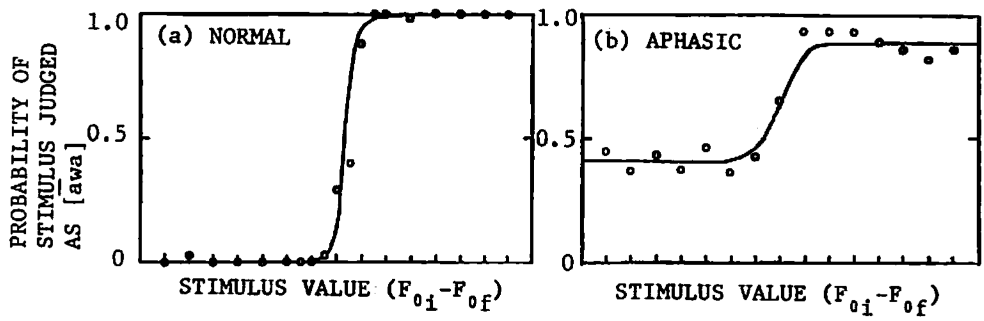


Fig. 3 Examples of the results of identification of verbal stimuli in the auditory modality for a normal subject (a) and an aphasic subject (b).

In the case of aphasic patients, however, this probability does not always asymptote to zero at one end of the stimulus value and/or to one at the other end (Fig. 3-(b)), suggesting that there exist other types of disturbances which cause errors, with a constant probability independent of the stimulus values. In the two-alternative-forced choice paradigms, these types of errors tend to occur when the subject fails to attend to the stimulus to which he must give a response and/or when he confuses one category symbol with the other.

These considerations led us to hypothesize the perceptual model illustrated in the upper panel of Fig. 4, in which the stimulus S_k is presented and the response is given by the category symbol '0' or '1'. The first block represents the attention process which is considered to be analogous to an electronic gate controlling a flow of stimulus information. The second block represents the perception/identification process which performs perception of the stimulus information sent from the attention process, and the categorical judgment based on the perceived stimulus information. As was remarked above, the categorical judgment can be disturbed by psychological noises. The last block is the response process which gives a response using the specified category symbols. The information loss in the response process occurs when the two category symbols are confused with each other.

Based on the perceptual model above, it is possible to calculate the probability of the response that a given stimulus belongs to one of two categories. Although the information loss in the attention and response processes can occur simultaneously, to simplify the formula it was hypothesized that the information loss takes place either in the attention process only or in the response process only.

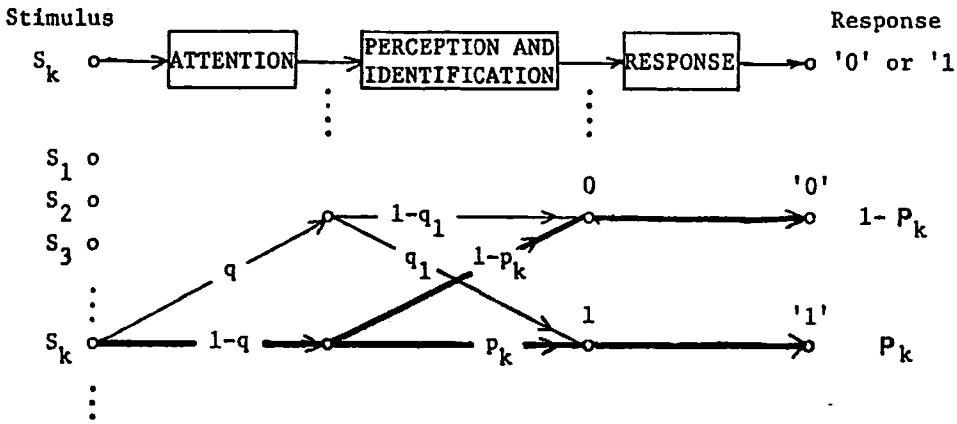


Fig. 4 Model for the processes responding to the stimulus S_k (upper panel) and flows of stimulus information (lower panel). In this case, no disturbance in the response process is hypothesized.

First, let us consider the information loss in the attention process. The flow of the stimulus information is shown in the lower panel of Fig. 4 by means of arrows. Assume that P_k is the probability of making the response that the stimulus S_k belongs to the category '1'. In addition, q is the probability of occurrence of attention failure, and q_1 the conditional probability that S_k to which the subject failed to attend is identified as '1'. Generally, q_1 is not equal to 0.5, since the subject tends to select one of two categories more frequently than the other. The symbol p_k represents the conditional probability that the normally attended stimulus is categorized as '1'. It is known that p_k is approximated by the cumulative normal distribution function[2] whose parameters (the mean μ and standard deviation σ) are the category boundary and an index of accuracy of identification, respectively[3]. Because no disturbance in the response process is hypothesized, P_k will be represented by equation ①:

$$P_k=(1-q)\cdot p_k+q\cdot q_1$$

$$=(1-2\cdot q_e)\cdot p_k+C_1 \qquad ①$$

$$q_e=q/2 \qquad ②$$

$$C_1=q\cdot q_1,$$

where q_e represents (the mean) error rate in the attention process.

Next we have to consider the case of information loss in the response process. As shown in the lower panel of Fig. 5, it is assumed that the disturbance in the response process may be represented by a binary noisy channel which is characterized by the probabilities q_{00}, q_{01}, q_{10} and q_{11}, where q_{ij} indicates the conditional probability that the response 'j' is selected when the input is 'i'. In general, q_{01} is not equal to q_{10} because of response bias. Then the probability P_k is given by equation ③:

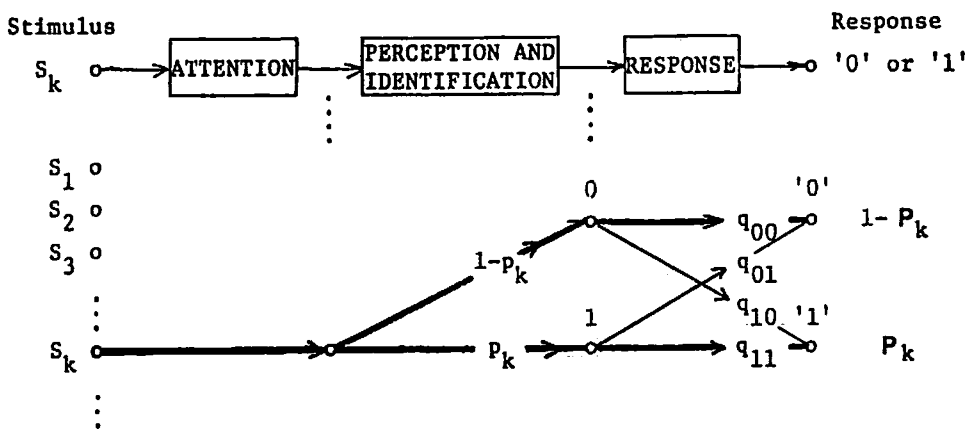


Fig. 5 Model for the processes responding to the stimulus S_k (upper panel) and flows of stimulus information (lower panel). In this case, it is hypothesized that all stimuli are attended normally.

$$P_k = p_k \cdot q_{11} + (1-p_k) \cdot q_{10}$$

$$= [1-(q_{01}+q_{10})] \cdot p_k + q_{01}$$

$$= (1-2 \cdot q_e) \cdot p_k + C_2 \qquad ③$$

$$q_e = (q_{01}+q_{10})/2 \qquad ④$$

$$C_2 = q_{01},$$

where q_e represents error rate in the response process.

It is of note here that equation ① is equivalent to equation ③. In fact, we will obtain essentially the same equation as ① (or ③) when disturbances are hypothesized to be presented in both the attention process and the response process.

Estimation of the theoretical value of P_k which gives the closest approximation to the measured data in the sense of the least-mean-squared error was based on the method of successive approximation. Since p_k is specified by its parameters μ and σ, a total of four parameters were obtained for each subject.

Figure 6 shows an example of a result of an aphasic patient for verbal stimuli in the auditory modality. The abscissa indicates the stimulus value or the difference of voice fundamental frequency $F_{0i}-F_{0f}$, and the ordinate the probability of making the response that the given stimulus is [āwa]. Circles represent the measured data, and the solid curve represents the closest approximation.

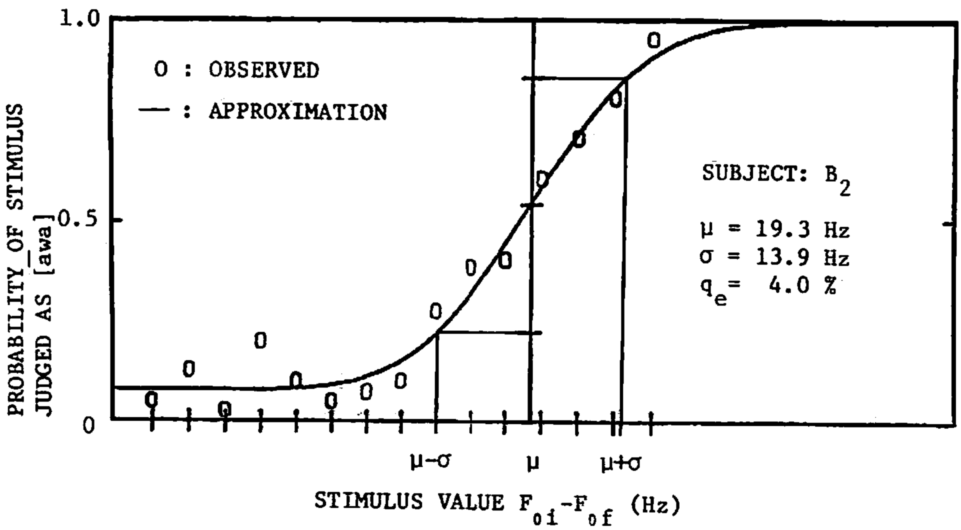


Fig. 6 Example of the results of auditory identification for verbal stimuli ([awā]/[āwa]) in an aphasic subject, and the approximate identification curve.

Because of the equivalence in equations ① and ③, it is impossible to determine, based on the value of q_e, which of the two processes (attention or response) is disturbed. Thus, analyses of the overall error patterns of responses are required to arrive at this answer. For example, occurrence of wrong responses in succession is the sign of information loss in the response process. Conversely, the random appearance of wrong responses may indicate information loss in the attention process.

RESULTS AND COMMENTS

In some aphasics, the probability of the response that a given stimulus belongs to one of the two categories remained almost constant regardless of the stimulus values, indicating that no approximation to the measured data can be obtained. We considered these patients to be incapable of identification. The test results obtained from these two types of patients will be described separately below.

1. Results for the aphasics capable of identification

(1) Category boundaries

The values of category boundaries in each of the identification tests are plotted in Fig. 7, in which μ's for verbal and nonverbal stimuli in the auditory modality, and μ's for kanji, kana and nonverbal stimuli in the visual modality are shown from left to right. B's and W's represent Broca's and Wernicke's aphasics, respectively, and circles normal subjects. The B's and W's located at the top of the graph represent those aphasics incapable of identification.

It can be seen that in the normals the category boundaries for verbal stimuli in the auditory modality are almost equal to those for their nonverbal counterparts, while in the aphasics they are not. In the aphasics, moreover, absolute values of category boundaries tend to be larger than those in the normals.

As for the visual modality, on the other hand, the category boundaries exhibited by the normals for the three types of test stimuli were quite different from each other, reflecting the different nature of each stimulus material. Similar results were obtained for the aphasics, with the exception of one Wernicke's aphasic (W_3) who showed considerably different values on kana and nonverbal tasks.

(2) Accuracy of identification

The results on the accuracy of identification for the five stimuli are shown in Fig. 8.

In the normals, σ's for the verbal and nonverbal stimuli in the auditory modality had approximately the same values. It should be hypothesized here that the perceptual process, represented in the second block in Fig. 4, consists of at least two processes connected in series, i. e., a nonspecific process common to the analysis of both verbal and nonverbal stimuli followed by a specific process for the analysis of the verbal stimuli only. Under this assumption, the finding of the equal values for the two σ's as well as two

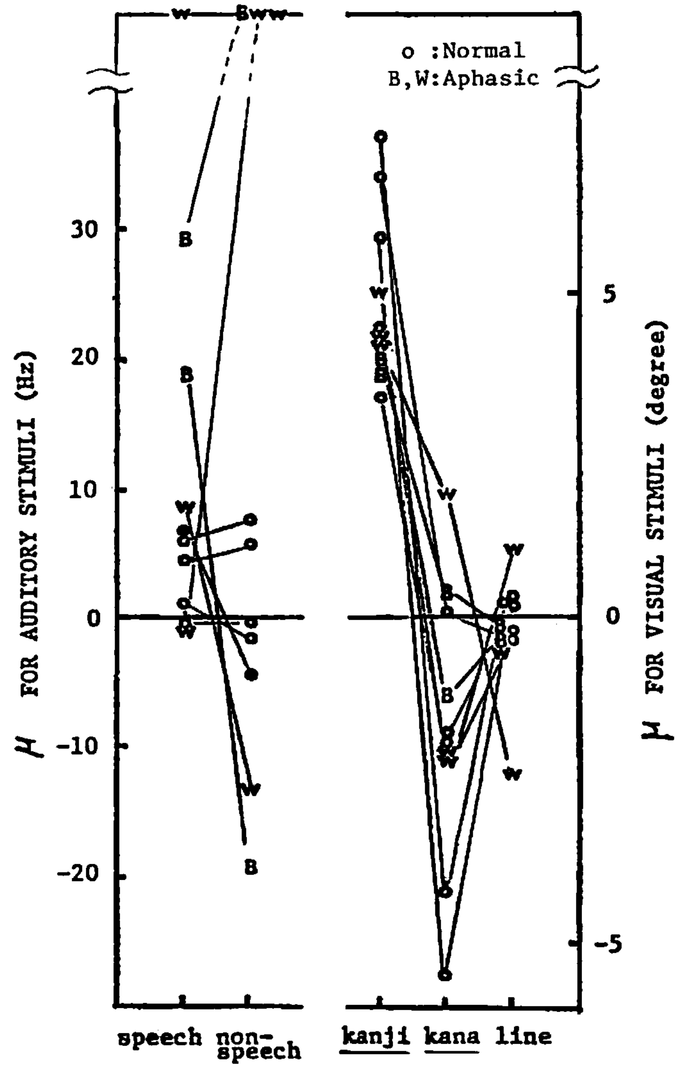


Fig. 7 Category boundary for each stimulus. Circles represent normal subjects, and B's and W's subjects with Broca's and Wernicke's aphasia, respectively.

μ's in the auditory modality may indicate that the values of these σ's and μ's are characterized mainly by the disturbances in the perceptual process common to the verbal and nonverbal stimuli.

The aphasics, however, showed remarkably different values for σ's between the verbal and nonverbal stimuli. In addition, their σ's for the verbal stimuli tended to be larger than those in normals. Similar tendencies were observed in the results of previous experiments[4] in which abilities of auditory identification of signal duration and visual identification of stroke length were investigated, suggesting that the locus of their impairment is in that process which is specific to the analysis of the verbal stimuli.

With visual stimuli, on the other hand, accuracies of identification exhibited by the normals for the three stimuli changed systematically with the geometric features characteristic to these visual stimuli. For all normals the accuracy of identification was highest for the nonverbal stimuli followed by kanji and kana stimuli, in that order. This can be interpreted to indicate that the stimuli with a horizontal reference stroke (i. e., the kanji and nonverbal stimuli) are identified more accurately than the stimuli without such reference (i. e., the kana stimuli). The difference in accuracy of identification among the two stimuli with a reference stroke also suggests that identification of a rise or fall of a stroke (as in the nonverbal stimuli) is more stable than that of a rise or horizontal of a stroke (as in the kanji stimuli).

Similarly, the accuracy of identification was highest for the nonverbal stimuli and lowest for the kana stimuli in all the aphasics. Furthermore, σ's for three visual stimuli were not beyond the normal range, with the exception of the patient with Wernicke's aphasia (W_3).

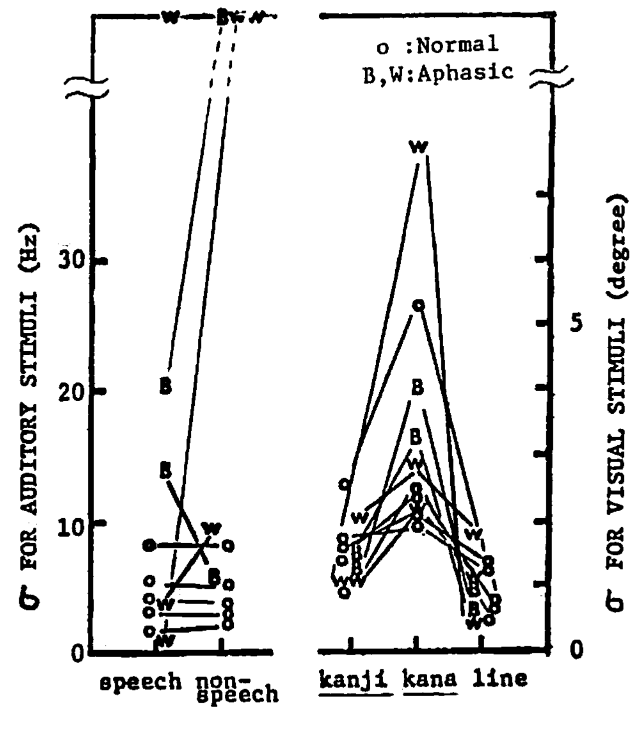


Fig. 8 Accuracy of identification for each stimulus.

(3) Error rate in the attention and/or response processes

The error rate in the attention and/or response processes q_e is shown in Fig. 9.

In the auditory modality, normals showed smaller q_e's for the verbal stimuli than those for the nonverbal stimuli. A similar trend was observed in the aphasics, though the values of q_e's were larger.

In the visual modality, on the other hand, q_e's for all three types of stimuli were distributed near zero percent for both the normals and the aphasics, with only one exceptional case, i.e., a Wernicke's aphasic who also showed abnormal μ's and σ for visual stimuli (W_3).

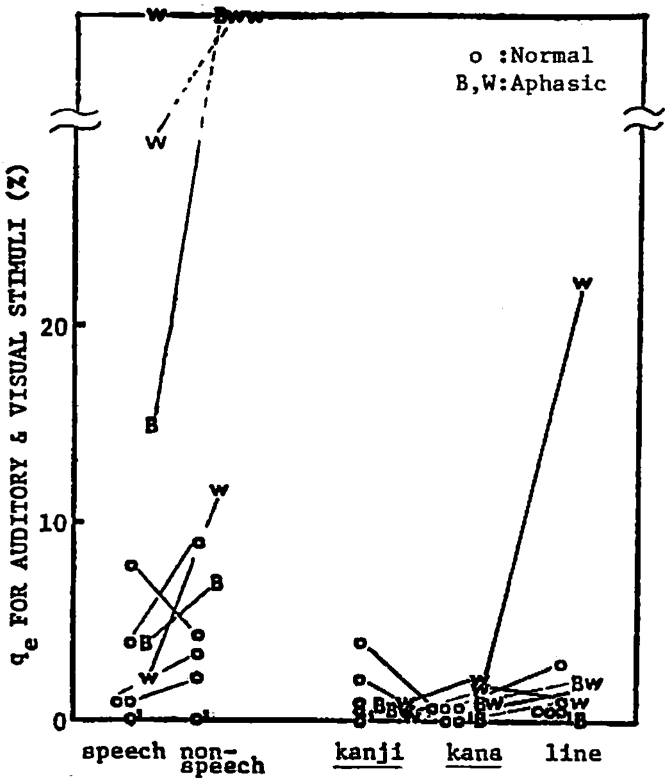


Fig. 9 Error rate in the attention and/or response processes for each stimulus.

As was stated earlier, for those patients with large q_e's, analyses of error patterns in a given response sequense would be necessary in order to determine which of the two types of errors, i.e., attention errors or response errors, is dominant. An example of such analysis with respect to the verbal stimuli in the auditory modality is presented in Table 3, in which the stimuli denoted by the numbers 1 through 15 and the responses to them given by one aphasic subject are recorded. The change of the stimuli from [awā] to [āwa] is indicated by an increase of the stimulus number, with the numbers in the upper row indicating the stimuli categorized as [āwa] while the numbers in the lower row indicating the stimuli categorized as [awā].

In this example the first 20 stimuli had a trend to be categorized correctly, i. e., the larger the number, the higher the probability of the stimulus having been categorized as [āwa]. However, a reversed response pattern was obtained in the next ten trials, indicating that it was the response process that is impaired and not the attention or the perception/identification processes. In the last ten trials the subject showed almost random response, indicating that the major loss of stimulus information took place in the attention process.

These trends were also found in the results of previous experiments on the auditory identification of signal duration and visual identification of stroke length[4].

Table 3. Sequence of responses to verbal stimuli in the auditory modality in an aphasic subject B_1. Stimuli represented by numbers 1 through 15 and responses to them are tabulated. Numbers in the upper and lower rows indicate the stimuli judged as [āwa] and [awā], respectively. The stimulus changes from typical [awā] to typical [āwa], as the stimulus number increases.

TRIAL #	CATEGORY	RESPONSES									
1-10	[āwa]		6		8	1		15	12		
	[awā]	5		3			10			7	5
11-20	[āwa]	14							11		13
	[awā]		9	4	1	3	2	6		8	
21-30	[āwa]		5		4	14		2		1	
	[awā]	8		12			15		11		7
31-40	[āwa]		9		10	6		15		5	
	[awā]	3		13			10		11		8
⋮						⋮					

2. Results for the aphasics incapable of identification

The instances of incapability of stimulus identification were restricted to the auditory modality, i. e., a total of four aphasic patients were found to be incapable of identifying verbal and/or nonverbal auditory stimuli. It is not clear what kind of impairment of information processing brings about the incapability of identification (i. e., attention error, perception/identification inaccuracy, or response error). It can be said, however, that a severe impairment in any one of these processes might be responsible for the incapability of identification.

Finally, the results of the present experiments showed no significant difference in performance between the two Wernicke's and three Broca's aphasics.

SUMMARY

In order to investigate the underlying mechanisms of language comprehension disorders in aphasia, identification tests were conducted using five aphasics and five normals. The tests consisted of auditory identification of fundamental frequency patterns for verbal and nonverbal stimuli, and visual identification of stroke directions of verbal and nonverbal stimuli.

The results indicated the deterioration of performances in the auditory tasks as against almost normal performances in the visual tasks in the aphasics. These observations were consistent with the results of our previous identification experiments with aphasics on signal duration in the auditory modality and stroke length in the visual modality, suggesting that in aphasia the primary impairment of perceptual abilities exists in the auditory modality. Based on analyses of the error patterns exhibited by the aphasics, it was suggested, further, that factors contributing to the observed impairment may involve not only disorders of perceptual processes per se, but also certain disturbances in the attentional and/or response processes in the auditory modality of aphasic patients.

REFERENCES

1) Fujisaki, H. (1977). Functional Models of Articulatory and Phonatory Dynamics, in Dynamic Aspects of Speech Production, M. Sawashima and F. S. Cooper (eds.), University of Tokyo Press, 347-368.
2) Thurstone, L. L. (1959). The Measurement of Values, University of Chicago Press.
3) Fujisaki, H., K. Nakamura and T. Imoto, (1973). Auditory Perception of Duration of Speech and Non-speech Stimuli, Ann. Bull. RILP, No. 7, 45-64.
4) Tatsumi, I. F., S. Sasanuma and H. Fujisaki (1978). Auditory and Visual Perception of Verbal and Nonverbal Stimuli in Aphasic Patients, Ann. Bull. RILP, No. 12, 157-166.