Vowel Similarity, Connectionist Models, and Syllable Structure in Motor Programming of Speech

ILAN YANIV

University of Chicago

DAVID E. MEYER

University of Michigan

PETER C. GORDON

Harvard University

AND

CAROL A. HUFF AND CHRISTINE A. SEVALD

University of Michigan

Using a response-priming procedure, five experiments examined the effects of vowel similarity on the motor programming of spoken syllables. In this procedure, subjects prepared to produce a pair of spoken syllables as rapidly as possible, but sometimes had to produce the syllables in reverse order instead. The spoken responses consisted of consonant–vowel–consonant (CVC) syllables whose medial vowels were /i/, /I/, /ʌ/, and /ɑ/. Performance was measured as a function of the phonetic relationship between the vowels in a syllable pair. Longer response latencies occurred for syllable pairs that contained similar vowels (e.g., /i/ and /I/) than for syllable pairs that contained dissimilar vowels (e.g., /i/ and /ʌ/). This inhibitory vowel-similarity effect occurred regardless of whether the initial consonants of the syllables in a pair were the same or different. However, it decreased substantially when the final consonants of the paired syllables were different. These results suggest that a lateral-inhibition mechanism may modulate the motor programming of vowels during speech production. They also provide evidence for the integrity of vowel-consonant (VC) subunits in syllables. © 1990 Academic Press, Inc.

Several models of speech production, speech perception, and visual word-recognition have represented linguistic units (e.g., letters, phonemes, and syllables) as nodes at various levels in a hierarchical interactive-activation network (Dell, 1985, 1986; MacKay, 1982; McClelland & Elman, 1986; McClelland & Rumelhart, 1981; Meyer & Gordon, 1985). According to such models, performance is controlled by interactions among large numbers of excitatory and inhibitory connections that modulate the flow of activation between the units of the network. In speech production, the selection of speech units is assumed to be determined simultaneously and interactively by the activation of nodes for higher units (e.g., phonemes, words) and lower

1

units (e.g., phonetic features, muscle commands) as well as by the types of connections between these nodes.

Typically, "vertical" excitatory connections between nodes provide pathways for top-down and bottom-up activation, whereas "horizontal" inhibitory connections influence the selection among competing units within the same level. During speech production, for example, this inhibition may help tune motor programs and guard against speech errors such as anticipations and exchanges between similar segments (Meyer & Gordon, 1985; cf. Fromkin, 1971). An important theoretical question concerns the psychological reality of the structural representations embodied in these connectionist models.

Some speech-production experiments by Meyer and Gordon (1985) provide results consistent with the hypothesis of "horizontal" inhibitory connections between representations of speech units. In their experiments, subjects prepared to produce a primary response (e.g., saying the syllable pair "ub–ut") but then sometimes had to produce a secondary response instead (e.g., saying the syllable pair "ut–ub"). The secondary responses always reversed the order of the syllables in the primary responses. An auditory response signal (high or low tone) specified the required response type (primary or secondary). The consonants in the primary and secondary responses varied systematically with respect to place-of-articulation and voicing features. For example, the syllables of the response "ub–up" differed in their voicing feature (i.e., voiced vs. voiceless), whereas the syllables of the response "up–ut" differed in their place-of-articulation feature (i.e., labial vs. alveolar). Response latency was measured as a function of relations between the phonetic features of the syllables in the responses.

Meyer and Gordon (1985) found that longer latencies and more errors occurred when the consonants of secondary-response syllable pairs shared either a place-of-articulation feature (e.g., "ut–ud") or a voicing feature (e.g., "ut–up") than when they shared neither of these features (e.g., "ut–ub"). Such results suggest that production of units at the phonemic level may involve an inhibitory process whose influence is greater for similar units than for dissimilar units. Lateral inhibition can provide an important tuning capacity in network systems that utilize both top-down and bottom-up activation (McClelland & Rumelhart, 1981; Meyer & Gordon, 1985). As noted earlier, top-down and bottom-up activation is transmitted through such systems via vertical connections between word nodes, syllable nodes, phoneme nodes, feature nodes, and motor-command nodes that control the speech articulators (cf. MacKay, 1982). Because the activation of nodes at a particular level may be relatively strong for several units that share some of the same features, a mechanism is needed to help determine which node(s) should control the ultimate output of the system. Inhibitory connections between the coordinate nodes in each level can fulfill this need.

For example, consider the pattern of activation and feedback for consonant nodes during the production of the utterance "ub–ud." Here the nodes /b/ and /d/ would have vertical connections to lower units such as the voicing-feature node. In programming the "ub" part of the utterance "ub–ud," activation would go from the /b/ node to the voicing-feature node. Furthermore, the voicing-feature node would be connected to *all* voiced phonemes, so a portion of its activation would be fed back upward to nodes of other potentially competing voiced phonemes, in particular, the /d/ phoneme. Thus, during the programming of the first syllable in an utterance, competition would occur from the second syllable if the two syllables share some feature. The role of lateral inhibitory connections is to keep the activation level of competing units below threshold, preventing inadvertent substitutions. A consequence of this inhibitory tun-

ing process is that the programming of utterances containing similar speech units may take longer than the programming of utterances containing dissimilar speech units.

An inhibitory tuning process would fit very naturally into various connectionist models that characterize psychological phenomena in terms of connections among simple processing units. In particular, inhibition among units within the same level of analysis can serve clear computational purposes (Rumelhart & McClelland, 1986), and lateral inhibition between nearby cells is a well-established neurophysiological characteristic of perceptual systems (Hartline & Ratliff, 1957). Yet while connectionist models provide a natural account of inhibition that has been observed in producing similar speech units (Meyer & Gordon, 1985), it could also potentially be explained within other frameworks. Therefore, throughout the remainder of this article, except when explicitly interpreting experimental results, we will adopt neutral terminology, referring for example to "vowel representations" or even simply "vowels," rather than "vowel nodes."

In this article, we extend our previous research on inhibition in the production of similar consonants (Meyer & Gordon, 1985) to vowels and other constituents of English syllables. Such an extension is important because consonants and vowels are separate, equally basic, classes of speech units that can be distinguished on articulatory (Catford, 1977; Öhman, 1966), acoustic (Ladefoged, 1971), and linguistic (Chomsky & Halle, 1968) grounds. One of our major goals here is to argue that a mechanism involving inhibition among similar speech-production units may underlie performance with vowels as well as consonants. We will also demonstrate that an explanation of this inhibition must take into account certain aspects of syllable structure (cf. Dell, 1985; Mackay, 1982).

The remainder of the article is organized as follows. First, we briefly describe some basic characteristics of vowels that are relevant to our research and that serve to distinguish them from consonants. Next, we outline a response-priming procedure used here for studying the motor programming of vowels in multi-syllabic utterances. Then, we report five experiments with this procedure, showing how the results bear on possible connectionist models of the programming process. More specifically, Experiment 1 provides evidence that preparing to produce a vowel involves inhibiting the production of similar vowels. Experiments 2 through 5 show that certain aspects of syllable structure modulate the extent of this inhibition. It appears, in particular, that significant inhibition occurs if, and only if, the vowels are embedded in syllables with identical final consonants. Such an outcome may have important implications both for theories of syllable structure and for the nature of inhibitory connections among subsyllabic speech units.

DESCRIPTION OF VOWELS

Comparison and Contrast with Consonants

One useful way to describe the major characteristics of vowels involves comparing them with consonants (Ladefoged, 1971). The action of the vocal tract differs in several respects for the production of vowels versus consonants. Vowels correspond to relatively open configurations of the vocal tract. They are produced by relatively slow, continuous alterations in the shape of the vocal tract achieved largely by repositioning the tongue body and lips. Distinctions between vowels result from global differences in vocal-tract shape. Consonants, on the other hand, correspond to relatively constricted configurations of the vocal tract. They are produced by relatively fast, complex movements of various articulators (e.g., tongue tip, lips, and velum) that require more temporal precision. Distinctions between consonants result from the location of the constriction, the manner

in which it is achieved, and the state of the larynx.

Significant functional separateness between vowels and consonants has been demonstrated by the finding that electromyographic (EMG) activity associated with producing vowels decreases at higher speaking rates, whereas EMG activity associated with producing consonants increases (Gay, Ushijima, Hirose, & Cooper, 1974). This pattern indicates that, at higher speaking rates, the precision of vowel articulation diminishes because of decreased muscular activity, whereas the greater precision of consonant articulation requires increased muscular activity.

Based on such differences, Fowler (1980) and Perkell (1969) have suggested that vowels and consonants are produced by different neuromuscular systems. Furthermore, it has been suggested that vowels are produced cyclically in a continuous stream somewhat separate from the production of consonants, with the leading and trailing edges of vowels being overlaid by consonants (Fowler, 1983). Given such suggestions, one would not necessarily expect that empirical results and theoretical hypotheses developed for consonant production should generalize completely to vowels. In particular, it remains an open question whether the evidence of lateral inhibition in the production of similar consonants (Meyer & Gordon, 1985) would generalize to vowels.

The Vowel Space

For present purposes, we have selected four English vowels to be included in our experiments. These vowels are /i/ as in PEAT, /I/ as in PICK, /ʌ/ as in PUCK, and /ɑ/ as in POT. The motivation for this selection may be understood in terms of a classification system devised by the phonetician Daniel Jones to describe the vowels of various languages (see Catford, 1977, Chap. 9, for a discussion of this system). In Jones' system, all vowels are characterized relative to a set of *cardinal vowels,* which serve as universal reference points for ar-

ticulatory displacements of the tongue within the mouth. The cardinal vowels vary on three features: vertical tongue position (high–low), horizontal tongue position (front–back), and lip position (rounded–unrounded). They provide a standard scheme for representing and comparing vowels regardless of the language from which they come. Figure 1 illustrates a primary subset of the cardinal vowel system.

With respect to Jones' system, two of the vowels selected for our experiments (/i/ and /ɑ/) are located at opposite extremes of the vowel space (Fig. 1). The vowel /i/ is similar to Cardinal Vowel 1, being the highest and most frontal English vowel. The vowel /ɑ/ is located at the other end of a diagonal line that extends from the upper left corner to the lower right corner of the vowel space. It is near Cardinal Vowel 5, being the lowest and furthest back vowel in English. The other two vowels selected for our experiments are paired with each of these first two, and come from the same regions of the vowel space as they do. /I/ is the closest English vowel to /i/, being frontal and mid-high. Similarly, /ʌ/ is near /ɑ/, being mid-back and mid-low. The vowels studied here thus incorporate multiple levels of similarity, ranging from high-similarity pairs like /i/–/I/ and /ʌ/–/ɑ/ to low-similarity pairs like /i/–/ʌ/ and /I/–/ɑ/.[1]

A comparable characterization of the similarity among these vowels may be reached by considering the phonological features that they share, following a classification scheme proposed by Chomsky and

---

[1] A third possible dimension for classifying vowels is lip position (the rounding feature). Vowels may be either "unrounded" or "rounded." Catford (1977, p. 178) has suggested that the vowel /ɑ/ is "very slightly rounded." However, in Ladefoged's (1971) classification, none of our four vowels (/i/, /I/, /ʌ/, and /ɑ/) is considered rounded. Also, electromyographic recordings of the orbicularis oris muscle, which mainly controls lip rounding, have revealed very little activity before and during production of these vowels (Fromkin, 1966). Hence, the two-dimensional spatial representation in Fig. 1 is probably an adequate description of the vowels chosen for present purposes.
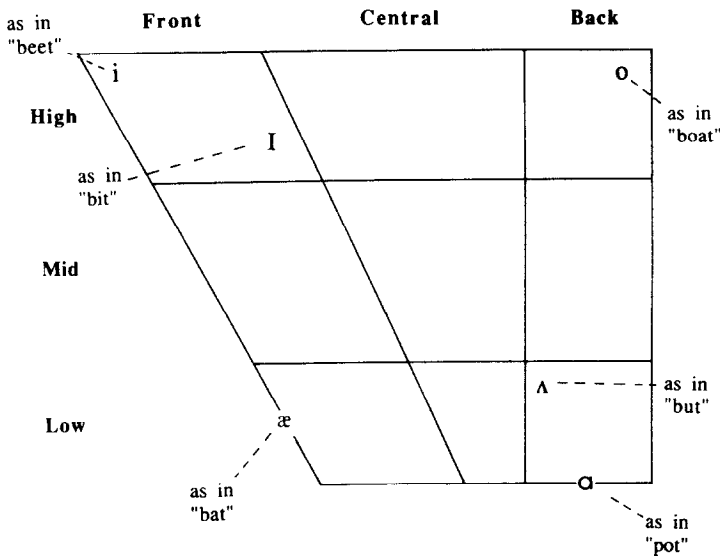
as in "beet"   Front   Central   Back

High   o   as in "boat"

i

I

as in "bit"

Mid

Low   æ

Λ — — — — as in "but"

as in "bat"   ɑ

as in "pot"

FIG. 1. The cardinal vowel system (adapted from J. C. Catford, 1977).

Halle (1968). According to this scheme, the vowels /i/ and /I/ are both [+high], [−back], [−low], [−round], and differ only in tenseness; /i/ is [+tense], /I/ is [−tense]. The vowels /ʌ/ and /ɑ/ are both [−high], [+back], [+low], [−round], and also differ only in tenseness: /ʌ/ is [−tense] and /ɑ/ is [+tense].

These vowels therefore seem suitable for testing hypotheses about inhibition among similar speech-production units. If inhibitory mechanisms affect the selection among similar speech units during motor programming, then this should be revealed by varying the similarity between vowels in pairs of response syllables used for our response-priming procedure.

THE RESPONSE-PRIMING PROCEDURE

Our response-priming procedure is related to ones used elsewhere in studies of spreading activation, response preparation, and motor programming (Meyer, Sternberg, Knoll, & Wright, 1978; Meyer, Yantis, Osman, & Smith, 1985; Posner & Snyder, 1975; Rosenbaum & Kornblum, 1982). The current version of the procedure most closely resembles that of Meyer and Gordon (1985, Experiment 1; see also Gordon

& Meyer, 1987), who have thoroughly discussed its rationale.

On each trial of the procedure, the subject is required to produce a spoken response consisting of a pair of monosyllabic English words or pseudowords (e.g., PEAT–PIT). The primary-response pair is specified visually on the display screen of a computer terminal, and the subject is instructed to prepare for producing it upon hearing an auditory primary-response signal. On half the trials, a *high pitch* tone then occurs, and the subject must produce the *primary response* as quickly as possible. This consists of uttering the specified pair of words in the same order that they appeared on the screen. On the other half of the trials, a *low pitch* secondary-response signal is given instead, and the subject must produce a *secondary response,* which consists of uttering the same two words in the reverse order (e.g., PIT–PEAT). We define the latency of the primary and secondary responses to be the amount of time between the onset of the response signal and the initiation of the corresponding vocal activity. The response duration is defined to be the amount of time between the moments when the vocal activity is initiated and termi-

nated. The latency, duration, and accuracy of the responses are recorded as a function of various factors that characterize the response.

Theoretically, primary-response latencies mainly reflect the amount of time required to initiate a response after it has been programmed for execution (Meyer & Gordon, 1985; Gordon & Meyer, 1987). The rationale for this interpretation is that preparation of a motor program for a primary response should be completed well before the onset of the response signal, because subjects are encouraged to produce the primary response as quickly as possible, and they are given ample opportunity to prepare for it in advance. In contrast, secondary responses presumably are not fully prepared unless a secondary-response signal occurs. That the primary responses are more highly prepared than the secondary responses can be inferred from their latencies, which are typically much shorter (e.g., in Meyer and Gordon, 1985, the difference between primary and secondary response latencies was on the order of 200 ms). Thus, secondary-response latencies could reflect the amount of time taken to change the primary-response program into a program appropriate for the secondary response, as well as the time to initiate the response after the program is ready. This time could depend on the relationship between the primary and secondary responses, and on the intrinsic difficulty of preparing the secondary-response program.

The utility of the response-priming procedure has been demonstrated by previous studies (Meyer & Gordon, 1985; Gordon & Meyer, 1987) where the relationship between primary and secondary responses had greater effects on secondary responses than on primary responses. However, the complete preparation of primary responses, uninfluenced by their relation to secondary responses, may be an idealization that is not always achieved. So in evaluating results from the response-priming procedure, one must check whether patterns of primary-response performance are consistent with the patterns of secondary-response performance.

EXPERIMENT 1

The objective of Experiment 1 was simply to determine whether producing a vowel involves inhibiting the production of other similar vowels. Using the response-priming procedure, we tested subjects with a set of paired CVC response words that differed from each other only in their medial vowels, thereby maximizing the likelihood that subjects' performance would depend on vowel similarity. The effects of vowel similarity were assessed by comparing performance measures (response latency, duration, and error rate) for word pairs that contained similar vowels (e.g., PIT–PEAT, PUTT–POT) versus word pairs that contained dissimilar vowels (e.g., PIT–POT, PUTT–PEAT).

Method

Subjects. Eight undergraduate students from the volunteer subject pool of the Human Performance Center at the University of Michigan served as paid subjects. All subjects were native speakers of English. None reported any speech or hearing defects. We screened them in terms of two criteria: mean primary-response latency and error rate. For inclusion in subsequent analyses, each subject was required to have mean primary-response latencies less than 450 ms and overall error rates less than 10%. These selection criteria helped to ensure that subjects prepared highly for the primary responses yet were successful at producing secondary responses. Based on the criteria, seven of the eight subjects were included in our sample. They received a salary of $4 per session plus a bonus, averaging $2 per session, determined from the speed and accuracy of their performance.

Apparatus. Visual messages, response words, and feedback were displayed on a video terminal (HP 2621A). Warning tones and response tones were played over Senn-

heiser HMD 224 headphones. A microphone attached to the headphones transmitted subjects' vocal responses to an A-to-D converter, which digitized the responses at a sampling rate of 10 kHz and stored them for subsequent analyses.

Responses. We used two sets of monosyllabic CVC English words whose medial vowels were /i/, /I/, /ʌ/, and /ɑ/. The first set consisted of the words PEAK, PICK, PUCK, and POCK. The second set consisted of the words PEAT, PIT, PUTT, POT. Primary and secondary responses were constructed by pairing the words within each set in all possible combinations (a total of 24 pairs, 12 from each set). For example, Table 1 illustrates the set of 12 pairs involving the initial consonant P and final consonant T. The members of each word pair varied in their medial vowel, but had the same initial and final consonants (e.g., PUTT–POT). A complete replication of this design included each of the word pairs twice, once as a primary response and once as a secondary response. Secondary responses always consisted of the primary-response words in reverse order.

In subsequent data analyses, we defined a factor called *vowel relation* for the pairs of response words. This factor had two levels: similar and dissimilar. Word pairs involving /i/ combined with /I/ and word pairs involving /ʌ/ combined with /ɑ/ were included in the first level (e.g., PEAT–PIT and PUTT–POT). All other pairs (e.g., PEAT–PUTT and PIT–POT) were included in the second level.

Design and procedure. Each subject participated in five sessions on separate days within a one-week span. A session lasted about one hour. The first session provided instruction and practice on the response-priming procedure. The remaining four test sessions yielded the data reported here. There were five to six trial blocks per session, with 32 trials per block. Half the trials required a primary response and half required a secondary response. The responses were sampled from the two sets of 12 word pairs with equal probabilities. The word pairs occurred equally often as primary and secondary responses. The orders of the different pairs and of the two alternative trial types (i.e., primary and secondary) were randomized within each block.

The sequence of events during a trial was as follows. First, the subject saw a pair of response words displayed side-by-side in uppercase letters at the center of the video terminal. The correct primary response was to say the two words in the same order as they appeared left-to-right on the screen. The secondary response was to say the two words in reverse order. The response pair remained visible for 2 s, during which the subject prepared mentally to produce the primary response as quickly as possible. Next, the response pair disappeared, and the subject heard three warning signals (417-Hz tones). Each warning signal lasted 100 ms and was followed by a silent interval of 400 ms. After the last warning signal and silent interval, there was a response signal. The response signal was either a high-pitch (833 Hz) tone or a low-pitch (208 Hz) tone, each having a probability of one-half. The high-pitch tone signaled the subject to produce the primary response immediately.

TABLE 1
SOME PAIRS OF RESPONSE WORDS IN EXPERIMENT 1

	Second vowel			
First vowel	/i/	/I/	/ʌ/	/ɑ/
/i/	—	Peat–pit	Peat–putt	Peat–pot
/I/	Pit–peat	—	Pit–putt	Pit–pot
/ʌ/	Putt–peat	Putt–pit	—	Putt–pot
/ɑ/	Pot–peat	Pot–pit	Pot–putt	—

Alternatively, the low-pitch tone signaled the subject to produce the secondary response instead. Following the response signal, a computer recorded vocal input from the subject's microphone for a period of 1300 ms, calculating the latency and duration of the response. After the subject responded, the correct response appeared on the experimenter's terminal. The experimenter coded the subject's actual response as being *correct* if it matched the response on the screen, and *incorrect* otherwise. In the latter case, an error message appeared on the subject's terminal. The experimenter transcribed the subject's utterance if it was incorrect.

During the initial instruction and practice session, it was strongly emphasized that each primary response should be produced as quickly as possible after the high-pitch response signal. We also arranged the events on a trial to increase the likelihood that the subject would prepare to produce the primary response. The regularity of the three warning signals encouraged the subject to anticipate the primary-response signal and to be in a high state of preparation when it occurred (Meyer & Gordon, 1985; Sternberg et al., 1978). In addition, the subject's completion time (the response latency plus the duration of the utterance) for primary responses was measured. Whenever it exceeded a preset deadline, the message "Response too slow" was shown on the subject's terminal screen. This deadline, which averaged 605 ms across subjects, was initially set at 750 ms and lowered as performance improved.

A one-minute rest period was given after each trial block. During the rest period, the experimenter reminded the subject about the task instructions and provided summary feedback about the mean completion times and error rates for the primary and secondary responses. Point scores were calculated so that the subjects could see how their performance progressed throughout the experiment.

Each subject received bonus points based on response speed and accuracy. For every correct primary response whose completion time fell below the deadline, 25 points were awarded. No points were awarded if a primary response exceeded the deadline, and 10 points were deducted for every incorrect primary response, regardless of its latency. Also, one point was deducted for every 100 ms in the latency of a secondary response. If an incorrect secondary response occurred, an additional 20 points were deducted. Thus, while the point system emphasized very fast and accurate performance on the primary responses, it did not neglect good performance on the secondary responses. A bonus of one cent was paid per each 10 points earned by the subject.

On-line speech analysis. Latencies and durations were measured through a software voicekey calibrated to compensate for low-level background noise in the audio equipment. The voicekey employed several statistical parameters derived from successive temporal windows of the speech waveform, including average power, frequency of zero crossings, and slopes. Here we used algorithms similar to those of other latency-measurement routines described in the literature (Meyer & Gordon, 1985; Rabiner & Samber, 1975; Sherak, 1982; Sternberg et al., 1978). The routines in the voicekey allowed the onset of both voiced and voiceless speech segments to be detected sensitively, despite differences in the energy of various segments.

During each test session, the experimenter periodically inspected visual displays of the subject's speech waveforms on a graphics terminal. The displays included were vertical bars that marked the beginning and end of each utterance, as determined by the voicekey. This allowed the experimenter to monitor the overall performance of the voicekey, ensuring that no significant portions of the speech waveform were truncated by its algorithms and that the decision criteria were maintained at appropriate levels.

Data analysis. The first session of the experiment was used to instruct the subjects and to provide practice; data from it were not analyzed. The initial blocks of later sessions served as warm-up, and were likewise excluded from the analysis. The remaining blocks from these sessions yielded the results reported below. Latencies less than 100 ms were classified as errors and not analyzed, because they involved premature activation of the voicekey. Latencies greater than 1200 ms were not analyzed because they resulted from failures to activate the voicekey. The excluded latencies constituted less than 1% of the data.

We evaluated the data with repeated-measures analyses of variance that included response type (primary vs. secondary) and vowel relation (similar vs. dissimilar) as fixed factors and subjects as a random factor. We also performed two-way analyses of variance on the primary-response data and secondary-response data, separately. In these analyses, the response latencies, durations, and error rates served as dependent variables.

Results

The principal results of Experiment 1 concern the mean latencies and durations of correct primary and secondary responses and the error rates for each response type. These are given in Table 2 as a function of the vowel relation between syllables. Equivalent patterns of results occurred for the two sets of paired response words (i.e., ones with the final consonants /k/ and /t/, respectively), so we have averaged the latencies, durations, and error rates across them.

Response latencies. The latencies of correct primary and secondary responses averaged 324 and 533 ms, respectively. This difference of 209 ± 20 ms was statistically significant ($F(1,6) = 114.6, p < .001$), indicating that subjects were more prepared for the primary response than for the secondary response.

There were significant effects of vowel

TABLE 2
MEAN RESPONSE LATENCIES (ms), DURATIONS (ms), ERROR RATES (%), AND STANDARD ERRORS FROM EXPERIMENT 1

Responses	Vowel relation		
	Similar	Dissimilar	Std. error[a]
Primary			
Latency	329	319	2
Duration	287	279	2
Error rate	7.9	3.3	1.6
Secondary			
Latency	550	517	6
Duration	309	301	1
Error rate	7.1	4.6	1.0

[a] These are approximate standard errors appropriate for pairwise comparisons of the means in each row.

relation on both primary and secondary response latencies. Primary responses that contained similar vowels were 10 ± 2 ms slower on average than ones that contained dissimilar vowels, $F(1,6) = 17.1, p < .01$. Secondary responses that contained similar vowels were 33 ± 6 ms slower on average than ones that contained dissimilar vowels, $F(1,6) = 27.5, p < .01$. The response-type by vowel-relation interaction was also significant; vowel relation had a 23 ± 8 ms greater effect on secondary responses than on primary responses, $F(1,6) = 8.92$, $p < .05$.

Response durations. The pattern of mean response durations paralleled the pattern of response latencies in most respects. The durations of correct primary and secondary responses averaged 283 and 305 ms, respectively. This yielded a difference of 22 ms, $F(1,6) = 104.8, p < .001$. Significant vowel-relation effects again occurred for both primary and secondary responses. Primary-response durations were 8 ± 2 ms longer for similar vowels than for dissimilar vowels, $F(1,6) = 11.9, p < .05$. Secondary-response durations were 8 ± 1 ms longer for similar vowels than for dissimilar vowels, $F(1,6) = 30.1, p < .01$. Unlike for the response latencies, however, the vowel-relation effect did not interact significantly

with the effect of response type (primary vs. secondary; $F(1,6) < 1.$).

Error rates. Errors occurred with roughly equal rates in primary and secondary responses, 5.6 vs. 5.9%, respectively, $F(1,6) < 1$. The vowel relation reliably affected error rates. Primary responses involving similar vowels contained more errors than those involving dissimilar vowels, 7.9 vs. 3.3%, $F(1,6) = 9.53$, $p < .05$. The same pattern occurred for secondary responses, 7.1 vs. 4.6%, $F(1,6) = 7.23$, $p < .05$. This suggests that the vowel-relation effect on response latency did not stem merely from a speed–accuracy trade-off. Again there was no significant response-type by vowel-relation interaction [$F(1,6) = 3.17$, $p > .10$].

Unidimensional scaling. To further clarify the correspondence between vowel similarity and the observed response latencies, we performed linear unidimensional scaling on the mean secondary-response latencies for the 12 vowel pairs included in Experiment 1 (Kruskal, 1964; Kruskal & Wish, 1978).[2] The obtained solution (Fig. 2) had a low stress value, indicating reasonably good fit. Of particular importance here is the fact that the proximity of the vowels in the derived scaling closely resembles their proximity in the cardinal vowel space (Fig. 1). This indicates that response latencies depended on the exact degree of similarity between the vowels of the responses.

Discussion

Experiment 1 showed that vowel similar-

[2] The original data matrix paralleled Table 1. To obtain a triangular (sub-diagonal) data matrix, we averaged each entry below the diagonal (top left to bottom right) with the corresponding entry above the diagonal (e.g., the mean latencies for PEAT–POT and POT–PEAT were averaged, and so on). Unidimensional scaling of these data is appropriate in that the vowels selected for our experiments essentially fall along a line embedded in the cardinal–vowel space (Fig. 1). Also, because the experiments included only four different vowels, there were not enough degrees of freedom here to conduct truly meaningful multidimensional scalings (e.g., one with two dimensions, analogous to the full cardinal-vowel space).
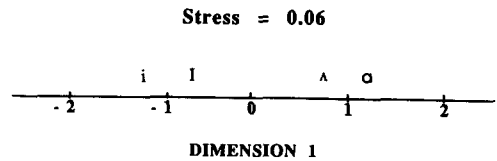
Stress = 0.06

DIMENSION 1

FIG. 2. One-dimensional scaling solution for representing the four vowels of Experiment 1 based on mean secondary-response latencies. (The obtained distances are similar to the theoretical representation of these vowels in terms of Jones' cardinal vowel space; cf. Fig. 1.)

ity increases response latency, duration, and error rates in the production of CVC syllable pairs. The magnitude of the inhibitory vowel-similarity effect correlates strongly with the proximity of the vowels in the cardinal-vowel space (cf. Figs. 1 and 2).

This finding rules out some simple possibilities for how vowel similarity influences speech production. One might have expected that vowel similarity would facilitate, not inhibit, the production of syllable pairs. Such an effect could arise peripherally because the tongue must travel a shorter distance between similar vowels like /i/–/I/ or /ʌ/–/ɑ/ than between dissimilar vowels like /i/–/ʌ/ or /I/–/ɑ/. Assuming that movement time increases directly with physical movement distance, as motor-control researchers have typically found (Fitts, 1954; Meyer, Abrams, Kornblum, Wright, & Smith, 1988; Meyer, Smith, Kornblum, Abrams, & Wright, 1990; Schmidt, Zelaznik, Hawkins, Frank, & Quinn, 1979), faster and more accurate performance should occur for similar vowels than for dissimilar vowels. However, just the opposite was found here. Similar vowels yielded slower and less accurate performance; the effect on response latency was especially pronounced in producing secondary vocal responses.

The results of Experiment 1 complement those of Meyer and Gordon (1985), who found that the motor programming of consonants during speech production involves inhibiting the production of other similar consonants (specifically ones that have matched voicing or place-of-articulation

features). Although vowels are produced differently than consonants in certain fundamental ways, it appears that similarity among vowels may inhibit motor programming much like consonant similarity does. This is consistent with a theoretical account based on central inhibitory processes that utilize a network of nodes to tune the output of subsyllabic speech units (Meyer & Gordon, 1985).

Of course, we do not mean to imply that more-peripheral processes are entirely immune to the inhibitory effects of vowel similarity. Our results did reveal modest but reliable effects on primary responses. The vowel-similarity effect on primary-response latencies could indicate that subjects neglected to program the primary responses fully before the latency interval began, even though they had the opportunity for full programming. However, another conceivable possibility is that vowel similarity also affects peripheral response execution at least somewhat, after motor programming has been completed. This could explain, for example, why vowel similarity increased not only the latencies but also the durations and error rates in the primary responses. (For further consideration of the last point, see General Discussion.)

THE VOWEL-SIMILARITY EFFECT AND SYLLABLE STRUCTURE

We have proceeded thus far as if the inhibitory effect of vowel similarity on speech production occurs at the level of internal representations (e.g., network nodes) for vowels *per se*. However, the vowels studied here, like those in natural speech, belong to various superordinate speech units, such as consonant–vowel (CV) clusters, vowel–consonant (VC) clusters, and whole syllables. The vowel-similarity effect could, perhaps, occur at one (or more) of these other levels. In particular, this possibility may be pursued with respect to four alternative hypotheses about the relationship between syllable structure and the vowel-similarity effect.

VC-cluster hypothesis. First, the vowel-

similarity effect may hinge at least partly on the similarity between the final VC clusters of CVC syllables. The impetus for this VC-cluster hypothesis comes from linguistic analyses of syllable structure, which postulate that syllables have a hierarchical internal structure with two major constituents, the *onset* and the *rime* (e.g., Halle & Vergnaud, 1980; Selkirk, 1982). The onset of a syllable corresponds to the initial consonant cluster, and the rime corresponds to the final vowel-consonant cluster, which in turn contains a vowel *peak* and consonant-cluster *coda*. There are two sorts of linguistic evidence for this decomposition (Selkirk, 1982). First, the range of phonotactic constraints seems to be well captured by the breakdown of the syllable into onset and rime, as well as the subsequent breakdown of the rime into its peak and coda. Second, formulation of some phonological rules, such as the heavy/light syllable distinction used in some prosodic analyses, is simplified if the rules can operate on the internal C–VC structure of a syllable.

Complementary psychological evidence for the VC-cluster hypothesis has been obtained by MacKay (1972, 1974), who presented a syllable-recoding theory wherein CVC syllables are recoded as C + VC. He analyzed speech errors such as the one in which the utterance "Don't shell" is produced as a blend of two synonyms: *shout + yell → shell*. His results showed that breaks within syllables, as in this example, tend to precede the vowel rather than follow it, suggesting that the final VC cluster is more cohesive than the initial CV cluster. Similarly, Treiman (1983) found that adults can more easily induce phonological rules that keep the onset and rime intact than rules that divide these units.

CV-cluster hypothesis. A second possibility is that the vowel-similarity effect hinges instead on the similarity between the CV clusters of CVC response syllables. Some evidence for this CV-cluster hypothesis comes from the work of Kozhevnikov and Chistovich (1965). Based on recordings of articulatory movements by Russian sub-

jects, they concluded that gestures associated with the first consonant in a consonant cluster are influenced by features of the subsequent vowel. This led them to suggest that the range of forward coarticulation mirrors the size of the programming unit in speech, and that articulatory movements are therefore organized in consonant-vowel groups such as CV, CCV, and CCCV (see also Benguerel & Cowan, 1974; Kent, 1976a; Kent & Minifie, 1977; MacNeilage & DeClerk, 1969; Perkell, 1969; Stetson, 1951). Kent (1976b) articulated a similar hypothesis according to which syllable structures are recoded into series of CV combinations with appropriate reductions of C and V segments. Thus, for example, a CVC syllable may be recoded as CV + CV, where the second CV consists of a consonant followed by a null unit (i.e., C + #).

Whole-syllable and independent-vowel hypotheses. Another possibility is that the vowel-similarity effect depends on the integrity of the whole CVC syllable. According to this hypothesis, the effect of vowel similarity could be modulated by an interaction of the vowel with both the initial and final consonants of a syllable. Any change in the consonant frame would then reduce the similarity between syllables and attenuate the similarity effect. Thus, the whole-syllable hypothesis provides the most extreme contrast with our initial characterization, which was that the vowel-similarity effect occurs entirely at the level of independent vowel representations. If the independent-vowel hypothesis were the case, then changes in the consonant frame should not influence the vowel-similarity effect.

Plan of subsequent experiments. To test these alternative hypotheses and clarify the vowel-similarity effect revealed by Experiment 1, we have conducted four additional experiments. They involved manipulating not only the vowels but also the initial and final consonants of CVC syllables. Assuming that a particular superordinate unit (e.g., CV cluster, VC cluster, or whole syllable) mediates the vowel-similarity effect,

this type of manipulation should attenuate or eliminate it, depending on which one of the preceding hypotheses is valid. In Experiments 2 and 3, we varied the *initial* consonants of the paired response syllables while keeping their final consonants unchanged (as in PICK–TUCK). If the VC-cluster hypothesis is valid, then we would expect results similar to those of Experiment 1, whereas this should not occur under the CV-cluster and whole-syllable hypotheses. In Experiments 4 and 5, we varied the *final* consonants of the paired response syllables but not the initial consonants (as in PIT–PUCK). If the VC-cluster hypothesis is valid, then the vowel-similarity effect might be expected to decrease. However, under the CV-cluster hypothesis, the opposite prediction would apply, namely, the vowel-similarity effect should recur as in Experiment 1. Of course, localization of the effect at the level of vowels *per se*, which are distinct from superordinate syllables and clusters, would imply that the vowel-similarity effect should persist across all our subsequent experiments.

EXPERIMENT 2

The specific goal of Experiment 2 was to determine whether the inhibitory vowel-similarity effect found in Experiment 1 still obtains when the initial consonants of paired CVC syllables are different rather than identical. Subjects had to produce monosyllabic word pairs like BEAT–PIT and BEAT–POT that contained either similar vowels (e.g., /i/ and /I/) or dissimilar vowels (e.g., /i/ and /a/), and that contained different initial consonants. The two initial consonants, /b/ and /p/, differed in the phonetic feature of voicing (/b/ is voiced, whereas /p/ is voiceless). The final consonant (/t/) was always the same for both syllables of a pair. Consequently, in proceeding from one syllable of a pair to the next, the VC clusters remained intact, whereas the CV clusters did not.

As explained previously, this manipulation provides a partial test of hypotheses

concerning the locus of the vowel-similarity effect. Under the VC-cluster and independent-vowel hypotheses, slower and less accurate secondary responses should occur in producing syllables that contain similar vowels and identical final consonants, even if the initial consonants differ. Such an effect would not be anticipated under the CV-cluster and whole-syllable hypotheses, because the present design disrupts these units.[3]

Our design also included monosyllabic word pairs that contained the same medial vowel (e.g., BIT-PIT). The inclusion of such identical-vowel pairs was possible here, unlike in Experiment 1, because primary and secondary responses in Experi-

[3] It is worth noting, however, that the VC clusters used here would not be characterized as rimes under all linguistic analyses. For the present application of the response-priming procedure, we sought to minimize the distance between our similar vowels and to maximize the distance between our dissimilar vowels in the vowel space (Fig. 1). Doing so required using long-short vowel pairs that may have induced effects at the level of either rimes or complex syllabic nuclei. In particular, Clements and Keyser (1983) have analyzed syllable structure into a center, the "nucleus," and two margins, the "onset" and "coda." According to them, the nucleus consists of a long vowel or a short vowel plus the subsequent consonant. The final consonants of syllables containing /i/, a long vowel, would therefore be characterized as the coda. On the other hand, the final consonants of syllables containing /I/, a short vowel, would be characterized as part of the nucleus. Thus, the present manipulation of VC clusters cannot distinguish among some competing linguistic analyses of syllable structure (Clements & Keyser, 1983; cf. Selkirk, 1982). In principle, the methodology developed here could, nevertheless, be used to such an end.

ment 2 could be adequately differentiated through their initial consonants, regardless of whether they contained identical, similar, or dissimilar final VC clusters. By including pairs whose members contained the same vowel, we established a baseline from which performance on the other types of pairs could be further assessed.

Method

Subjects. Six new individuals who satisfied our speed and accuracy criteria as described in Experiment 1 were recruited from the same subject pool. They were again paid on the basis of their performance.

Responses. We constructed the primary and secondary responses from two sets of monosyllabic CVC English words and pseudowords: (1) PEAT, PIT, PUTT, POT; and (2) BEAT, BIT, BUT, BOT. Each response contained one member from the first set and one member from the second set. The 16 possible combinations presented in each of two orders provided a total of 32 distinct response pairs for use in the experiment. Table 3 illustrates one set of 16 pairs. One quarter of them involved the same vowels (e.g., PIT-BIT, BUT-PUTT); one quarter involved similar vowels (e.g., BIT-PEAT, POT-BUT); and one half involved dissimilar vowels (e.g., PIT-BOT, BUT-PEAT).

Apparatus, procedure, and design. The apparatus and procedure were the same as in Experiment 1. Except for changes in the selection of syllable pairs, the design was also the same. Subjects were tested during

TABLE 3
SOME PAIRS OF RESPONSE WORDS IN EXPERIMENT 2

First vowel	Second vowel			
	/i/	/I/	/ʌ/	/ɑ/
/i/	Peat–beat	Peat–bit	Peat–but	Peat–bot
/I/	Pit–beat	Pit–bit	Pit–but	Pit–bot
/ʌ/	Putt–beat	Putt–bit	Putt–but	Putt–bot
/ɑ/	Pot–beat	Pot–bit	Pot–but	Pot–bot

four sessions with 6 to 7 blocks of trials per session.

Results

Table 4 summarizes the principal results from this experiment as a function of response type and vowel relation.

Response latencies. Latencies of correct primary and secondary responses averaged 284 and 558 ms, respectively, $F(1,5) = 596.3, p < .001$. No significant effects of the vowel relation on mean primary-response latencies were obtained, $F(2,10) < 1$, but there were such effects on mean secondary-response latencies, $F(2,10) = 7.67, p < .01$. The interaction between vowel relation and response type was reliable, $F(2,10) = 5.27$, $p < .05$.

Post hoc comparisons, using the Newman–Keuls method, revealed that the latencies of secondary responses with similar vowels exceeded the latencies of those with dissimilar vowels by 18 ± 9 ms, $q(2,10) = 2.95, p < .05$. Secondary responses that involved identical vowels were 16 ± 9 ms faster than those that involved dissimilar vowels, $q(2,10) = 2.58, p < .05$.

Response durations. The mean duration of primary responses was 24 ± 6 ms shorter than the mean duration of secondary responses, 283 versus 307 ms, $F(1,5) = 15.8$, $p < .05$. Vowel relation had no significant effects on either primary responses, $F(2,10)$

$= 3.05, p > .05$, or secondary responses, $F(2,10) < 1$. Nor was the interaction between vowel relation and response type significant, $F(2,10) < 1$.

Error rates. The mean error rate was lower on primary responses than on secondary responses, 5.2 vs. 9.6%, $F(1,5) = 13.8$, $p < .05$. Responses that contained similar vowels exhibited the highest error rates, and responses that contained identical vowels exhibited the lowest. However, the effect of vowel relation did not reach a statistically significant level in either primary responses, $F(2,10) = 2.96, p > .05$, or secondary responses, $F(2,10) < 1$. Vowel relation and response type did not interact significantly, $F(2,10) < 1$.

Discussion

Experiment 2 produced two major results. First, it replicated the inhibitory vowel-similarity effect on secondary responses. As before, performance was slower and somewhat less accurate when the syllables in a pair contained similar vowels and VC clusters than when these subunits were dissimilar. This occurred even though the initial consonants of the paired syllables always differed from each other, disrupting the integrity of the CV clusters from syllable to syllable, while preserving the integrity of the VC clusters. In addition, we found a complementary facili-

TABLE 4
MEAN RESPONSE LATENCIES (ms), DURATIONS (ms), ERROR RATES (%), AND STANDARD ERRORS FROM EXPERIMENT 2

	Vowel relation			
Responses	Identical	Similar	Dissimilar	Std. error[a]
Primary				
Latency	281	288	283	6
Duration	280	286	284	2
Error rate	3.8	7.2	4.8	1.4
Secondary				
Latency	541	575	557	9
Duration	307	308	307	2
Error rate	8.4	11.0	9.6	2.0

[a] These are approximate standard errors appropriate for pairwise comparisons of the means in each row.

tative effect; when the vowels and VC clusters within a syllable pair were the same, rather than merely being "similar" to each other, performance was fastest and most accurate. This is what one would expect from linguistic and psycholinguistic theories of syllable structure that decompose CVC syllables into onset (C) and rime (VC) subunits (Halle & Vergnaud, 1980; Mac-Kay, 1972, 1974; Selkirk, 1982; Treiman, 1983). Alternative hypotheses based on CV clusters and whole syllables are not supported by our findings.

The locus of the facilitative vowel-identity effect remains to be determined. One possibility is that it occurs during the same phase of speech production as the inhibitory vowel-similarity effect does. Perhaps both effects are localized in a single process that mediates the programming of vowels and VC clusters for output. For example, residual activation of a particular vowel following the production of a syllable that contains it could carry over to facilitate the production of another subsequent syllable that also contains this vowel. Alternatively, the vowel-identity effect may occur during some later phase of processing, perhaps even extending to a peripheral execution (e.g., overt articulatory gesture) phase. Moving from the articulatory positions for one consonant to those for the next may be easier when the global shape of the vocal tract, as dictated by the identities of medial vowels, remains essentially constant.

EXPERIMENT 3

The goal of Experiment 3 was to replicate the results of Experiment 2 using a somewhat different set of primary and secondary responses. We again wanted to determine whether the vowel-similarity and vowel-identity effects would occur when subjects rapidly produce paired CVC syllables whose initial consonants differ while their final consonants are the same. However, we changed the particular way in which the syllables' initial consonants differed from each other. Rather than varying with re-

spect to the phonetic feature of voicing (i.e., voiced versus voiceless), they varied with respect to place of articulation. For each syllable pair, one of the initial consonants was *labial* (made with the lips, as in /p/) and the other was alveolar (made with the tongue tip and alveolar ridge, as in /t/). This yielded pairs of syllables such as PEAK–TICK and PEAK–TOCK, whose initial consonants and vowel relation (identical, similar, or dissimilar) both varied, but whose final consonant always stayed the same. As before, we reasoned that if the vowel-similarity and vowel-identity effects take place at the level of vowels or VC-clusters, then this manipulation should maintain these effects, because it does not disrupt the cross-syllable integrity of the key speech subunits. On the other hand, if the effects are localized at the level of CV clusters or whole syllables, then the effects should be attenuated or eliminated here, because the present manipulation does disrupt the integrity of these units.

Such disruption would, on *a priori* grounds, seem more likely to occur in Experiment 3 than in Experiment 2. This is because of the change in place of articulation between the different consonants. For example, consider the pair of syllables PEAK and TICK, whose initial consonants have labial and alveolar places of articulation, respectively. To produce the CV cluster /tI/ of TICK, the tongue's position in the oral cavity must change markedly from where it is while producing the CV cluster /pi/ of PEAK. The required change makes the successive CV clusters (i.e., /pi/ and /tI/) of these syllables articulatorily dissimilar, even though they contain similar vowels. So one might expect a substantial reduction in cross-syllable integrity for this case. By contrast, the reduction achieved through changing the voicing feature of the initial consonants, going from voiced to voiceless (as in PEAT–BIT), seems relatively slight. Changes of voicing do not markedly influence required sequences of tongue positions. This could have limited

the power of Experiment 2, which manipulated the voicing feature, relative to what Experiment 3 provides with its changes in place of articulation.

Method

Subjects. Eight new subjects who satisfied our response speed and accuracy criteria were recruited from the same pool used for Experiments 1 and 2.

Responses. Syllable pairs used for primary and secondary responses were constructed from the following two sets of monosyllabic English words and pseudowords: (1) TEAK, TICK, TUCK, TOCK; (2) PEAK, PICK, PUCK, POCK. As in Experiment 2, one quarter of the pairs incorporated the same vowel, one quarter incorporated similar vowels, and the remaining half incorporated dissimilar vowels. The consonant and vowel relations within the pairs were analogous to those of Experiment 2. However, the initial consonants of the syllables in each pair had different places of articulation (labial /p/ vs. alveolar /t/) instead of different voicing. Also, the final consonant was /k/ instead of /t/, because /t/ now served as an initial consonant. This helped further generalize the results of Experiments 1 and 2.

Apparatus, design, and procedure. The apparatus, design, and procedure were the same as in Experiment 2.

Results

The results from Experiment 3 are summarized in Table 5.

Response latencies. The latencies of correct primary and secondary responses averaged 277 and 441 ms, respectively. This yielded a difference of 164 ± 21 ms, $F(1,7) = 62.0$, $p < .001$. Vowel relation did not affect primary-response latencies, $F(2,14) = 1.02$, $p > .10$, but it strongly affected secondary-response latencies, $F(2,14) = 11.6$, $p < .01$. The interaction between vowel relation and response type was reliable, $F(2,14) = 5.55$, $p < .05$.

The vowel-similarity and vowel-identity effects were tested further in two separate orthogonal contrasts. The first contrast compared performance for syllable pairs with similar vowels versus performance for pairs with dissimilar vowels. The second contrast compared performance for pairs with identical vowels versus the average of performance for pairs with similar and dissimilar vowels. Planned comparisons were warranted here because the results from Experiments 1 and 2 strongly predicted the existence of these effects in the present data. We found that secondary responses involving similar vowels were 24 ± 8 ms slower on the average than secondary responses involving dissimilar vowels, $F(1,14) = 9.86$, $p < .01$. Secondary responses involving identical vowels were 25

TABLE 5
MEAN RESPONSE LATENCIES (ms), DURATIONS (ms), ERROR RATES (%), AND STANDARD ERRORS FROM
EXPERIMENT 3

Responses	Vowel relation			
	Identical	Similar	Dissimilar	Std. error[a]
Primary				
Latency	277	280	274	4
Duration	307	328	318	5
Error rate	4.4	9.0	4.1	1.6
Secondary				
Latency	424	461	437	8
Duration	330	351	341	5
Error rate	7.0	13.6	6.7	2.1

[a] These are approximate standard errors appropriate for pairwise comparisons of the means in each row.

± 7 ms faster on average than secondary responses involving nonidentical vowels, $F(1,14) = 13.02, p < .01$.

Response durations. The mean duration of primary responses was 23 ± 6 ms shorter than the mean duration of secondary responses, 318 vs. 341 ms, $F(1,7) = 16.97, p < .01$. Vowel relation had completely analogous effects on primary and secondary responses, $Fs(2,14) = 10.33, 9.56, p < .01$. The interaction between vowel relation and response type was not reliable, $F(2,14) < 1$.

Planned comparisons revealed that both primary and secondary responses involving similar vowels had 10 ± 5 ms longer mean durations than did those involving dissimilar vowels, $Fs(1,14) = 4.08, 4.45, .05 < p < .1$. Responses involving identical vowels had 11 ± 5 ms shorter durations than did those involving dissimilar vowels, $Fs(1,14) = 4.94, 5.21, p < .05$.

Error rates. The pattern of errors paralleled the latency and duration data. Subjects tended to make fewer errors on primary responses than on secondary responses (5.8 vs. 9.1%), but the difference did not reach statistical significance, $F(1,7) = 2.85, p > .1$. Vowel relation affected error rates in both primary and secondary responses, $F(2,14) = 5.82, p < .05; F(2,14) = 7.24, p < .01$. Again, the interaction between vowel relation and response type was not reliable, $F(2,14) < 1$.

Planned comparisons revealed that similar vowels yielded more errors than did dissimilar vowels for each response type, $Fs(1,14) = 8.96, 10.9, p < .01$. The rate of errors was also a bit greater on identical vowels than on dissimilar vowels, but this effect did not reach significance for either response type, $Fs(1,14) < 1$.

Discussion

The results of Experiment 3 replicate and extend those found in Experiments 1 and 2. Using our response-priming procedure with CVC syllable pairs, we again obtained marked inhibitory effects of vowel similarity on secondary responses. As before,

there was also a facilitative vowel-identity effect. These effects persisted even though the initial consonants of the paired syllables differed with respect to place of articulation, not voicing as in Experiment 2. Experiment 3 therefore lends additional credibility to the conclusion that programming processes based on representations of the vowels or VC-clusters mediate the production of speech output. Given that changes of the initial consonants across the paired syllables did not substantially attenuate the vowel-similarity or vowel-identity effects, one may further question the validity of the alternative CV-cluster and whole-syllable hypotheses.

EXPERIMENT 4

In Experiment 4, subjects had to produce primary and secondary vocal responses consisting of monosyllabic (CVC) word pairs like PEAT–PICK and PEAT–POCK. The initial consonant remained the same within each pair, while the medial vowels and final consonants varied systematically. According to the CV-cluster and independent-vowel hypotheses, such changes should leave the inhibitory vowel-similarity effect and the facilitative vowel-identity effect unattenuated, because the integrity of the key subunits is maintained across the paired words. The VC-cluster hypothesis predicts, on the other hand, that some attenuation should take place here, because the integrity of the VC clusters is markedly disrupted.

Method

Subjects. Six new subjects who satisfied our response speed and accuracy criteria were recruited from the same pool as for the previous experiments.

Responses. Syllable pairs used for primary and secondary responses were constructed from the following two sets of monosyllabic words and pseudowords: (1) PEAT, PIT, PUTT, POT; (2) PEAK, PICK, PUCK, POCK. Each member of the first set was paired with each member of the

second set and vice versa. One quarter of the resulting pairs contained the same vowel (e.g., PUTT–PUCK), one quarter contained similar vowels (e.g., PIT–PEAK), and the remaining half contained dissimilar vowels (e.g., POT–PEAK). The initial consonant (/p/) of the pairs remained the same throughout, whereas the final consonants (/t/ and /k/) always differed within pairs.

Apparatus, design, and procedure. The apparatus, design, and procedure were identical to those of Experiments 2 and 3, except that the experiment lasted five sessions.

Results

The principal results of Experiment 4 are summarized in Table 6.

Response latencies. Correct primary and secondary responses had latencies that averaged 252 and 446 ms, respectively. This yielded a difference of 194 ± 34 ms, $F(1,5) = 33.0$, $p < .01$. Vowel relation did not affect mean primary-response latencies significantly, $F(2,10) = 3.42$, $p > .05$, but it had a strong effect on secondary-response latencies, $F(2,10) = 12.7$, $p < .01$. There was a significant interaction between vowel relation and response type, $F(2,10) = 11.5$, $p < .01$. Planned comparisons revealed that secondary responses with identical vowels were 49 ± 12 ms faster on average than

those with dissimilar vowels, $F(1,10) = 16.7$, $p < .01$. However, secondary responses with similar vowels differed by only 2 ± 12 ms on average from those with dissimilar vowels, $F(1,10) < 1$.

Response durations. The mean durations of primary responses were 18 ± 6 ms shorter than those of secondary responses (248 vs. 266 ms); $F(1,5) = 8.45$, $p < .05$. Vowel relation had no significant effect on either primary or secondary response durations, $F(2,10) < 1$, nor did vowel relation interact significantly with response type, $F(2,10) = 2.16$, $p > .10$.

Error rates. There was no significant effect of response type (primary vs. secondary) on error rates, 4.6 vs. 6.6%, $F(1,5) = 3.40$. Vowel relation affected error rates on primary responses, $F(2,10) = 4.11$, $p < .05$, and secondary responses, $F(2,10) = 16.1$, $p < .01$. On primary responses, similar vowels caused more errors than did dissimilar vowels, $F(1,10) = 7.47$, $p < .05$. On secondary responses, identical vowels caused fewer errors than did dissimilar vowels, $F(1,10) = 20.8$, $p < .01$.

Discussion

Experiment 4 yielded results that are striking in several respects. For the first time, we did not obtain a large inhibitory vowel-similarity effect on our performance measures. The effect's absence was especially salient in secondary responses,

TABLE 6
MEAN RESPONSE LATENCIES (ms), DURATIONS (ms), ERROR RATES (%), AND STANDARD ERRORS FROM
EXPERIMENT 4

Responses	Vowel similarity			
	Identical	Similar	Dissimilar	Std. error[a]
Primary				
Latency	248	257	253	3
Duration	247	249	247	2
Error rate	3.9	6.9	2.8	1.5
Secondary				
Latency	413	464	462	12
Duration	268	267	265	3
Error rate	3.5	8.4	7.6	0.9

[a] These are approximate standard errors appropriate for pairwise comparisons of the means in each row.

where syllable pairs with similar and dissimilar vowels had latencies (and durations) whose means differed by only 2 ms. Changing the final consonants across syllables markedly reduced the impact of vowel similarity compared to what was obtained previously. In Experiments 1, 2, and 3, the effect's size was respectively 33, 18, and 24 ms on the average. A two-way analysis of variance of the vowel-similarity effect, with experiment as a between-subjects factor, revealed a significant difference in the effect's size, $F(3,23) = 3.84, p < .05$. A post hoc comparison confirmed that the size of the effect was significantly smaller in Experiment 4 than in Experiments 1, 2, and 3, $F(1,23) = 8.97, p < .01$.

This outcome reinforces the VC-cluster hypothesis concerning the locus of the vowel-similarity effect. Unlike in previous experiments, here the VC clusters of syllable pairs were relatively dissimilar, because their final consonants always changed from one syllable to the next. It does not appear that the vowel-similarity effect results from a level of representation involving initial CV clusters. Had the effect's locus been there, then some inhibition should have remained during Experiment 4, given that the present manipulation did not markedly disrupt the integrity of these clusters across the syllables of each pair. Nor does it appear that the vowel-similarity effect stems from a level of representation involving vowels as separate subunits. Had the locus been there, then some inhibition should likewise have remained.

Although the inhibitory vowel-similarity effect disappeared during Experiment 4, the facilitative vowel-identity effect did not. Subjects were still fastest and most accurate in producing secondary responses whose paired syllables contained the same vowel. As a result, there appears to be a dissociation between the two types of effect. This supports the possibility that vowel similarity and vowel identity have their effects at different loci in the speech-production system. On the basis of present

evidence, one must attribute the effect of vowel identity to some level of processing other than VC-clusters, since disrupting the integrity of these clusters did not eliminate the effect. Perhaps vowel identity instead has its effect at either a separate level of independent vowel representations or some other place more peripheral than this.

EXPERIMENT 5

The purpose of Experiment 5 was to replicate our finding that, in producing pairs of CVC syllables, an inhibitory vowel-similarity effect occurs when the syllables have the same final consonant but not when they have different final consonants. We wanted to verify this important pattern of results using a within-subjects design, where subjects would experience both same-final-consonant syllable pairs and different-final-consonant syllable pairs. If the vowel-similarity effect occurs at the level of VC clusters, then the results obtained with the first of these pair types (same final consonant) should parallel those from Experiments 1 through 3 (Tables 2, 4, and 5), while the results obtained with the second of these pair types (different final consonants) should parallel those from Experiment 4 (Table 6).

Method

Subjects. Nine new subjects who satisfied our speed and accuracy criteria were recruited from the same pool as in previous experiments.

Responses. Syllable pairs used for primary and secondary responses were assigned to two distinct conditions: *same final consonant*, and *different final consonants*. In the same-final-consonant condition, we used a set of 12 monosyllabic word pairs from Experiment 1, whose initial and final consonants were /p/ and /t/, respectively (e.g., POT–PUTT). In the different-final-consonants condition, we used a set of 12 monosyllabic word pairs from Experiment 4, whose initial consonant was always /p/, and whose final consonants included /t/

and /k/ as part of each pair (e.g., PEAT–PICK). Vowel similarity varied systematically under each condition.

Apparatus, design, and procedure. The apparatus and procedure were the same as before. There were five experimental sessions. Each session included three trial blocks with word pairs from the same-final-consonant condition and three trial blocks with word pairs from the different-final-consonants condition. The blocks for these two conditions alternated. The results obtained under each condition are summarized in Table 7.

Results: Same-Final-Consonant Condition

Response latencies. In the same-final-consonant condition, the primary-response latencies averaged 313 ms. The corresponding secondary-response latencies averaged 510 ms, 197 ± 23 ms longer than the primary-response latencies, $F(1,8) = 70.2, p < .001$. There was no significant effect of vowel relation on mean primary-response latencies when the syllables had

the same final consonant, $F(1,8) = 1.20, p > .10$. Vowel relation did affect mean secondary-response latencies significantly; secondary responses that contained similar vowels were 11 ± 2 ms slower than those that contained dissimilar vowels, $F(1,8) = 21.9, p < .01$. However, these results must be interpreted somewhat cautiously, because the vowel-relation effect was not significantly larger for the secondary responses than for the primary responses; $F(1,8) = 2.45, p > .10$.

Response durations. The mean duration of correct primary responses in the same-final-consonant condition was 26 ± 6 ms shorter than the mean duration of secondary-responses, 265 vs. 291 ms, $F(1,8) = 19.1, p < .01$. Vowel relation had no significant effect on the durations of either primary or secondary responses, $F(1,8) = 3.60, p > .05$, and $F(1,8) < 1$, respectively. There was no significant interaction between vowel relation and response type either, $F(1,8) < 1$.

Error rates. There was no statistically

TABLE 7
MEAN RESPONSE LATENCIES (ms), DURATIONS (ms), ERROR RATES (%), AND STANDARD ERRORS FROM
EXPERIMENT 5

Responses	Vowel similarity		
	Similar	Dissimilar	Std. error[a]
Same final consonant			
Primary			
Latency	316	312	4
Duration	269	263	3
Error rate	7.0	1.8	1.0
Secondary			
Latency	517	506	2
Duration	292	290	3
Error rate	6.6	5.0	1.4
Different final consonants			
Primary			
Latency	327	314	4
Duration	267	264	2
Error rate	3.2	2.0	0.6
Secondary			
Latency	523	524	5
Duration	284	288	2
Error rate	6.2	7.2	1.3

[a] These are approximate standard errors appropriate for pairwise comparisons of the means in each row.

significant effect of response type (primary vs. secondary) on error rate, although more errors occurred on secondary responses than on primary responses, 5.7 vs. 4.4%, $F(1,8) = 2.21$. On primary responses, similar vowels yielded significantly more errors than did dissimilar vowels, 7.0 vs. 1.8%, $F(1,8) = 28.1, p < .01$. On secondary responses, similar vowels tended to yield more errors than did dissimilar vowels, but the difference was not significant, 6.6 vs. 5.0%, $F(1,8) = 1.31$. The interaction between vowel relation and response type was significant, $F(1,8) = 6.61, p < .05$.

Results:
Different-Final-Consonants Condition

Response latencies. In the different-final-consonants condition, correct primary-response latencies averaged 320 ms. The corresponding secondary-response latencies averaged 524 ms, 204 ± 24 ms greater than the mean primary-response latencies, $F(1,8) = 69.8, p < .001$. Primary responses that contained similar vowels were 13 ± 4 ms slower than those that contained dissimilar vowels, $F(1,8) = 11.4, p < .01$. Vowel relation had virtually no effect on the mean secondary-response latencies, $F(1,8) < 1$. The difference between the effects of vowel relation on primary and secondary responses was statistically significant, $F(1,8) = 7.84, p < .05$.

Response durations. The mean duration of correct primary responses was 22 ± 6 ms shorter than the mean duration of correct secondary responses, $F(1,8) = 15.0, p < .01$. Vowel relation did not affect the mean durations of either primary or secondary responses significantly, $Fs(1,8) = 2.29$ and 3.93, $p > .05$. However, there was a small but reliable interaction between vowel relation and response type, $F(1,8) = 6.13, p < .05$; similar vowels tended to increase the durations of primary responses slightly (3 ± 2 ms) while decreasing the durations of secondary responses slightly (4 ± 2 ms).

Error rates. There was a tendency to make fewer errors on primary responses than on secondary responses, 2.6 vs. 6.7%,

$F(1,8) = 4.59, .05 < p < .1$. Vowel relation did not affect error rate significantly on either primary or secondary responses, $F(1,8) = 3.63, p > .05$, and $F(1,8) < 1$.

Discussion

The results of Experiment 5 largely confirm those from our previous four experiments. As in Experiments 1 through 3, vowel similarity had an inhibitory effect on performance for paired CVC syllables whose final consonants were the same. This effect was exhibited in secondary-response latencies, a measure that has consistently revealed effects of similarity relationships within secondary responses (Experiments 1 through 3; Meyer & Gordon, 1985) and between primary and secondary responses (Gordon & Meyer, 1987). An inhibitory vowel-similarity effect was also revealed in the pattern of primary-response errors, where response pairs with similar vowels exhibited higher error rates than response pairs with dissimilar vowels. This corroborates the pattern shown in secondary-response latencies.

As in Experiment 4, the inhibitory effect of vowel similarity was substantially reduced when the paired CVC syllables had different final consonants. For such pairs, no effect occurred on the secondary-response latencies. The difference in the sizes of the vowel-similarity effect across the same-final-consonant and different-final-consonants conditions was significant, $t(8) = 2.23, p < .05$, one tail. This further supports the hypothesis that vowel similarity has its effect at the level of VC-clusters. On the other hand, the CV-cluster and whole-syllable hypotheses are further weakened by these results.

It should be noted, nevertheless, that the overall picture is complicated somewhat by the finding that, in Experiment 5, a significant inhibitory vowel-similarity effect on primary-response latencies did emerge for syllable pairs with different final consonants. This could be viewed as evidence that inhibitory effects of vowel similarity take place at the level of CV clusters or indepen-

dent vowels. However, the corresponding effect in Experiment 4 was considerably less (viz. 4 ms) and not reliable (Table 6); over the last two experiments, it averaged under 9 ms. In light of the secondary-response latencies, which have consistently reflected the nature of programming processes (Experiments 1 through 3; Gordon & Meyer, 1987; Meyer & Gordon, 1985), the firmest conclusion here is that the vowel-similarity effect declines markedly when the final consonants of paired CVC syllables differ from each other.

GENERAL DISCUSSION

The present research examined the inhibitory effects of vowel similarity on the motor programming of vocal responses (CVC syllable pairs). Using a response-priming procedure, we measured speech latencies, durations, and error rates for producing paired CVC syllables as a function of the similarity between their medial vowels and other subsyllabic units. In Experiment 1, where only the medial vowels varied, utterances containing similar vowels took longer and were more error prone than utterances containing dissimilar vowels (Table 2). This inhibition was most pronounced in the latencies of secondary responses, which had to be prepared after the measured latency interval began, suggesting that some central process associated with utterance planning or motor programming depends on vowel similarity. Furthermore, primary responses exhibited small but reliable inhibitory vowel-similarity effects, suggesting either that they were not fully prepared before the latency interval, or that vowel similarity may affect response execution *per se* as well as motor programming.

Our subsequent experiments replicated these results and extended them to determine the locus of the inhibition as a function of syllable structure. In Experiments 2 and 3, both the initial consonants and medial vowels of the CVC syllables varied, but the inhibitory vowel-similarity effect remained (Tables 4 and 5). However, in Experiments 4 and 5, where both the final consonants and medial vowels of the CVC syllables varied, there was no significant effect of vowel similarity on secondary-response latencies.

An overall picture of the pattern obtained from the five experiments appears in Table 8. The upper half of the table shows results averaged across Experiments 1, 2, 3, and the first part of Experiment 5, where the response syllables in each pair contained the same final consonant. The lower half of the table shows results averaged across Experiment 4 and the second part of Experiment 5, where the response syllables in each pair contained different final consonants.

These results suggest that the vowel-similarity effect stems mainly from the similarity between component VC clusters used to form successive CVC syllables.

TABLE 8
MEAN RESPONSE LATENCIES (ms), DURATIONS (ms), AND ERROR RATES (%), OBTAINED FROM AVERAGING ACROSS THE RESULTS OF EXPERIMENTS 1 THROUGH 5

| | Vowel similarity | |
Responses	Similar	Dissimilar
Same final consonant		
Primary		
Latency	303	297
Duration	293	286
Error rate	7.8	3.3
Secondary		
Latency	526	504
Duration	315	310
Error rate	9.6	6.5
Different final consonant		
Primary		
Latency	292	283
Duration	258	256
Error rate	5.1	2.4
Secondary		
Latency	494	493
Duration	276	277
Error rate	7.3	7.4

Note. The upper half of the table shows results averaged across Experiments 1, 2, 3, and the first part of Experiment 5. The lower half of the table shows results averaged across Experiment 4 and the second part of Experiment 5.

The similarity between component CV clusters did not modulate the vowel-similarity effect nearly as much, especially where secondary-response latencies were involved. A possible interpretation of this asymmetry is that VC clusters are more cohesive subsyllabic units than are CV clusters (MacKay, 1972; MacNeilage & De-Clerk, 1969). The apparent cohesiveness of VC clusters supports linguistic and psycholinguistic theories of syllable structure (Halle & Vergnaud, 1980; MacKay, 1972, 1974; Selkirk, 1982; Treiman, 1983) that decompose CVC syllables into onset (C) and rime (VC) subunits (i.e., CVC → C + VC). Some alternative theories involving other types of syllabic decomposition (e.g., Benguerel & Cowan, 1974; Kent, 1976a, 1976b; Kozhevnikov & Chistovich, 1965) are not supported here.

Relation to Connectionist Models of Speech Production

As outlined earlier, the effect of vowel similarity and the role of syllable structure may be mediated by an interactive-activation network for speech production. In such a network, nodes at different levels would represent a hierarchy of linguistic units, including syllables, phonemes, and phonetic features (Dell, 1985, 1986; MacKay, 1982; Meyer & Gordon, 1985). Lateral inhibitory connections among nodes at the same level of the network could then modulate the selection of competing subunits (cf. Rumelhart & McClelland, 1986). More specifically, the present results suggest that such a network must include an intermediate level at which VC clusters are represented and that the strength of inhibitory connections between the nodes for these subsyllabic units (i.e., rimes) is a direct function of whether they contain similar vowels and consonants.

This hypothesized mechanism could play an important role in speech production. The stream of speech flows through a succession of rapid, finely timed movements of the articulators aimed at vocalic target positions that are specified as part of a motor program (MacNeilage, 1970). During the output sequence, inhibition among similar speech units provides a "fine tuning" capacity that reduces the chances of inadvertent slips from one chosen speech unit to other neighboring units (Fromkin, 1971). In contrast, the likelihood of exchange errors between dissimilar speech units is low. Less lateral inhibition would therefore be needed to deal with them.

We should reiterate, moreover, that other levels of the production system and stages of the production process are presumably affected by similarity factors. Along with the inhibitory vowel-similarity effect found here, Experiments 2 through 4 revealed a facilitative vowel-identity effect. Secondary responses consisting of monosyllabic word pairs like PEAT–PEAK, which contained the same medial vowel, were produced more quickly and accurately than pairs like PEAT–PICK or PEAT–POCK, which contained different (similar or dissimilar, but not identical) vowels. This facilitation, unlike the preceding inhibition, occurred even when the paired items had different final consonants (Table 6). Maintaining the integrity of the final VC clusters was not essential for the facilitation to occur.

It is not yet entirely clear what underlies this dissociation between the facilitative effect of vowel identity and the inhibitory effect of vowel similarity. Perhaps the dissociation can be understood with respect to "three-dimensional phonology," in which strict dominance relations between levels of analysis are not preserved and in which segments (such as vowels) are represented on different levels from other subsyllabic constituents (Clements & Keyser, 1983; Halle & Vergnaud, 1980). Alternatively, the vowel-identity effect may instead take place less centrally as part of overt peripheral (articulatory movement) processes.

Relation to Manual Motor Performance

Some potential relations also exist between the present research on speech production and work on manual movement

performance (i.e., movements of the hands and arms). In particular, it has been demonstrated that the time (T) taken by manual movements to reach a specified target region varies logarithmically with the ratio of the target's distance (D) and width (W). This speed–accuracy trade-off, which was discovered originally by Fitts (1954) and is known as *Fitts' law* (for a review, see Meyer et al., 1990), can be expressed as $T = A + B \log_2(2D/W)$. The monotonic relation between movement time and the target distance–width ratio could have some manifestation in speech production as well.

More specifically, an analogy may be drawn between manual movements as studied by Fitts (1954) and articulatory (e.g., tongue and jaw) movements as found in the continuous production of spoken vowels. It has been proposed that the parameters for controlling the articulators in vowel production are target regions, not canonical target points (Stevens, 1972). According to this proposal, producing a selected vowel quality involves moving the tongue to any point within the confines of its target region. The width of the target region provides some freedom for achieving appropriate tongue and jaw configurations (Fowler, 1983; MacNeilage, 1970). Such flexibility enables, for example, articulatory compensation in "pipe speech," where a speaker must talk with some artificial object (e.g., a pipe) clenched between his or her teeth. When the jaw is constrained in fixed positions, speakers can readily re-adjust their tongue positions to produce acceptable vowel quality (Lindblom, Lubker, & Gay, 1979). Also, sudden perturbations of the jaw during utterances yield virtually immediate compensatory movements by the tongue and lips to restore the intended speech sounds (Kelso, Tuller, Vatikiotis-Bateson, & Fowler, 1984).

Given the analogy between target widths in manual movements and in spoken vowel production, we suspect that the times taken to produce vowels depend at least somewhat on the sizes of the vocalic target regions involved. This dependence could contribute to the inhibitory vowel-similarity effect observed here. It would explain, for example, why there were small but persistent inhibitory effects of vowel similarity on the accuracy and durations of primary responses, even though subjects had ample time to program these responses in advance. Perhaps the target regions for forthcoming vowels are adjusted dynamically as a function of other vowels with which they are paired. In particular, the target region may be narrowed if the vowels are similar, so as to maintain clearer perceptual and motor distinctions between them.

Combined with Fitts' law, the hypothesized narrowing would yield slower movements and longer response durations, as observed here. Furthermore, occasional failures to achieve a sufficient degree of narrowing could yield increased error rates, which we also observed.

Relation to Speech Perception

The inhibitory effect of vowel similarity on speech production may also have a counterpart in the perceptual domain. Contrast and similarity affect perceptual identification, not just production processes. Because of this, natural vowel systems tend to contain distinct vowels drawn from different regions of the cardinal vowel space. For example, three-vowel systems typically include {i,ɑ,u}, and four-vowel systems include {i,e,ɑ,u}. Liljencrants and Lindblom (1972) argued that a principle of maximal perceptual contrast partially accounts for the structure of natural vowel systems. It states that vowels in natural languages are distributed throughout the vowel space such that they have maximal "distances" between them. Using a numerical simulation model based on this principle, Liljencrants and Lindblom (1972) successfully predicted the distributions of vowels found in various natural languages.

According to Liljencrants and Lindblom (1972, p. 856), a further improvement of

their theory might be obtained if one could quantify and incorporate ease of production in it. The results from the present experiments provide preliminary evidence that ease of production, as reflected by response latencies and error rates, can be quantified through our response-priming procedure. A greater distance between articulatory target regions appears to be associated with greater "ease of production"; and, vice versa, utterances incorporating similar vowels appear more difficult to pronounce under time pressure. Thus, "ease of perception" and "ease of production" may be complementary characteristics of natural language.

REFERENCES

BENGUEREL, A.-P., & COWAN, H. A. (1974). Coarticulation of upper lip protrusion in French. *Phonetica*, 30, 41–55.

CATFORD, J. C. (1977). *Fundamental problems in phonetics*. Bloomington, IN: Indiana University Press.

CHOMSKY, N., & HALLE, M. (1968). *The sound pattern of English*. New York: Harper & Row.

DELL, G. S. (1985). Positive feedback in hierarchical connectionist models: Applications to language production. *Cognitive Science*, 9, 3–23.

DELL, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, 93, 283–321.

FITTS, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, 47, 381–391.

FOWLER, C. A. (1980). Coarticulation and theories of extrinsic timing. *Journal of Phonetics*, 8, 113–133.

FOWLER, C. A. (1983). Converging sources of evidence on spoken and perceived rhythms of speech: Cyclic production of vowels in monosyllabic stress feet. *Journal of Experimental Psychology: General*, 112, 386–412.

FROMKIN, V. A. (1966). Neuromuscular specification of linguistic units. *Language and Speech*, 9, 170–199.

FROMKIN, V. A. (1971). The non-anomalous nature of anomalous utterances. *Language*, 47, 27–52.

GAY, T., USHIJIMA, T., HIROSE, H., & COOPER, F. S. (1974). Effect of speaking rate on labial consonant-vowel articulation. *Journal of Phonetics*, 2, 47–63.

GORDON, P. C., & MEYER, D. E. (1987). Control of serial order in rapidly spoken syllable sequences. *Journal of Memory and Language*, 26, 300–321.

HALLE, M., & VERGNAUD, J-R. (1980). Three dimensional phonology. *Journal of Linguistic Research*, 1, 83–105.

HARTLINE, H. K., & RATLIFF, F. (1957). Inhibitory interaction of receptor units in the eye of *Limulus*. *Journal of General Physiology*, 40, 357–376.

KELSO, J. A. S., TULLER, B., VATIKIOTIS-BATESON, E., & FOWLER, C. A. (1984). Functionally specific articulatory cooperation following jaw perturbations during speech: Evidence for coordinative structures. *Journal of Experimental Psychology: Human Perception and Performance*, 10, 812–832.

KENT, R. D. (1976a). Models of speech production. In N. J. Lass (Ed.), *Contemporary issues in experimental phonetics*. New York: Academic Press.

KENT, R. D. (1976b). Syllabic complexity and syntagmatic rules for syllable production—A comment on "Aspects of the syntax of behavior" by D. G. MacKay, *Quarterly Journal of Experimental Psychology*, 28, 483–489.

KENT, R. D., & MINIFIE, F. D. (1977). Coarticulation in recent speech production models. *Journal of Phonetics*, 5, 115–133.

KOZHEVNIKOV, V. A., & CHISTOVICH, L. A. (1965). *Speech: Articulation and perception*. Washington, DC: Joint Publications Research Service.

KRUSKAL, J. B. (1964). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrica*, 29, 1–27.

KRUSKAL, J. B., & WISH, M. (1978). *Multidimensional scaling*. Sage University Paper series on Quantitative Applications in the Social Sciences, 7–11. Beverly Hills/London: Sage Publications.

LADEFOGED, P. (1971). *Preliminaries to linguistic phonetics*. Chicago, IL: The University of Chicago Press.

LILJENCRANTZ, J., & LINDBLOM, B. (1972). Numerical simulation of vowel quality systems: The role of perceptual contrasts. *Language*, 48, 839–862.

LINDBLOM, B., LUBKER, J., & GAY, T. (1979). Formant frequencies of some fixed-mandible vowels and a model of speech motor programming by predictive simulation. *Journal of Phonetics*, 7, 147–161.

MACKAY, D. G. (1972). The structure of words and syllables: Evidence from errors in speech. *Cognitive Psychology*, 3, 210–227.

MACKAY, D. G. (1974). Aspects of the syntax of behavior: Syllable structure and speech rate. *Quarterly Journal of Experimental Psychology*, 26, 642–657.

MACKAY, D. G. (1982). The problems of flexibility, fluency, and speed-accuracy trade-off in skilled behavior. *Psychological Review*, 89, 483–506.

MACNEILAGE, P. F. (1970). Motor control of serial ordering of speech. *Psychological Review*, 77, 182–196.

MacNeilage, P. F., & DeClerk, J. L. 1969. On the motor control of coarticulation in CVC monosyllables. *Journal of the Acoustical Society of America*, 45, 1217–1233.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, 18, 1–86.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive-activation model of context effects in letter perception. Part I. An account of basic findings. *Psychological Review*, 88, 375–407.

Meyer, D. E., Abrams, R. A., Kornblum, S., Wright, C. E., & Smith, J. E. K. (1988). Optimality in human motor performance: Ideal control of rapid aimed movements. *Psychological Review*, 95, 340–370.

Meyer, D. E., & Gordon, P. C. (1985). Speech production: Motor programing of phonetic features. *Journal of Memory and Language*, 24, 3–26.

Meyer, D. E., Smith, J. E. K., Kornblum, S., Abrams, R. A., & Wright, C. E. (1990). Speed–accuracy tradeoffs in aimed movements: Toward of theory of rapid voluntary action. In M. Jeannerod (Ed.), *Attention and performance XIII*. Hillsdale, NJ: Erlbaum.

Meyer, D. E., Sternberg, S., Knoll, R. L., & Wright, C. E. (1978, May). *Memory retrieval and motor programming of related-word sequences*. Paper presented at the meeting of the Midwestern Psychological Association, Chicago.

Meyer, D. E., Yantis, S., Osman, A. M., & Smith, J. E. K. (1985). Temporal properties of human information processing: Tests of discrete versus continuous models. *Cognitive Psychology*, 17, 445–518.

Öhman, S. E. G. (1966). Coarticulation in VCV utterances: Spectrographic measurements. *Journal of the Acoustical Society of America*, 39, 151–168.

Perkell, J. (1969). *Physiology of speech production: Results and implications of a cineradiographic study*. Cambridge, MA: MIT Press.

Posner, M. I., & Snyder, C. R. R. (1975). Facilitation and inhibition in the processing of signals. In P. M. A. Rabbitt & S. Dornic (Eds.), *Attention and performance V*. London: Academic Press.

Rabiner, L. R., & Sambur, M. R. (1975). An algorithm for determining the endpoints of isolated utterances. *Bell System Technical Journal*, 54, 297–315.

Rosenbaum, D. A., & Kornblum, S. (1982). A priming method for investigating the selection of motor responses. *Acta Psychologica*, 51, 223–243.

Rumelhart, D. E., & McClelland, J. L. (1986). *Parallel distributed processing. Volume 1*. Cambridge, MA: MIT Press.

Schmidt, R. A., Zelaznik, H., Hawkins, B., Frank, J. S., & Quinn, J. T. (1979). Motor output variability: A theory for the accuracy of rapid motor acts. *Psychological Review*, 86, 415–451.

Selkirk, E. O. (1982). The syllable. In H. van der Hurst & N. Smith (Eds.), *The structure of phonological representations (Part II)*. Dordrecht: Foris Publications.

Sherak, R. (1982). A real-time software voicekey and an application. *Behavioral Research Methods and Instrumentation*, 14, 124–127.

Sternberg, S., Monsell, S., Knoll, R. L., & Wright, C. E. (1978). The latency and duration of rapid movement sequences: Comparisons of speech and typewriting. In G. Stelmach (Ed.), *Information processing in motor control and learning*. New York: Academic Press.

Stetson, R. H. (1951). *Motor phonetics*. Amsterdam: North Holland.

Stevens, K. N. (1972). The quantal nature of speech: Evidence from articulatory-acoustic data. In E. E. David & P. B. Denes (Eds.), *Human communication: A unified view*. New York: McGraw-Hill.

Treiman, R. (1983). The structure of spoken syllables: Evidence from novel word games. *Cognition*, 15, 49–74.