Neuropsychologia, 1977, Vol. 15, pp. 371 to 383. Pergamon Press. Printed in England.

THE PERCEPTION AND PRODUCTION OF VOICE-ONSET TIME IN APHASIA

SHEILA E. BLUMSTEIN*, WILLIAM E. COOPER†, EDGAR B. ZURIF‡
and ALFONSO CARAMAZZA§

(*Received* 17 *May* 1976)

Abstract—This study assessed aphasic and right brain-damaged non-aphasic patients' ability to label and discriminate a synthetic speech continuum differing in voice-onset time (VOT). We investigated these abilities in relation to type of aphasia and language comprehension facility, and explored the relation between perception and production of voice-onset time distinctions. Results of the perception tasks indicated that if a subject could not discriminate the stimuli, he could not reliably label them; however, a subject with a normal discrimination function might nonetheless be unable to label the stimuli reliably. These results were interpreted in relation to two levels of processing; one based upon the function of a set of property detectors, the other making use of these properties for linguistic processing. It is this latter level which seems to be selectively impaired in aphasia. Analysis of VOT production indicated that anterior aphasics have a deficit in the articulatory programming of speech sounds, whereas posterior patients have a deficit in selecting the appropriate phoneme. Finally, performance on the perception and production of VOT may be partially dissociated from language comprehension ability.

IN RECENT years, studies in the perception of speech have been concerned with determining the properties of the acoustic signal used by the listener in discriminating and identifying the phonetic dimensions of speech sounds. Results of these studies have shown that such phonetic dimensions as place of articulation and voicing can be defined quasi-independently by changes in specific parameters of the acoustic signal.

Granting that gaining the meaning of an utterance depends, at least in part, on an ability to perceive speech, the perception of acoustic parameters is critical for normal auditory language comprehension. Yet, with respect to aphasia, the issue remains whether the perception of the phonetic and ultimately the phonological form of a word can be disrupted by left brain-damage. That is, to what extent do language comprehension deficits reflect speech perception deficits?

Research on aphasics' ability to process auditory input has focused both on their capacity to make auditory discriminations of non-linguistic stimuli and to discriminate acoustic parameters critical to the speech signal. With respect to the former, aphasic patients have been found to be impaired in temporal order judgments [1] and auditory intensity discriminations [2]. With respect to discriminations based on parameters of the speech signal, aphasics show difficulty discriminating temporal cues [3], formant transition cues [4], and in categorically judging segmental duration cues [5]. However, it has not been determined whether aphasics who have impairments in discriminating such acoustic

---

*Department of Linguistics, Brown University, Providence, R.I. and Boston Aphasia Research Center.
†Massachusetts Institute of Technology, Cambridge, Massachusetts, U.S.A.
‡Boston Aphasia Center and Boston University School of Medicine, Boston, Massachusetts, U.S.A.
§Johns Hopkins University, Baltimore, Maryland, U.S.A.

parameters may nonetheless be able to categorize these stimuli linguistically or whether these impairments relate in any way to language comprehension ability.

In this study, we have attempted to answer these questions. We have focused upon the acoustic parameter of voice-onset time* (VOT), as it has been carefully investigated in normals both in terms of perception and production, and it is a parameter that is found in nearly all languages of the world and as a consequence seems to be a fundamental attribute in human speech perception.

The perception of voice-onset time as a cue to voiced–voiceless contrasts has been investigated in normals by synthesizing acoustic continua ranging systematically along this dimension. When asked to label the stimuli, subjects perceive the VOT continuum in terms of two discrete phonetic categories, i.e. either voiced or voiceless consonants. When asked to discriminate differences between stimuli taken from the same continuum, they can only discriminate reliably those stimuli which belong to different phonetic categories; those stimuli which are labelled as the same phonetic category cannot be discriminated reliably. This relation between the peaks of discrimination functions and category boundaries has been called categorical perception [8, 9].

A subsidiary question of concern to us was the extent to which speech perception abilities relate to speech production abilities. Specifically, do patients who manifest deficits in the perception of voice-onset time have similar deficits in producing these distinctions? Studies on VOT production in normals have revealed that English-speaking subjects produce two distinct non-overlapping ranges of VOT responses for voiced and voiceless stops sharing the same place of articulation, e.g. /d/–/t/ [7, 10]. In aphasia, patients commonly make phonological substitutions in their speech. These substitutions seem to be of two types, the first in which subjects make clear-cut substitutions of one phoneme for another, e.g. 'tack' →/dæk/, and the second in which the aphasic's production is a phonetic distortion of the target phoneme, e.g. [p^{h}Il] → [p$^-$Il] [11, 13]. Voice-onset time measurements can serve as an index of the two types of paraphasia described here. A phonemic substitution error would be reflected by a VOT value lying within the normal VOT range of the category opposite the target. A phonetic distortion error, on the other hand, would be reflected by a VOT value lying between the normal range of VOT for the voiced and voiceless categories. Our intention, therefore, was to determine whether such production deficits would coincide with the perception data in any meaningful way.

A final concern was to determine whether any deficits in the perception and production of the VOT continuum would reflect a general brain-damage effect or whether they could be attributable specifically to pathology in the language dominant hemisphere. Accordingly, we investigated both left brain-damaged aphasic subjects and non-aphasic right brain-damaged subjects.

METHOD

Perception experiment

Subjects. A total of 24 subjects participated in the perception experiment. These included: 4 non-brain-damaged control subjects taken from the medical and orthopedic services of the Boston V.A. Hospital;

---

*Voice-onset time is defined as the timing relation between release of the burst in a stop consonant and and onset of glottal pulsing. Simultaneous burst release and glottal pulsing is perceived as a voiced consonant, whereas a delay of glottal pulsing on the order of 30 or 35 msec in relation to the release of the burst is perceived as a voiceless consonant. Although voice-onset time generally refers to this timing relation, it in fact represents a constellation of acoustic attributes including the intensity and duration of the burst release, cutback and starting frequency of the first formant, and the presence or absence of friction noise upon consonantal release (cf. [6], [7]).

4 right brain-damaged non-aphasic patients; and 16 left brain-damaged aphasic patients. Of the aphasic patients, 8 had anterior brain-damage and 8 had posterior brain-damage. These patients represented several diagnostic categories—including among the anterior aphasics, 6 Broca's, 2 Mixed Anteriors, and among the posterior aphasics, 4 Wernicke's, 2 Conduction, and 2 Anomics. The diagnostic category to which each patient was assigned was based upon a composite aphasia examination including psychological, language, and neurological tests presented and discussed at the Aphasia Rounds of the Boston V.A. Hospital. The auditory language comprehension ability for all aphasic patients was assessed by performance on 4 sub-tests of the Boston Diagnostic Aphasia Battery [14]. All patients participating in this experiment were pre-screened for normal hearing in both ears for frequencies in the human speech range.

Stimuli. A total of 11 stimuli signalling the phonetic categories [da] and [ta] were synthesized by means of a computer-controlled parallel resonance synthesizer made available at Haskins Laboratories. Each stimulus consisted of 3 formant patterns in which the F1 transition started at 200 Hz, the F2 transition at 1800 Hz, and the F3 transition at 3200 Hz followed by a vowel with formants at frequencies of 800, 1250, and 3500 Hz respectively. The continuum ranged in VOT from −20 to +80 msec in 10 msec steps. Spectrograms of stimuli with VOT's of −20, 0, and +80 are shown in Fig. 1.

Two test tapes were constructed; an identification tape and a discrimination tape. The identification tape consisted of the random presentation of ten occurrences of each stimulus with an interstimulus interval of 4 sec. The discrimination tape consisted of the random presentation of stimulus pairs in which the members of each pair were either the same or differed from each other by a VOT of 20 msec. The members of each pair were separated by 1 sec of silence and the intertrial interval was 3·5 sec. The discrimination test consisted of 6 parts; 3 separate random orders, each presented twice. In each part there were a total of 47 discrimination trials—in 36 of these the members of the stimulus pair were different and in 11 they were the same. On the basis of results with normals, the critical discrimination pairs were between stimuli with VOT's of +10 and +30, +20 and +40, and +30 and +50. As a consequence, these were presented 8 times each. The remaining discrimination pairs containing different members occurred twice each. All discrimination pairs containing the same members were presented once each.

Procedure. Each subject was tested individually in a quiet room. The test-tapes were played on a high quality tape recorder and the signal was presented through Koss Pro 4-AA headphones at a comfortable listening level.

For the identification test, subjects were required to identify each stimulus presented auditorily by pointing to the appropriate card printed with *da* or *ta*. In order to insure that the subject could perform the task, he was first asked to point to the appropriate card upon hearing the examiner orally produce a good prototype of either /da/ or /ta/. When the examiner felt that the subject understood the task and could appropriately respond to the written letters, the subject was then presented with at least 6 test items taken from the actual test series which were unambiguously labelled by normals as /d/ or /t/. If the subject responded correctly in 4 of the 6 trials, he was included in the study.

In the discrimination test, subjects were required to determine whether the stimulus pairs were the same or different by pointing to a card marked 'yes' if the pair members were identical and 'no' if they were not. As in the identification test, subjects were pre-screened by the examiner to insure that they could read the cards appropriately and could understand the task demands. In this procedure, subjects were at first required to point to the appropriate card as the examiner orally produced good prototypes of the stimulus pairs and then to respond correctly to 4 out of 6 trials from the test series.

Production experiment

Subjects. A total of 12 of the 24 subjects who participated in the perception experiment were included in the production experiment. These included 3 control patients, 1 right brain-damaged patient, and 8 aphasic patients, including 3 Broca's, 1 Mixed Anterior, 2 Anomics, 1 Wernicke, and 1 Conduction. Four additional subjects participated only in the production experiment including 1 control non-brain-damaged subject, 1 Broca, 1 Anomic, and 1 Wernicke aphasic.

Stimuli. Six real-word monosyllabic stimuli were used to evaluate the subject's production of the VOT parameter in initial pre-stressed position. All stimuli included alveolar stops, three with the voiceless stop [t], 'Tom', 'tot', and 'top', and 3, the voiced stop [d], 'dock', 'Don', and 'dot'. In addition, 6 filler words were included beginning with bilabial and velar stop consonants. All words were presented in large capital letters on 3 × 5 cards for presentation to the subject. These words were presented in random order four times, totalling 48 items in the test. Only the alveolar pairs were analyzed for VOT.

RESULTS AND DISCUSSION

Perception of VOT

Labelling and discrimination functions were plotted for each subject. The subject's ability to perform the two tasks was assessed by (1) noting the locus of the VOT boundary and steepness of the identification function in the labelling task; (2) determining if there was a peak or several peaks in discrimination at the locus of the phonetic boundary.

As a measure of the subject's response bias in the discrimination task, we computed the % correct YES trials for each subject, i.e. the percent correct responses in which the discrimination pair contained identical members. The mean percent correct 'yes' trials for the control group was 96% with a standard deviation of 7·1%, and 98% for the right-brain damaged group with a SD of 3·7%. For the aphasic group, the mean percent correct 'yes' trials was 82% with a SD of 14·1%. Only one subject (a Broca's aphasic) with 52% correct 'yes' trials was eliminated on the basis that he was performing at chance level. The performance of the remaining subjects fell within two standard deviations above or below the mean.*

The performance of the control group replicated the results reported in the literature literature [8, 9]. The identification and discrimination functions of a typical subject can be seen in Fig. 2. Note that he labels stimuli ranging from a VOT of −20 to + 25 as a [d], and from +40 msec to +80 msec as a [t]. Moreover, as the figure shows, this subject can only discriminate stimuli with a VOT of +20 (labelled by him as a [d]) from +40 msec (labelled by him as a [t]), and stimuli with a VOT of +30 (labelled by him as a [t]), from +50 (labelled by him as a [d]). Comparisons between stimuli lying within the [d] range, e.g. VOT of 0 and +20 or within the [t] range, VOT of +40 and +60, could not be reliably discriminated.

Similar results were obtained for all subjects in the right brain-damaged group. The normal performance of the right brain-damaged group suggests that any deficits on these tasks manifested by the left brain-damaged aphasic patients can be attributed to the effects of damage specific to the dominant language hemisphere.

The performance of the aphasic patients on these tasks was characterized by three distinct patterns. These results are summarized in Fig. 3. In the first pattern, patients performed normally on both identification and discrimination tasks. Eight aphasic patients comprised this group including 3 Broca's, 1 Mixed Anterior, 1 Wernicke, 1 Conduction, and 2 Anomics. The performance of one of these subjects can be seen in Fig. 3a and 3b. In the second group, consisting of 1 Mixed Anterior, 1 Broca, and 1 Conduction aphasic, patients were unable to either label or discriminate the stimuli reliably. The performance of such a patient is represented in Fig. 3c and 3d. Note that, although this patient labelled consistently stimuli with a VOT of −10, 0, and +10 msec as a /d/, and stimuli with a VOT of +20, +60, and +70 as a /t/, he nonetheless did not consistently label the stimuli comprising the continuum, nor was there a sharp category boundary between /d/ and /t/ normally found at about 30 msec. Although this pattern of results differs dramatically from that found in our control patients, we cannot be certain that the failure of these patients is attributable to left brain-damage, as it is sometimes the case that normals are also unable to respond to

---

*In addition to measuring the overall accuracy of each patient's performance, a correction for guessing based on signal detection theory analysis was also applied to the raw discrimination data. D′ for each discrimination pairing was computed and converted to percent correct using 2-alternative forced-choice tables [15]. On the whole, the pattern of results obtained for the aphasic patients based on this analysis remained the same. However, for those Ss whose discrimination accuracy was 100% or close to 100% for paired stimuli that were physically identical, the application of the signal detection analysis produced a heightened measure of the S's discrimination capacity. For example, on the basis of analysis of the raw data, a normal subject with 100% accuracy in 'same' judgments showed discrimination peaks between the stimulus pairs +20 and +40, and +30 and +50. After the application of the signal detection analysis, the same subject showed peaks of discrimination at +10 and +30, +20 and +40, +30 and +50, and +60 and +80. The peaks of discrimination normally found only at the phoneme boundary now occurred at several other points along the continuum. As a consequence, the results of the discrimination task are based upon the original (raw) discrimination data.

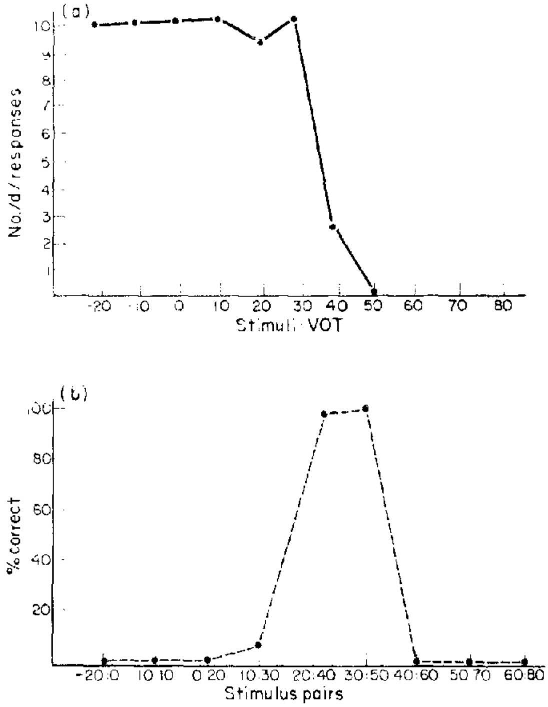

FIG. 2. The identification function (a) and discrimination function (b) for a VOT continuum for a normal control subject. In (a) the abscissa represents the stimulus values of the continuum ranging from −20 to +80 VOT. The ordinate represents the total number of d-responses (maximum is 10). In (b), the abscissa represents the 'different' stimulus pairs; each pair was distinguished by a VOT of 20 msec. The ordinate represents the percent correct different responses.

synthetic stimuli as speech-like and perform randomly both in discrimination and labelling tasks. The third group, including 3 Wernicke and 1 Broca aphasics, performed normally on the discrimination task, but were unable to identify the test stimuli reliably. The identification and discrimination functions of an aphasic of this type are presented in Fig. 3e and 3f. Note that although he tended to label more stimuli as a /d/ at short VOT values and stimuli as a /t/ at longer VOT values, his identification function did not indicate a sharp boundary function between the /d/ and /t/ categories. Nevertheless, his discrimination function indicated peaks between stimuli with a VOT of +20 and +40 and +30 and +50 corresponding to the phoneme boundary typically found in normal English-speaking subjects.

In summary, the performance of the aphasic patients has demonstrated an interesting relation between the identification and discrimination of the VOT continuum. Namely, if a subject is able to identify reliably the stimuli along this continuum, he is also able to make accurate discrimination judgments at the phoneme boundary. Discrimination ability was reflected by peaks at the phoneme boundaries even for those subjects who were unable to label the stimuli reliably. No cases were found in which a patient was unable to discriminate but was able to label the stimuli.

Although it is clear that some deficits in the perception of VOT are found in nearly all

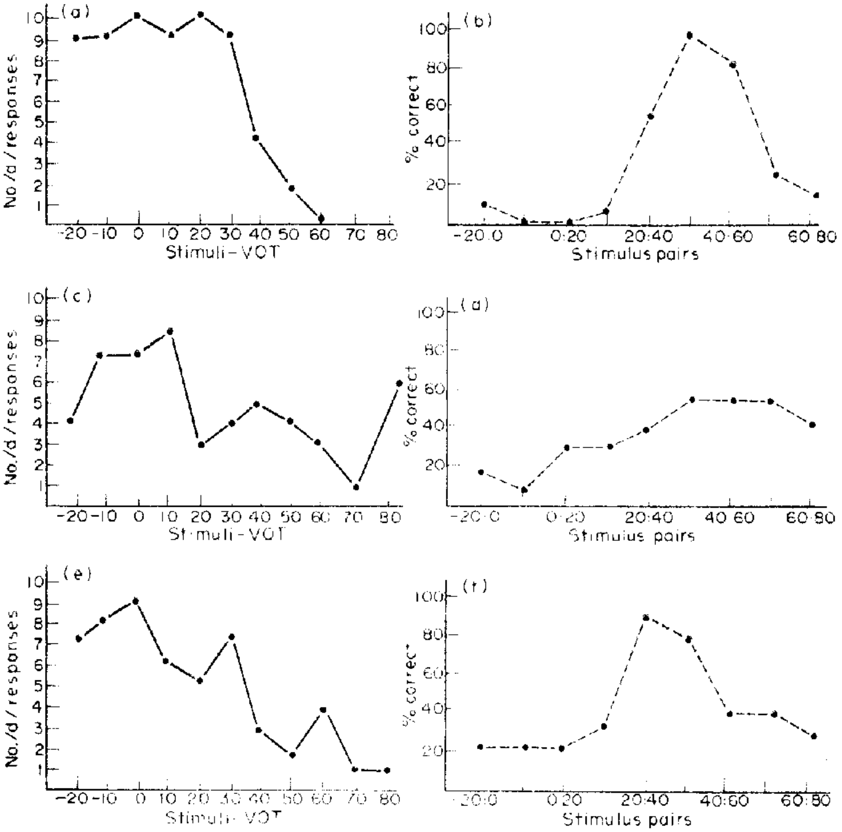

FIG. 3. The three patterns of labelling and discrimination of VOT shown by the aphasic patients. Each pattern (a,b), (c,d), (e,f) is represented by the performance of one aphasic patient. See also legend for Fig. 2.

aphasic groups represented, performance was analyzed in relation to clinical type in order to determine whether there was any overall performance difference across groups. Table 1 summarizes the discrimination and labelling performance of each subject in relation to type of aphasia. The only consistent pattern of deficit among the groups is found in the Wernicke aphasics. Here, three of the four patients were able to perform the discrimination task but were unable to label the stimuli consistently.

Finally, in order to determine to what extent language comprehension facility was correlated with performance on the perception tasks, all aphasics were given the auditory comprehension sub-test of the Boston Diagnostic Aphasia Battery [14]. The mean performance score, converted to a z-score ratio, was computed for each subject and is plotted in Fig. 4 in relation to performance on the perception tasks. As the figure shows, there seems to be little relation between the patient's auditory language comprehension and his ability to perform the VOT perception tasks. Although most of the subjects who had good comprehension were able to perform both the discrimination and labelling tasks, there were three subjects with high comprehension scores (a z-score better than 0) who demonstrated deficits in either labelling or both labelling and discrimination. Perhaps more revealing is the performance of those subjects who had poor comprehension. Ability to perform the per-

Table 1. Performance for each subject on discrimination and labelling of a VOT continuum in relation to type of aphasia. A plus indicates normal performance, a minus impaired performance. Normal performance for discrimination is defined by peaks of discrimination at the phoneme boundaries, and for labelling as consistent (70%) identification of stimuli with a steep identification slope between the two category values

Type of Aphasia	Discrimination	Labelling
Broca		
1	+	+
2	+	+
3	+	+
4	+	−
5	−	−
Mixed Anterior		
1	+	+
2	−	−
Conduction		
1	+	+
2	−	−
Wernicke		
1	+	+
2	+	−
3	+	−
4	+	−
Anomic		
1	+	+
2	+	+

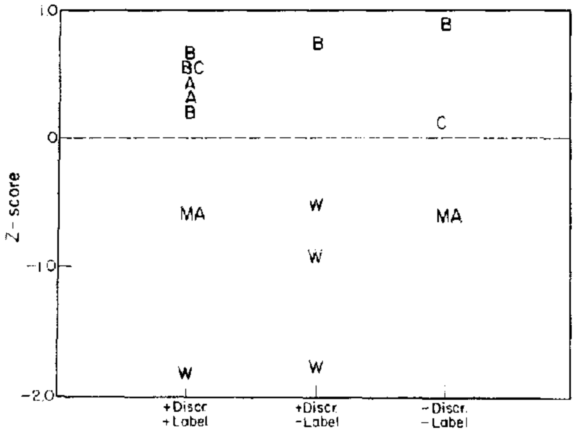

FIG. 4. Performance of aphasic subjects on the discrimination and labelling of a VOT continuum in relation to level of auditory comprehension. The aphasic subjects represented include Broca's (B), Mixed Anterior (MA), Conduction (C), Anomic (A), and Wernicke (W) aphasics. The abscissa represents performance on the discrimination and labelling tasks. A plus (+) indicates normal performance, a minus (−) impaired performance (see also legend for Table 1). The ordinate represents the aphasic's performance on the auditory comprehension subtest of the Goodglass and Kaplan aphasia test battery converted to z-scores (see text).

ception tasks ranges from normal performance on labelling and discrimination to an inability to perform either task. In fact, the Wernicke aphasic with the most severe comprehension deficit performed normally on the perception of VOT.

The production of VOT

Each utterance produced by the patient was measured for VOT oscillographically using a computer-controlled program developed by A. W. F. Huggins at the M.I.T. Research Laboratory of Electronics. All alveolar stop productions produced by the patient, including repeated attempts at the target word, were measured for VOT. Note was also made of the particular target word attempted. In this way, we could compare the distribution of VOT responses for both voiced and voiceless targets.

The frequency distribution of the VOT responses was plotted individually for each subject. Figure 5 shows the distribution of VOT productions for a normal control. Like the findings on normal subjects reported by LISKER and ABRAMSON [7], the subject's voiced and voiceless productions are distributed into two areas of VOT—a short VOT ranging between −10 and +27 msec, representing the voiced targets, and a long VOT, ranging between +55 and +76 msec, representing the voiceless targets. Note also that there is no overlap between the two phonetic categories. Similar results were found for the 3 Anomic aphasics and the one right brain-damaged subject.

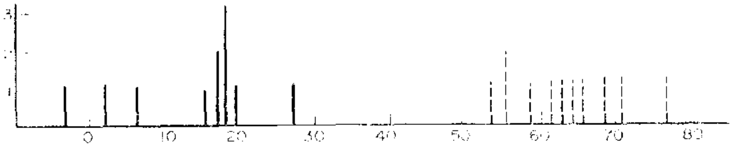

FIG. 5. Distribution of VOT production responses for alveolar stop consonants in initial pre-stressed position for a normal control. The abscissa represents the VOT values for each stimulus, and the ordinate the total number of responses. The solid lines indicate the voiced-target /d/, and dashed lines the voiceless-target /t/.

A comparison of the VOT production distributions for one Wernicke and one Broca aphasic can be seen in Fig. 6. Note that the Wernicke aphasic produced several paraphasic errors; that is, several of the target words beginning with the voiced stop /d/ were produced with VOT values clearly in the voiceless range. These words were perceptually clear-cut

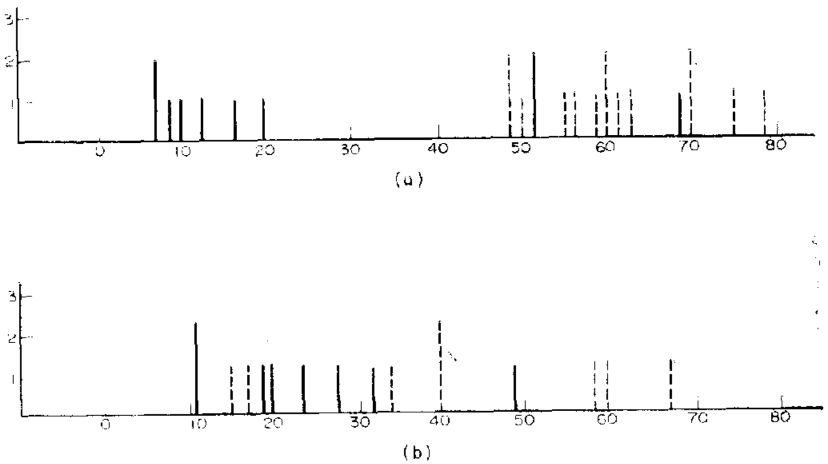

FIG. 6. Distribution of VOT production responses for alveolar stop consonants in initial pre-stressed position for a Wernicke (a) and a Broca (b) aphasic. See also legend for Fig. 5.

substitutions, e.g. [dan] ⟶ [t^han]. Nevertheless, the VOT productions of this patient indicated no overlap between the VOT values for the two phonetic categories [d] and [t]. Thus, although he produced phonological errors, they fell within the range of either the voiced or voiceless phonetic categories, not midway between.

This distribution is in sharp contrast to the one represented by the Broca's aphasic. Two important characteristics of this patient's performance are noteworthy. First, there do not seem to be two distinct VOT regions. Rather, both voiced and voiceless productions overlap in the VOT range between about +25 and +40, the range where few VOT responses were found in the normals. Second, some target values for voiceless stops were produced which clearly fell within the range of voiced stops, i.e. target /t/ with a VOT measurement of +50. These productions then seemed to be clear-cut phonemic substitutions, in the sense that they fell within the normal range of either the voiced or voiceless VOT regions. The occurrence of these so-called phonemic substitutions were found much less frequently in the production of the anterior aphasics than those errors in which the VOT values spanned the area between the normal voiced and voiceless categories. The latter type of error in which both voiced and voiceless targets occurred in the VOT range typically found between the voiced and voiceless phonetic categories, was found in all but one of the anterior aphasics tested.*

A comparison of perception and production of VOT

Performance on both VOT production and perception tasks was compared for those aphasics who participated in both parts of the study to determine the degree to which the ability to produce normal distributions of VOT correlated with the ability to perceive distinctions along this dimension. Results of this comparison can be seen in Table 2. It is quite clear, at least for the anterior aphasics (Broca and Mixed Anterior), that the ability to perceive the VOT continuum relates in no way to the ability to produce voiced and voiceless stops. Thus, the anterior aphasics maintain the ability to perceive this distinction, but make both phonemic as well as phonetic substitutions. For the two posterior aphasics, there are insufficient data to make any statements concerning the relation between production and perception of VOT.

GENERAL DISCUSSION

The results of this study provide some insights into the nature of perceptual processing in the brain-damaged subject as well as some idea of the nature of perceptual processes in the normally functioning individual. The aphasic's ability to perceive the VOT continuum appears to be based upon a particular relation between the discrimination and labelling tasks. Thus, subjects could label the test stimuli only if they could also discriminate them, and furthermore, discrimination ability reflected peaks at the phoneme boundaries even for those subjects who were unable to label the stimuli reliably. The aphasics' impairment then seems to reflect an inability to maintain a stable configuration or category label despite the fact that he retains the ability to discriminate phonemically distinct categories. These results suggest that discrimination ability as tested here underlies phoneme

---

*The VOT responses of the patient whose pattern did not include VOT responses between the voiced and voiceless categories nevertheless produced a pathological distribution of responses. All of this subject's VOT responses for both voiced and voiceless targets were in the range between 56 and 97 msec, i.e. a long VOT lag. This may reflect the subject's difficulty in initiating speech.

Table 2. Performance for each aphasic subject on the production of a VOT continuum in relation to the perception of a VOT continuum. Production performance is divided into 3 categories: $\begin{bmatrix}+\text{Phonetic}\\+\text{Phonemic}\end{bmatrix}$ representing the presence of both phonetically and phonemically based errors, $\begin{bmatrix}-\text{Phonetic}\\+\text{Phonemic}\end{bmatrix}$ representing the presence only of phonemic substitutions, and $\begin{bmatrix}-\text{Phonetic}\\-\text{Phonemic}\end{bmatrix}$ representing normal performance, i.e. the absence of both phonetic and phonemic errors. Perception of VOT is divided into 3 categories: $\begin{bmatrix}+\text{D}\\+\text{L}\end{bmatrix}$ indicating normal performance, i.e. normal discrimination and labelling functions; $\begin{bmatrix}+\text{D}\\-\text{L}\end{bmatrix}$ signifies normal discrimination and impaired labelling, and $\begin{bmatrix}-\text{D}\\-\text{L}\end{bmatrix}$ indicates impaired performance on both discrimination and labelling tasks (see also Table 1 and Text)

	Production of VOT		
Perception of VOT	+Phonetic +Phonemic	−Phonetic +Phonemic	−Phonetic −Phonemic
+D +L	2 Broca 1 Mixed Anterior		2 Anomics
+D −L		1 Broca 1 Wernicke	
−D −L		1 Conduction	

perception, and that the use of linguistic categories as discrete phonemic entities is based upon the sensitivity of auditory property detectors to a limited range of acoustic stimuli.* It may well be that it is this limited set of properties to which the perceptual system is sensitive that defines the linguistically significant categories used in natural language. Recent research from infant speech perception supports such an interpretation [16].

For the aphasic patient, discrimination ability need only reflect the capacity to perceive differences between acoustic parameters. The ability to use these sounds in a linguistically relevant way seems to be the basis for the perceptual difficulties of many of these patients.

The perception of speech sounds then seems to be based upon the integration of two distinct levels of processing. The first and more basic is a pre-linguistic level in which selective differences between auditory stimuli may be discriminated. The second level may be properly termed linguistic as the acoustic categories derived from the pre-linguistic level are used to distinguish functionally different speech sounds within the linguistic system. It is only when these two levels are integrated that the fundamental relation between sound and meaning requisite for the comprehension of linguistic labels and ultimately the linguistic system may be fully realized.

The relation between sound and meaning may well be the basis for the comprehension deficit of the Wernicke's aphasics. Although it is difficult to attribute a selective deficit in the perception of VOT to any one clinical group on the basis of the findings in this study, the Wernicke aphasics did show the most consistent pattern of deficit; namely, an inability to reliably label the stimuli, despite the ability to discriminate them. These results support recent findings that suggest that the Wernicke's aphasics do not have a selective

---

*It should be noted that this argument runs counter to the original account of categorical perception [8] in which a peak of discriminability for normals is ascribed to the presence of a well-defined category boundary. Our results, showing the ability of aphasics to discriminate without the ability to identify, suggest instead that a peak in discrimination is not due to a constraint on the listener's ability to assign category labels. In our view, categorical perception is attributed to a discrimination constraint on identification.

deficit in phonemic hearing ([17] but cf. [18]), but rather seem to be unable to assign a stable category label to a speech stimulus, i.e. to use phonological information in a linguistically relevant way [19,20].

Thus, an inability to maintain the relation between sound and meaning would be reflected by a deficit in pointing to or naming appropriately concrete objects or assigning a linguistic label to speech sounds, despite the fact that the ability to recognize differences between names and sounds may be intact.

Nevertheless, it does seem surprising that ability to perform the labelling and discrimination tasks bears little relation to auditory language comprehension. If it is the case as suggested here that speech perception is at the very least a necessary pre-requisite to auditory language comprehension, then a failure to perform the tasks in this study should correlate with low auditory language comprehension scores. It is certainly easy to reconcile a patient's ability to discriminate and label VOT despite poor language comprehension. This suggests that the disorder is at some other level of linguistic processing. What is more difficult to interpret is a patient's inability to perform the perceptual tasks in the face of good language comprehension. Several factors may be operative. First, the synthetic stimuli contained a minimal number of acoustic cues normally present in natural speech. Thus, the subject is presented with a skeletal signal stripped of the redundancies normally found in the acoustic signal. Second, language is typically spoken in the context of meaningful utterances such that the listener is provided with situational cues of both a linguistic and non-linguistic nature. In effect, the listener may use semantic data to resolve phonetic doubt. Finally, it is not clear to what extent a subject's awareness of segments tested in the labelling task is a necessary condition for normal auditory language comprehension. The comprehension of words may be based on an ability to apprehend unique auditory patterns rather than to recognize the identity of individual segments.

The analysis of VOT production in aphasia has provided an operational measure of the distinction between phonetic and phonemic disintegration, with phonetic disintegration found in the anterior aphasic. The overlap between the voiced and voiceless categories shown by the anterior aphasics suggests that they have a deficit in the articulatory programming of speech sounds. Thus, the timing relation between the release of the stop consonant and the beginning of glottal pulsing seems to be one manifestation of this disorder. The analysis of other articulatory parameters would presumably reveal further degrees of impairment (cf. [12, 21]). In addition to the presence of so-called phonetic errors in VOT, the measurements of VOT productions in the anterior aphasics also indicated that they may also target incorrectly a particular speech sound, resulting in a clear-cut shift in phonetic category, at least along the VOT dimension. This type of error occurs much less frequently and seems to be on the whole less characteristic of their articulatory pattern. In contrast, the VOT measurements of the posterior aphasics revealed only the presence of phonemic substitutions, i.e. clear-cut shifts in phonetic category from the target phoneme. It would seem then that these errors reflect a deficit in selecting the appropriate phoneme or underlying phonological form, and subsequently, programming correctly the articulatory commands for the substituted phoneme.

Acknowledgements—This research was supported in part by USPHS Grants NS07615 to Clark University, NS06209 and NS11408 to Boston University School of Medicine. Many thanks to A. W. F. HUGGINS for his technical assistance and to K. N. STEVENS for his critical comments and suggestions on an earlier draft of this paper.

REFERENCES

1. EFRON, R. Temporal perception, aphasia, and *déjà vu. Brain* **86,** 403–424, 1963.
2. SWISHER, L. P. Auditory intensity discrimination in patients with temporal-lobe damage. *Cortex* **2,** 179–193, 1967.
3. CARPENTER, R. L. and RUTHERFORD, D. R. Acoustic cue discrimination in adult aphasia. *J. Speech Hear. Res.* **16,** 534–544, 1973.
4. TALLAL, P. and PIERCY, M. Developmental aphasia: the perception of brief vowels and extended consonants. *Neuropsychologia* **13,** 69–74, 1975.
5. SASANUMA, S., TATSUMI, I. F., KIRITANI, S. and FUJISAKI, H. Auditory perception of signal duration in aphasic patients. *Research Institute of Logopedics and Phoniatrics, Annual Bulletin* **7,** 65–72, 1973.
6. STEVENS, K. N. and KLATT, D. H. Role of formant transitions in the voiced–voiceless distinction for stops. *J. acoust. Soc. Am.* **55,** 653–659, 1974.
7. LISKER, L. and ABRAMSON, A. A cross-language study of voicing in initial stops: acoustical measurements. *Word* **20,** 384–422, 1964.
8. LIBERMAN, A. M., HARRIS, K. S., HOFFMAN, H. S. and GRIFFITH, B. C. The discrimination of speech sounds within and across phoneme boundaries. *J. exp. Psychol.* **54,** 358–368, 1957.
9. LIBERMAN, A. M., COOPER, F. S., SHANKWEILER, D. P. and STUDDERT-KENNEDY, M. Perception of the speech code. *Psychol. Rev.* **74,** 431–461, 1967.
10. LISKER, L. and ABRAMSON, A. S. Some effects of context on voice onset time in English stops. *Language and Speech* **10,** 1–28, 1967.
11. BLUMSTEIN, S. E. *A Phonological Investigation of Aphasic Speech.* Mouton, The Hague, 1973.
12. ALAJOUANINE, TH., OMBREDANE, A. and DURAND, M. *Le Syndrome de la Désintégration Phonétique dans l'Aphasie.* Masson & Cie, Paris, 1939.
13. LECOURS, A. R. and LHERMITTE, F. Phonemic paraphasias: linguistic structures and tentative hypotheses. *Cortex* **5,** 193–228, 1968.
14. GOODGLASS, H. and KAPLAN, E. *The Assessment of Aphasia and Related Disorders.* Lea & Febiger, Philadelphia, 1972.
15. SWETS, S. A. (Editor). *Signal Detection and Recognition by Human Observers: Contemporary Readings.* Wiley, New York, 1964.
16. EIMAS, P. D., SIQUELAND, E. R., JUSCZYK, P. and VIGORITO, J. Speech perception in infants. *Science* **171,** 303–306, 1971.
17. BLUMSTEIN, S. E., BAKER, E. and GOODGLASS, H. Phonological factors in auditory comprehension in aphasia. *Neuropsychologia* **15,** 19–30, 1977.
18. LURIA, A. R. *Traumatic Aphasia.* Mouton, The Hague, 1970.
19. HARRIS, C. M. Phonemic errors made by aphasic subjects in the identification of monosyllabic words. Unpublished doctoral dissertation, University of Minnesota, 1970.
20. SCHUELL, H., JENKINS, J. J. and JIMENEZ-PABON, E. *Aphasia in Adults.* Harper & Row, New York, 1964.
21. LEHISTE, I. *Some Acoustic Characteristics of Dysarthria.* Biblioteca Phonetica, Switzerland, 1968.

Résumé :

On a examiné dans cette étude la capacité des malades aphasiques et des malades non aphasiques avec lésions droites à étiqueter verbalement et discriminer un continuum de parole synthétique différant selon le moment du début des vibrations vocales (voice-onset time : VOT). Nous avons examiné ces capacités dans leur relation avec le type d'aphasie et la compréhension du langage et nous avons exploré la relation entre la perception et la production des distinctions de VOT. Les résultats des tâches de perception indiquaient que si un sujet ne pouvait pas discriminer des stimulus, il ne pouvait pas de manière fiable les étiqueter; mais un sujet avec une fonction de discrimination normale pouvait néanmoins être incapable de les étiqueter. On interprète ces résultats en admettant 2 niveaux de traitement : l'un fondé sur la fonction des séries de détecteurs de propriété, l'autre utilisant ces propriétés pour un traitement linguistique. C'est ce dernier niveau qui semble être sélectivement troublé dans l'aphasie. L'analyse de la production de VOT indique que les aphasiques antérieurs ont un déficit de la programmation articulatoire des sons verbaux tandis que les malades avec lésions postérieures ont un déficit de la sélection du phonème approprié. Enfin, les performances de perception et de production de VOT peuvent être partiellement dissociées de l'incapacité de compréhension du langage.

Deutschsprachige Zusammenfassung:

In dieser Untersuchung wurde die Fähigkeit von aphasischen und nichtaphasischen rechtshirnig geschädigten Kranken erfaßt, ein Kontinuum synthetischer Sprache zu kennzeichnen und zu diskriminieren. Das sprachliche Angebot wies Unterschiede bezüglich der Stimmeinsatzzeit (VOT) auf. Wir untersuchten die Fähigkeiten in Bezug auf Aphasietyp und vorhandenes Sprachverständnis und prüften die Beziehung zwischen Perception und Produktion von Stimmeinsatzzeitunterschieden. Die Ergebnisse der Perceptionsaufgaben zeigten, daß, wenn ein Patient die Reize nicht zu diskriminieren vermochte, er sie auch nicht zuverlässig kennzeichnen konnte; dagegen mochte ein Proband mit normaler Diskriminationsleistung nichtsdestoweniger unfähig sein, die Reize zuverlässig zu kennzeichnen. Diese Ergebnisse wurden in Bezug auf 2 Ebenen des Prozessing interpretiert. Eine betraf die Funktion eines Analysatorensatzes für Eigenschaften, die andere benützte diese Eigenschaften für das weitere sprachliche Prozessing. Es ist diese zweite Ebene, die bei der Aphasie besonders gestört zu sein scheint. Die Analyse der VOT-Produktion weist darauf hin, daß die präcentralen Aphasiker ein Defizit im artikulatorischen Programmieren von Sprachlauten haben, während die postzentralen Patienten ein Defizit im Selektieren der relevanten Phonomene aufweisen. Schließlich kann die Leistung im der Perception und Produktion der VOT teilweise unabhängig von der Sprachverständnisfähigkeit sein.