

Parallel processing of auditory and phonetic information in speech discrimination*

CHARLES C. WOOD†

Neuropsychology Laboratory, VA Hospital, West Haven, Connecticut 06516
and Yale University, New Haven, Connecticut 06510

Recent experiments using a variety of techniques have suggested that speech perception involves separate auditory and phonetic levels of processing. Two models of auditory and phonetic processing appear to be consistent with existing data: (a) a strict *serial* model in which auditory information would be processed at one level, followed by the processing of phonetic information at a subsequent level; and (b) a *parallel* model in which auditory and phonetic processing could proceed simultaneously. The present experiment attempted to distinguish empirically between these two models. Ss identified either an auditory dimension (fundamental frequency) or a phonetic dimension (place of articulation of the consonant) of synthetic consonant-vowel syllables. When the two dimensions varied in a completely correlated manner, reaction times were significantly shorter than when either dimension varied alone. This "redundancy gain" could not be attributed to speed-accuracy trades, selective serial processing, or differential transfer between conditions. These results allow rejection of a completely serial model, suggesting instead that at least some portion of auditory and phonetic processing can occur in parallel.

Auditory and Phonetic Levels of Processing in Speech Perception

Current theories of speech perception generally view the process by which linguistic information is extracted from an acoustic speech signal as a hierarchy of logically distinct levels or stages (see, for example, Fant, 1967; Fry, 1956; Stevens & Halle, 1967; Stevens & House, 1972; Studdert-Kennedy, in press). Recently, we began a series of experiments designed to investigate possible levels of processing involved in the perception of isolated consonant-vowel (CV) syllables (Day & Wood, 1972a, b; Wood, 1973). The results of these experiments, together with those using other procedures (cf. Studdert-Kennedy & Shankweiler, 1970; Studdert-Kennedy, Shankweiler, & Pisoni, 1972; Tash & Pisoni, 1973), provide empirical support for a basic distinction between two levels of processing in phonetic perception: (a) an *auditory* level, in which an acoustic speech signal is analyzed into a set of corresponding auditory parameters; and (b) a *phonetic* level, in which abstract phonetic features are extracted from the results of the preliminary auditory analysis.

The basic paradigm used by Day and Wood (1972a, b) and Wood (1973) was a two-choice speeded classification task similar to that employed by Garner and Felfoldy (1970) to study patterns of interaction

between stimulus dimensions. Ss were presented a sequence of synthetic CV syllables which varied randomly between two levels of a given dimension and were required to identify which level of that dimension occurred on each trial.¹ Reaction time (RT) for the identification of each dimension was measured under two conditions: (a) a single-dimension *control* condition, in which only the target dimension to be identified varied in the stimulus sequence; and (b) a two-dimension *orthogonal* condition, in which both the target dimension and the irrelevant nontarget dimension varied orthogonally. For a given target dimension, the only difference between the control and orthogonal conditions was the presence or absence of irrelevant variation in the nontarget dimension. Therefore, a comparison of the RTs from these two conditions indicates the degree to which each dimension may be processed independently of irrelevant variation in the other dimension.

The initial experiments in this series analyzed the interactions between the following pairs of dimensions: (a) an auditory and a phonetic dimension, (b) two auditory dimensions, and (c) two phonetic dimensions. Day and Wood (1972a) and Wood (1973, Experiment 1) compared an auditory dimension, fundamental frequency, with a phonetic dimension, place of articulation of voiced stop consonants. For convenience, these dimensions will be referred to as pitch and place, respectively. In both experiments, irrelevant variation in pitch produced substantial interference with the processing of place, but irrelevant variation in place produced minimal interference with the processing of pitch. Thus, when required to identify the auditory dimension, Ss were able to ignore or "filter" irrelevant variation in the phonetic dimension to a considerable degree. In contrast, when required to identify the

*This research was supported by National Institute of Mental Health Grant M-05296 and National Science Foundation Grants GB-3919 and GB-5782 to W. R. Goff, and by National Institute of Child Health and Human Development Grant HD-01994 to the Haskins Laboratories. The author wishes to thank W. R. Goff, R. S. Day, W. R. Garner, J. R. Jennings, and J. D. Wicke for their generous advice and for comments on the manuscript, F. S. Cooper for making available the facilities of Haskins Laboratories for the generation and analysis of the synthetic speech stimuli, and T. C. Fisher and J. Jasiorkowski for technical assistance.

†Present address: Division of Neuropsychiatry, Walter Reed Army Institute of Research, Washington, D.C. 20012.

phonetic dimension, Ss could not ignore irrelevant variation in the auditory dimension.

A different pattern of results was obtained in experiments where the two dimensions were from the same class: both auditory or both phonetic. Wood (1973, Experiment 2) compared the same levels on the pitch dimension used in the initial experiments with another auditory dimension, overall intensity. In this experiment, neither auditory dimension could be processed without substantial interference from irrelevant variation in the other dimension. Finally, for the case of two phonetic dimensions, Day and Wood (1972b) compared place with formant positions for the vowel in the CV syllables. The pattern of interaction between these dimensions was similar to that obtained for the two auditory dimensions: neither phonetic dimension could be processed without substantial interference from irrelevant variation in the other.

These results may be summarized in the following way. When both dimensions were members of the same class, both auditory or both phonetic, the interaction between them in this paradigm was a mutual or symmetric interaction. This result is typical of that obtained for "integral" stimulus dimensions in the experiments of Garner and Felfoldy (1970), and is what would be expected if the two dimensions were extracted by a single perceptual process or by multiple processes which are strongly dependent upon each other. In contrast, the interaction between a phonetic dimension and an auditory dimension was a unidirectional or asymmetric interaction. This kind of interaction suggests: (a) that auditory and phonetic dimensions are not extracted by a single perceptual process, and (b) that the component processes for phonetic information are in some way dependent upon those for auditory information.

Serial vs Parallel Organization of Auditory and Phonetic Levels

Two process models appear to be consistent with the unidirectional interaction between place and pitch: (a) a strict *serial* or sequential model, in which auditory information would be processed at one level, followed by the processing of phonetic information at a subsequent level; and (b) a *parallel* model, in which at least some portion of auditory and phonetic processing could proceed simultaneously.²

A serial organization of auditory and phonetic levels provides a parsimonious account of the unidirectional interference between place and pitch by virtue of the position of each component process in the sequence. According to such a serial model, a response could be based on the output of the initial auditory level alone without processing by the subsequent phonetic level. In contrast, a response based on the phonetic level could occur only after both auditory and phonetic processes had been completed.

A parallel model, however, does not itself provide such a direct and parsimonious account of the unidirectional interference between auditory and phonetic dimensions. Nevertheless, while a parallel model is less parsimonious, it cannot be rejected on the basis of existing data. In fact, a potential advantage of the parallel model over the serial is the finding of the initial experiments that RTs for place and pitch were not significantly different in the control conditions.³

The purpose of the present experiment was to provide more direct evidence concerning serial vs parallel organizations of auditory and phonetic processing. An efficient way of distinguishing between serial and parallel models in the context of the RT experiments described above is to include, in addition to the control and orthogonal conditions, a third condition in which the two dimensions are completely correlated. In such a *correlated* condition, both the target dimension and the redundant nontarget dimension provide sufficient information for a correct response. The outcome of interest is simply whether the additional information provided by the redundant dimension can be used by the S to facilitate performance (i.e., to decrease RT relative to the control conditions without sacrificing accuracy).⁴

A strict serial model would predict no decrease in RT in the correlated condition, since according to this model the pitch dimension would always be processed first, followed by the processing of place at a subsequent level. In contrast, given certain reasonable assumptions, the parallel model would predict a decrease in RT (usually referred to as a "redundancy gain") in the correlated conditions relative to the control conditions for both dimensions. According to this model, the processing of both dimensions could occur simultaneously and the response on each trial would be based on whichever component process was completed first. The decrease in RTs in the correlated conditions would therefore occur statistically from the fact that the RT distribution that results from selecting the faster outcome of two "competing" distributions is faster than either "competing" distribution alone.⁵

METHOD

Subjects

Each of six Ss (five volunteers and the E) served in an experimental session lasting approximately 2 h. Ss were between the ages of 20 and 35 and reported no history of hearing disorder.

Stimuli

The acoustic stimuli were the four synthetic CV syllables /bae/-104 Hz, /bae/-140 Hz, /gae/-104 Hz, and /gae/-140 Hz, corresponding to the two levels on the two dimensions of place and pitch. These two-formant stimuli were generated by the Haskins Laboratories parallel resonance synthesizer and were prepared to be equal in all acoustic parameters other than the two dimensions explicitly varied for experimental purposes. All four stimuli had identical fundamental frequency contours (falling), intensity contours (falling), duration (300 msec), and formant frequencies appropriate for the vowel /ae/. Pairs of

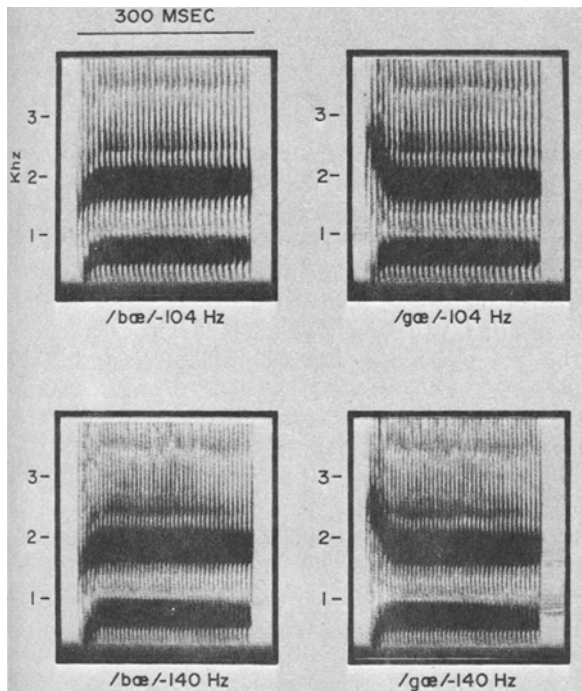


Fig. 1. Spectrograms of the four synthetic syllables. Stimuli differing in place (/bae/ vs /gae/) differed in the direction and extent of the F2 transition (left vs right half of the figure), while stimuli differing in pitch (104 Hz vs 140 Hz) differed in fundamental frequency (upper vs lower half). The four stimuli were identical in all acoustic parameters other than place and pitch.

stimuli differing on the place dimension differed in the direction and extent of the second formant (F2) transition (Liberman, Delattre, Cooper, & Gerstman, 1954; Delattre, Liberman, & Cooper, 1955), while pairs of stimuli differing on the pitch dimension differed in fundamental frequency (initial fundamental frequencies of 104 Hz vs 140 Hz). Spectrograms of the four stimuli are shown in Fig. 1.

Identification Task, Dimensions, and Conditions

Ss listened to blocks of trials in which the place dimension, the pitch dimension, or both dimensions could vary within a block. One dimension was specified as the *target dimension* for each block of trials, and Ss were required to indicate as rapidly as possible which of the two levels on that target dimension occurred on each trial. Ss made their responses by pressing one of two response buttons with either the index or middle finger of the preferred hand. Each button was assigned to the same level on each dimension throughout the experiment.

Place and pitch were each specified as the target dimension in three different conditions, with each condition presented in a separate block of 64 trials. With respect to the target dimension, these three conditions were identical; that is, in all three conditions, the target dimension varied randomly between its two levels and Ss had to identify which level on that dimension occurred on each trial. The only difference between conditions was the status of the nontarget dimension. In the control condition, the nontarget dimension was held constant at one of its two levels throughout the entire block of trials. For half the Ss, the nontarget dimension was held constant at one level, and for the remaining half, it was held constant at the other level. In the orthogonal condition, the target dimension again varied randomly but, in this case, the nontarget dimension varied orthogonally. Thus, all four stimuli occurred randomly in the orthogonal condition. The control and orthogonal conditions

were therefore identical to those employed in the previous RT experiments described above (Day & Wood, 1972a, b; Wood, 1973). Finally, in the correlated condition, both dimensions again varied, but in this case were completely correlated. The stimuli for this condition were /bae/-104 Hz and /gae/-140 Hz, regardless of which dimension was specified as the target dimension.

Apparatus

The stimuli were presented binaurally through Koss Pro-4AA earphones from a Precision Instruments FM tape recorder, through computer-controlled relays and a Grason-Stadler electronic switch. A 64-item series of each of the four stimuli was recorded on a separate channel of the stimulus tape, synchronized under computer control. On a given trial, any of the four stimuli could be presented to the S by closing a relay between the appropriate stimulus channel and the S's earphones. For each block of trials, a LINC computer read a pseudorandom sequence of predetermined stimulus codes from paper tape and closed the appropriate relays in that specified sequence. In this way, the same four-channel stimulus tape was used for all six conditions.

The S's identification responses and RTs were recorded to the nearest millisecond by an external clock (Beckman-Berkley Universal Counter-Timer). The clock was triggered simultaneously with stimulus onset and was halted by the S's response. Following the response on each trial, the LINC read the RT from the counter and punched the stimulus code, response, and the RT for that trial on paper tape for later analysis.

Procedure

Each S received one block of 64 trials in each of the six conditions (2 dimensions by 3 conditions per dimension), in an order specified by a balanced Latin square.

At the beginning of the experimental session, each S was informed of the general nature of the experiment, the stimuli and dimensions to be presented, and the identification tasks to be required. Both speed and accuracy were strongly emphasized in all conditions. For the control condition, Ss were instructed that only the target dimension would vary in that block of trials, and on each trial they were to identify which level on that dimension occurred. For the orthogonal condition, they were told that one dimension would be the target dimension but that the other dimension would also vary. In this condition, they were instructed to identify the target dimension and to ignore variations in the irrelevant nontarget dimension. Finally, in the correlated condition, Ss were told that the two dimensions would be completely correlated, with the levels on both dimensions specifying the same response. In this condition, they were again instructed to identify the target dimension as rapidly and accurately as possible, and to use the extra information provided by the redundant dimension if possible.

Prior to the block of trials for each condition, Ss received at least eight practice trials under conditions identical to those they would receive in the following block. These practice trials were designed to stabilize RT performance and allow Ss to become familiar with the stimulus set and identification task they would receive in that block.

Data Analysis

As in the initial experiments, Ss made very few errors, averaging less than 3% over the entire experiment. Since there were no significant differences in the number of errors between any of the six conditions (Mann-Whitney U tests), error scores will not be considered in detail in the analysis below.

For statistical analysis of the RT data, a complete four-way factorial analysis of variance was computed (Ss by Conditions by Dimensions by Within Cell). The data entered into this analysis were the untransformed RTs, with the single exception that all

Table 1
Mean Reaction Time (in Milliseconds) for
Each Dimension and Condition

| Dimension | Condition | | |
|-----------|-----------|------------|------------|
| | Control | Correlated | Orthogonal |
| Place | 386.8 | 342.8 | 444.1 |
| Pitch | 381.4 | 346.1 | 385.2 |

Note—According to the Scheffé method for individual comparisons, a difference between any pair of means ≥ 29.7 msec is significant at $p < .001$.

values greater than 1 sec were set equal to 1 sec. This procedure eliminated the few very long RTs due to Ss' failure to press the response buttons sufficiently to make electrical contact, etc. Subsequent individual comparisons among main effect and interaction means were made according to the Scheffé procedure (Winer, 1962).

RESULTS

Before examining the correlated conditions for evidence of serial or parallel processing, it is important to establish that the results obtained in the control and orthogonal conditions of the present experiment were similar to those in the corresponding conditions of the initial experiments. These data are presented in Table 1. For place, there was an increase in RT of 57.3 msec from the control to the orthogonal condition, while the increase between conditions for pitch was 3.8 msec. In the analysis of variance, the effects of conditions, dimensions, and the Condition by Dimension interaction were significant, $F(2,126) = 154.3$, $p < .001$, $F(1,63) = 53.73$, $p < .001$, and $F(2,126) = 26.12$, $p < .001$, respectively. According to the results of the Scheffé analysis on the Condition by Dimension interaction means in Table 1, a difference ≥ 29.7 msec was significant at $p < .001$. Thus, the control and orthogonal data are consistent with the unidirectional interference between place and pitch obtained in the initial experiments. Therefore, an analysis of the correlated conditions for evidence of parallel processing is appropriate for these dimensions (cf. Note 4).

In the correlated conditions (Table 1), there were substantial decreases in RT for both dimensions: 44.0 msec for place and 35.3 msec for pitch ($p < .001$). These significant redundancy gains are clearly in accord with the predictions of a parallel model.

However, before concluding that place and pitch can actually be processed in parallel, several alternative explanations for the redundancy gain must be ruled out by further analyses of the data. First, as mentioned above, there were no significant differences in errors between any conditions of the experiment, therefore eliminating the possibility that speed-accuracy trades could be responsible for the observed reduction in RT in the correlated conditions.

A second way in which a redundancy gain could be obtained in the absence of parallel processing is the

strategy of "selective serial processing" ("SSP," cf. Biederman & Checkosky, 1970; Garner, 1969; Garner & Felfoldy, 1970; Felfoldy & Garner, 1971; Morton, 1969). According to the SSP strategy, the underlying mode of processing is strictly serial but Ss are presumed to have the ability to select which dimension they actually process in tasks with redundant dimensions. An apparent redundancy gain could therefore be produced if each S performed the correlated task based on the faster of the two single dimensions, regardless of which was specified as target dimension by the instructions.

In an explicit analysis of the SPP strategy, Felfoldy and Garner (1971) suggested that either of the following two conditions must be met in order for SPP to be effective: (a) that RTs for one dimension are significantly faster than the other across all Ss (as would be the case if the two dimensions differed greatly in discriminability); or (b) that each S is able to identify one dimension significantly faster than the other, but the faster dimension is different for different Ss. In either case, the SSP strategy could produce decreased RTs in the correlated relative to the control conditions, but without true parallel processing. The first of these two conditions for SSP was clearly not met in the present experiment, since mean RT for place and pitch in the control conditions was not significantly different (Table 1). However, there was a significant interaction of Ss by Dimensions, $F(5,315) = 25.24$, $p < .001$, indicating that there were reliable differences in RTs between the two dimensions for individual Ss. Therefore, it is logically possible that SSP could have produced the redundancy gains.

Whether or not Ss actually used the SSP strategy can be evaluated directly by comparing the RTs from the correlated conditions to RTs from the faster of the two control conditions for each S (Biederman & Checkosky, 1970; Garner, 1969; Garner & Felfoldy, 1970; Felfoldy & Garner, 1971; Morton, 1969). For each S, optimal performance under the SSP model would result in the RT of the correlated conditions being equal to the RT of that S's faster control condition. In contrast, the parallel model predicts that RTs in the correlated conditions will be faster than the control, even after correcting for the possible use of SSP. In the present experiment, the mean of the faster control RTs for each S was 369.4 msec. In contrast, mean RTs in the correlated conditions were 342.8 msec for place and 346.1 msec for pitch (Table 1). A separate analysis of variance and subsequent individual comparisons among these means (Scheffé method) showed that RTs in both correlated conditions were significantly faster than each S's faster control RTs, $F(2,126) = 14.05$, $p < .001$. Therefore, the obtained redundancy gain cannot be attributed to the SSP strategy.

A third source of an apparent redundancy gain in the absence of parallel processing is the possibility of differential transfer between control and correlated conditions. As pointed out by Biederman and

Checkosky (1970), greater positive transfer between the two correlated conditions than between the two control conditions might tend to reduce artificially the RTs in the correlated relative to the control conditions. This would be true both for parallel processing and for SSP, since in either case the two correlated conditions would actually be repetitions of the same task while the two control conditions would always be different. To examine this possibility of differential transfer, the control conditions received first and second in sequence by each S were compared to the correlated conditions received first and second by each S, without the regard to the actual target dimension in each condition. For the control conditions, mean RT for the second block was 1.3 msec faster than the first, while for the correlated conditions, the second block was 2.9 msec slower. Thus, there was minimal transfer between the two blocks of trials within each condition, and the direction of the obtained differences favored the control and not the correlated conditions.⁶

DISCUSSION

The results of the present experiment have implications both for models of speech perception in a narrow sense and for the broader question of how multidimensional stimuli of any kind are perceived. The relation of the present results to both these problem areas will be discussed below.

Relation to Multidimensional Information Processing

The way in which humans process multidimensional stimuli has been the subject of considerable theoretical and experimental effort (cf. Egeth, 1966, 1967; Egeth & Pachella, 1969; Garner, 1962, 1970, 1973, in press; Lockhead, 1966, 1970, 1972; Nickerson, 1971; Posner, 1964). The large number of experiments devoted to this question can be classified into two distinct approaches or patterns of major emphasis: (a) those which emphasize *stimulus concepts* and focus upon the nature of the stimulus dimensions that make up the multidimensional stimuli, and (b) those which emphasize *processing concepts* and focus upon the nature of the processes by which the multidimensional stimuli are perceived.

Major emphasis upon stimulus concepts has come from the distinction between integral and separable stimulus dimensions (Garner & Felfoldy, 1970; Garner, 1970, 1973, in press). Expanding upon previous suggestions by Torgerson (1958), Attneave (1962), Shepard (1964), Lockhead (1966), and Hyman and Well (1968), Garner and Felfoldy (1970) argued that the concept of integral dimensions could best be defined by converging experimental operations (Garner, Hake, & Eriksen, 1956): "Integral dimensions are those which lead to a Euclidean metric in direct distance scaling, produce a redundancy gain when the dimensions are correlated and some measure of speed of classification is

used, and produce interference in speed of classification when selective attention is required with orthogonal stimulus dimensions [p. 238]." A fourth converging operation upon the integral-separable distinction is provided by the experiment of Handel and Imai (1972) based on free classification tasks. Since the present experiment employed two of the converging operations listed above, conditions directly analogous to those of Garner and Felfoldy (1970), the present results may be compared to those expected of integral and separable dimensions.

When place was the target dimension, the results corresponded exactly to those of one member of an integral pair of dimensions: a significant redundancy gain in the correlated condition and significant interference in the orthogonal condition. However, when pitch was the target dimension, the results were consistent with neither integral nor separable dimensions. In this case, there was again a significant redundancy gain in the correlated condition, but there was minimal interference in the orthogonal condition.

These results pose two difficulties for a strict dichotomous distinction between integral and separable dimensions. First, the interference between place and pitch in the orthogonal conditions was unidirectional or asymmetric. Previously, the integral-separable distinction has always been considered to be symmetric, with either equal interference or no interference between a pair of dimensions. However, as discussed in detail by Garner (1973), recent data suggest that integrality and separability may be more accurately considered as the two ends of a continuum rather than a strict dichotomy. The present data support this suggestion.

The second difficulty for the integral-separable distinction posed by the present results is that the processing of the place and pitch dimensions was affected differentially by the correlated and orthogonal conditions. That is, instead of strictly obeying the converging operations notion of integrality and separability presented above, place and pitch produced one pattern of results in the correlated conditions and another pattern in the orthogonal conditions. These results suggest that Ss appear to have some option regarding the way in which place and pitch are processed: (a) they can process the two dimensions in parallel as demonstrated by the redundancy gain for both dimensions in the correlated conditions, or (b) they can in some way process the pitch dimension selectively (e.g., filter or attenuate the processing of place) as demonstrated by the minimal interference produced by irrelevant variation in place in the orthogonal condition.⁷

The processing options for place and pitch also create difficulties for existing process models. A number of authors have attempted to develop a single serial or parallel model which could account for a wide variety of information processing tasks (see discussions by Biederman & Checkosky, 1970; Egeth, 1966; Grill,

1971; Hawkins, 1969; Nickerson, 1971; Saraga & Shallice, 1973; Smith, 1968; Sternberg, 1969; Townsend, 1971). Basic serial and parallel models have been modified to include distinctions between exhaustive and self-terminating processes, fixed or random orders of search, fixed or variable durations for each component process, and a number of others. In addition, Lockhead (1972) presented a "blob" or holistic processing model in which a multidimensional stimulus must first be processed holistically with subsequent serial processes as required by the task. The results of the control and correlated conditions of the present experiment are consistent with either a parallel model or Lockhead's "blob" model. However, when the data from the orthogonal conditions are also included, it is clear that all of the present results cannot be accounted for by any model which specifies a particular *mandatory* process—serial, parallel, or holistic. These data provide further evidence for recent suggestions made by Garner (in press), Grill (1971), Nickerson (1971), and Townsend (1971), that no single stimulus distinction or process model may be universally appropriate. Rather, under different conditions, Ss may employ various strategies, including serial processing, parallel processing, holistic processing, and various combinations thereof, depending upon the constraints of the stimuli and perceptual tasks involved. A useful strategy for future experiments would appear to be: (a) the empirical distinction between processes that are mandatory and those that are optional; and (b) the detailed analysis of the stimulus properties and task conditions related to each (cf. Garner, in press; Grill, 1971; Nickerson, 1971; Townsend, 1971).

Relation to Models of Speech Perception

The idea that speech perception may involve some form of parallel processing has been suggested on logical grounds by a number of investigators (cf. Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967; Stevens & House, 1972; Studdert-Kennedy, in press; Wickelgren, 1972). The results of the present experiment provide empirical evidence that at least some portion of auditory and phonetic processing can occur in parallel, and provide a starting point for the investigation of possible parallel processing of other kinds of linguistic and nonlinguistic information in speech perception.

Superficially, these results appear to contradict the widely held assumption that auditory and linguistic processes are organized serially, with the linguistic processes dependent upon those performed by the general auditory system. This idea is clearly summarized by Stevens and House (1972): "All acoustic signals undergo some common peripheral processing, and up to a certain point in the auditory system the nature of this early processing is the same whether the signal is speech or is not speech [p. 9]." The idea that linguistic processes are dependent upon the peripheral auditory system must be correct in some form, since any acoustic

signal must initially be transduced by the receptor apparatus before subsequent processing can occur. This anatomical constraint indicates that at some level auditory and phonetic processes must be organized serially, since some subset of auditory processes (e.g., receptor transduction) must occur before phonetic processing can begin. The present results are not at all in conflict with this requirement. Rather, the present data require only that phonetic processing can occur in parallel with some *other* subset of auditory processes. Some important issues remaining to be determined are the functional and anatomical loci at which phonetic processing begins, and the specific auditory processes that occur before and after the onset of phonetic processing.

The general organization of auditory and phonetic processes that seems necessary to account for these observations would involve both serial and parallel processing, and would include the following three components: (a) a common "peripheral" component for the transduction and preliminary analysis of all acoustic signals, (b) a "central" auditory component for the additional processing of nonlinguistic auditory information, and (c) a "central" phonetic component for the extraction of phonetic features from the results of the preliminary auditory analysis. The two "central" components would be capable of functioning in parallel, but both would be dependent upon the output of the preliminary auditory analysis. From this point of view, the common peripheral stage could not be directly manipulated in RT tasks like those of the present experiment, since this stage would be mandatory for all auditory processing tasks. As a working hypothesis, this serial-parallel or "hybrid" organization would be consistent with the capability for parallel processing of auditory and phonetic information demonstrated by the present experiment, and with the previous experiments which distinguish between auditory and phonetic processes in speech perception (cf. Liberman et al, 1967; Mattingly, Liberman, Syrdal, & Halwes, 1971; Studdert-Kennedy & Shankweiler, 1970; Studdert-Kennedy et al, 1972; Tash & Pisoni, 1973; Wood, Goff, & Day, 1971; Wood, 1973).

REFERENCES

- Attneave, F. Perception and related areas. In S. Koch (Ed.), *Psychology: A study of a science*. Vol. 4. New York: McGraw-Hill, 1962. Pp. 619-659.
- Biederman, I., & Checkosky, S. F. Processing redundant information. *Journal of Experimental Psychology*, 1970, 83, 486-490.
- Day, R. S., & Wood, C. C. Interactions between linguistic and nonlinguistic processing. *Journal of the Acoustical Society of America*, 1972a, 51, 79 (A).
- Day, R. S., & Wood, C. C. Mutual interference between two linguistic dimensions of the same stimuli. *Journal of the Acoustical Society of America*, 1972b, 52, 175 (A).
- Delattre, P. C., Liberman, A. M., & Cooper, F. S. Acoustic loci and transitional cues for consonants. *Journal of the Acoustical Society of America*, 1955, 27, 769-773.
- Egeth, H. E. Parallel versus serial processes in multidimensional stimulus discrimination. *Perception & Psychophysics*, 1966, 1, 245-252.

- Egeth, H. E. Selective attention. *Psychological Bulletin*, 1967, 67, 41-57.
- Egeth, H. E., Jonides, J., & Wall, S. Parallel processing of multielement displays. *Cognitive Psychology*, 1972, 3, 674-698.
- Egeth, H. E., & Pachella, R. Multidimensional stimulus identification. *Perception & Psychophysics*, 1969, 5, 341-346.
- Fant, G. Auditory patterns of speech. In W. Wathen-Dunn (Ed.), *Models for the perception of speech and visual form*. Cambridge: M.I.T. Press, 1967. Pp. 111-125.
- Fry, D. B. Perception and recognition in speech. In M. Halle, H. G. Lunt, and C. H. van Schooneveld (Eds.), *For Roman Jakobson*. The Hague: Mouton, 1956. Pp. 169-173.
- Felfoldy, G. L., & Garner, W. R. The effects on speeded classification of implicit and explicit instructions regarding redundant dimensions. *Perception & Psychophysics*, 1971, 9, 289-292.
- Garner, W. R. *Uncertainty and structure as psychological concepts*. New York: Wiley, 1962.
- Garner, W. R. Speed of discrimination with redundant stimulus attributes. *Perception & Psychophysics*, 1969, 6, 221-224.
- Garner, W. R. The stimulus in information processing. *American Psychologist*, 1970, 25, 350-358.
- Garner, W. R. Paul M. Fitts memorial lectures. Presented at the University of Michigan, 1973.
- Garner, W. R. Attention: The processing of multiple sources of information. In E. C. Carterette and M. P. Friedman (Eds.), *Handbook of perception*. Vol. 2, in press.
- Garner, W. R., & Felfoldy, G. L. Integrality of stimulus dimensions in various types of information processing. *Cognitive Psychology*, 1970, 1, 225-241.
- Garner, W. R., Hake, H. W., & Erickson, C. W. Operationism and the concept of perception. *Psychological Review*, 1956, 63, 149-159.
- Grill, D. P. Variables influencing the mode of processing of complex stimuli. *Perception & Psychophysics*, 1971, 10, 51-57.
- Handel, S., & Imai, S. The free classification of analyzable and unanalyzable stimuli. *Perception & Psychophysics*, 1972, 12, 108-116.
- Hawkins, H. L. Parallel processing in complex visual discrimination. *Perception & Psychophysics*, 1969, 5, 56-64.
- Hyman, R., & Well, A. Perceptual separability and spatial models. *Perception & Psychophysics*, 1968, 3, 161-165.
- Liberman, A. M., Cooper, F. S., Shankweiler, D., & Studdert-Kennedy, M. Perception of the speech code. *Psychological Review*, 1967, 74, 431-461.
- Liberman, A. M., Delattre, P. C., Cooper, F. S., & Gerstman, L. J. The role of consonant-vowel transitions in the perception of the stop and nasal consonants. *Psychological Monographs*, 1954, Whole No. 68.
- Lockhead, G. R. Effects of dimensional redundancy on visual discrimination. *Journal of Experimental Psychology*, 1966, 72, 95-104.
- Lockhead, G. R. Identification and the form of multidimensional discrimination space. *Journal of Experimental Psychology*, 1970, 85, 1-10.
- Lockhead, G. R. Processing dimensional stimuli: A note. *Psychological Review*, 1972, 79, 410-419.
- Mattingly, I. G., Liberman, A. M., Syrdal, A. K., & Halwes, T. Discrimination in speech and nonspeech modes. *Cognitive Psychology*, 1971, 2, 131-157.
- Morton, J. The use of correlated stimulus information in card sorting. *Perception & Psychophysics*, 1969, 5, 374-376.
- Nickerson, R. S. Binary-classification reaction time: A review of some studies of human information-processing capabilities. *Psychonomic Monograph Supplements*, 1971, 4(17, Whole No. 65).
- Posner, M. I. Information reduction in the analysis of sequential tasks. *Psychological Review*, 1964, 71, 491-504.
- Posner, M. I. Abstraction and the process of recognition. In G. Bower and J. T. Spence (Eds.), *The psychology of learning and motivation*. Vol. 3. New York: Academic Press, 1969. Pp. 43-100.
- Posner, M. I., & Mitchell, R. F. Chronometric analysis of classification. *Psychological Review*, 1967, 74, 392-409.
- Saraga, E., & Shallice, T. Parallel processing of the attributes of single stimuli. *Perception & Psychophysics*, 1973, 13, 261-270.
- Shepard, R. N. Attention and the metric structure of the stimulus space. *Journal of Mathematical Psychology*, 1964, 1, 54-87.
- Smith, E. E. Choice reaction time: An analysis of the major theoretical positions. *Psychological Bulletin*, 1968, 69, 77-110.
- Sternberg, S. The discovery of processing stages: Extensions of Donders' method. In W. G. Koster (Ed.), *Attention and performance II*. Amsterdam: North Holland, 1969. Pp. 276-315.
- Stevens, K. N., & Halle, M. Remarks on analysis by synthesis and distinctive features. In W. Wathen-Dunn (Ed.), *Models for the perception of speech and visual form*. Cambridge: M.I.T. Press, 1967. Pp. 88-102.
- Stevens, K. N., & House, A. S. Speech perception. In J. V. Tobias (Ed.), *Foundations of modern auditory theory*. New York: Academic Press, 1972. Pp. 1-62.
- Studdert-Kennedy, M. The perception of speech. In T. A. Sebeok (Ed.), *Current trends in linguistics*. Vol. XII. The Hague: Mouton, in press. (Also appeared in Haskins Laboratories Status Report on Speech Research, SR-23, 1970, pp. 15-48.)
- Studdert-Kennedy, M., & Shankweiler, D. Hemispheric specialization for speech perception. *Journal of the Acoustical Society of America*, 1970, 48, 579-594.
- Studdert-Kennedy, M., Shankweiler, D., & Pisoni, D. Auditory and phonetic processes in speech perception: Evidence from a dichotic study. *Cognitive Psychology*, 1972, 3, 455-466.
- Tash, J., & Pisoni, D. B. Auditory and phonetic levels of processing as revealed by reaction time. Paper presented at the 85th Meeting of the Acoustical Society of America, Boston, Massachusetts, 1973.
- Torgerson, W. S. *Theory and methods of scaling*. New York: Wiley, 1958.
- Townsend, J. T. A note on the identifiability of parallel and serial processes. *Perception & Psychophysics*, 1971, 10, 161-163.
- Wickelgren, W. A. Discussion paper on speech perception. In J. H. Gilbert (Ed.), *Speech and cortical functioning*. New York: Academic Press, 1972. Pp. 237-262.
- Winer, B. J. *Statistical principles in experimental design*. New York: McGraw-Hill, 1962.
- Wood, C. C. Levels of processing in speech perception: Neurophysiological and information-processing analyses. Unpublished doctoral dissertation, Yale University, 1973.
- Wood, C. C., Goff, W. R., & Day, R. S. Auditory evoked potentials during speech perception. *Science*, 1971, 173, 1248-1251.

NOTES

1. The term "dimension" is used in the present paper to refer to aspects or properties of stimuli that are varied in a given experiment. It should be emphasized that this term explicitly does not imply that the stimulus property in question is singular or unitary in a perceptual sense. The latter is an empirical question (cf. Garner, 1970, 1973, in press; Garner & Felfoldy, 1970).

2. Nickerson (1971) pointed out an important limitation of the terms "simultaneous" and "parallel" that is directly relevant to the present experiment: "What may appear to be simultaneous activities at one level of analysis may prove to be the result of an efficient switching process when the analysis is carried to a more microtemporal level [p. 276]." Following Nickerson, the terms "simultaneous" and "parallel" will be used in the present paper to indicate: "... that the processes in question are proceeding concurrently *relative to the time scale on which they are measured*, which is to admit the possibility of intermittent switching of attention between one process and another on a more microtemporal scale [p. 277]."

3. According to a strict serial model, any task which requires a response based on information about place would also require prior processing of information about pitch, therefore resulting in longer processing times for place than for pitch (for data which seem to satisfy such a model, see Posner & Mitchell, 1967; Posner, 1969). As described above, RTs in the orthogonal condition of the initial experiments were indeed significantly longer for place than for pitch. However, according to a serial model, this result should be true for the control condition as well, since in both conditions identification of pitch would require only one level of processing, while identification of place would involve processing at both levels. The control condition RTs were in fact slightly longer for place than for pitch in both of the initial experiments (mean differences of 10.0 and 3.4 msec), although in neither case was the difference statistically significant. Thus, the existing data are equivocal with respect to the serial-parallel distinction. These results suggest either: (a) that processing of auditory and phonetic information can occur in parallel, or (b) that processing is serial and the equal control RTs were the result of other factors, such as the relative discriminability of the particular levels on the place and pitch dimensions employed in the initial experiments.

4. Garner (1970; see also Garner & Felfoldy, 1970) has argued that the distinction between "integral" and "separable" stimulus dimensions must logically precede any consideration of serial vs parallel processing, since the serial-parallel question is inappropriate for pairs of integral dimensions. The unidirectional interference between place and pitch in the initial experiments is clear evidence that these dimensions are not completely integral, and therefore that the serial-parallel question is appropriate.

5. For a detailed analysis of the assumptions and predictions of various serial and parallel models in this and other paradigms, see Biederman and Checkosky (1970), Egeth (1966), Egeth, Jonides, and Wall (1971), Grill (1971), Hawkins (1969), Lockhead (1972), Nickerson (1971), Smith (1968), Saraga and Shallice (1973), and Townsend (1971). Biederman and Checkosky (1970) and Lockhead (1972) present particularly clear discussions of the statistical rationale for the prediction of a parallel model that RTs in the correlated condition should be faster than those of the single-dimension control conditions.

6. The question of explicit instructions to the Ss regarding the correlation between dimensions should also be considered as a possible source of misinterpretation. It might be argued that such instructions could bias Ss toward responding faster in the correlated conditions, leading to a false conclusion about a redundancy gain. In their analysis of the necessary conditions for the SSP strategy to be effective, Felfoldy and Garner (1971) compared the effects of explicit instructions about the correlation between dimensions with implicit instructions consisting of exposure to only the control and correlated stimulus sequences. Using stimulus dimensions which do not produce a redundancy gain under conditions of neutral instructions (Garner & Felfoldy, 1970), Felfoldy and Garner (1971) showed that the explicit instructions produced extensive use of SSP by all Ss, but no evidence of a redundancy gain beyond that attributable to SSP alone. That is, RTs in the correlated conditions were significantly faster than the control conditions, but were not significantly different from each S's faster control condition.

Explicit instructions were therefore used in the present

experiment to maximize the ability to discriminate between a redundancy gain due to parallel processing and one produced by SSP. Indeed, a result demonstrating selective serial processing of auditory and phonetic information would have implications for models of speech perception as important as those of parallel processing. The correction for SSP described above rules out possible biased reduction of RTs in the correlated condition *per se*, since it demonstrates that correlated RTs were faster than would be possible according to serial processing alone. However, this correction does not completely eliminate the possibility that the explicit instructions produced artificially inflated RTs in the control conditions, thereby producing relatively faster RTs in the correlated conditions than the control. This possibility is unlikely, since: (a) the control RTs for both dimensions in the present experiment averaged 25 msec faster than those of extremely well-practiced Ss in the experiment of Wood (1973; Experiment 1) under similar conditions, and (b) the differences in RT between control and orthogonal conditions in this experiment were virtually identical to those of Wood (1973; Experiment 1). Thus, if Ss inflated the control RTs, they would have also had to inflate the orthogonal RTs by a precisely equal amount. These observations make the biased inflation of the control RTs extremely unlikely.

7. The present experiment provides little information concerning the precise nature of this filtering or attenuation process. This question is currently being investigated.

(Received for publication September 26, 1973;
revision received January 7, 1974.)