

604

CONSONANT-VOWEL COHESIVENESS IN SPEECH PRODUCTION AS REVEALED BY INITIAL AND FINAL CONSONANT EXCHANGES

Carol A. FOWLER

Dartmouth College, Hanover, NH 03755, U.S.A.

and

Haskins Laboratories, New Haven, CT 06511, U.S.A.

Received 8 January 1987

Revised 10 April 1987

Abstract. Two experiments use a procedure developed by Carter and Bradshaw (*Speech Communication*, Vol. 3 (1984) pp. 347-360) to examine the role of syllable structure in speech production. In the procedure, subjects exchange phonological segments in corresponding positions of a pair of visually-presented words or nonwords and to produce the resulting words or nonwords as quickly as possible. Carter and Bradshaw have shown that the pattern of latencies mirrors that of frequencies of exchange errors in natural speech. The first experiment of the present study shows that initial consonant exchanges are promoted by phonetic similarity of the exchanging consonants and they reflect a bias for producing real words. With these influences controlled, Experiment 2 replicates and extends the finding of Carter and Bradshaw that initial consonant exchanges are made more rapidly than final consonant exchanges. The discussion relates the latency difference between these conditions to a difference in the "cohesiveness" of initial and final consonants with their vowel. In particular, in Experiment 2, more than one-third of errors made on final-consonant or vowel exchanges are exchanges of the whole syllable rhyme (VC), whereas just 10% of errors made on initial consonant or vowel exchanges are exchanges of the initial CV of the word. Various explanations for the difference in cohesiveness are examined in post hoc analyses.

Zusammenfassung. In zwei Experimenten wird eine Methode von Carter und Bradshaw (*Speech Communication*, Vol. 3, (1984) pp. 347-360) angewendet, um die Rolle der Silbenstruktur beim Sprechen zu untersuchen. Die Aufgabe der Versuchspersonen ist, in einem Paar visuell dargebotener sinnvoller oder sinnloser Wörter phonologische Segmente, die in gleicher Position stehen, gegeneinander auszutauschen und das Resultat so schnell wie möglich auszusprechen. Carter und Bradshaw fanden heraus, dass das Muster der Reaktionszeit dem der Häufigkeiten von Austauschfehlern in natürlicher Sprache ähnlich ist. Experiment 1 demonstriert, dass der Austausch von Konsonanten am Wortanfang durch phonetische Ähnlichkeit erleichtert wird und dass schnellere Reaktionszeiten auch auftreten, wenn sinnvolle Wörter auszusprechen sind. Experiment 2 kontrolliert diese Einflüsse und bestätigt das Resultat von Carter und Bradshaw, dass Konsonanten am Wortanfang schneller austauschbar sind als am Wortende. Dieser Unterschied hat vermutlich eine Beziehung zu der "Kohärenz" von Konsonanten mit ihren vorangehenden oder nachfolgenden Vokalen. Mehr als ein Drittel der Fehler bei finale Konsonanten- oder Vokal-Austausch bedeutet den Ersatz des gesamten Silben-Reimes (VC), während nur 10% der Fehler bei initialem Konsonanten- oder Vokal-Austausch die gesamte initialen CV-Struktur des Wortes erfasst. Verschiedene Erklärungen für den Grad der Kohärenz werden analytisch untersucht.

Résumé. Deux expériences utilisent une procédure développée par Carter et Bradshaw (*Speech Communication*, Vol. 3 (1984) pp. 347-360) pour examiner le rôle de la structure de la syllabe sur la production de la parole. Dans cette procédure, les sujets intervertissent des segments phonologiques dans des positions correspondantes d'une paire de mots ou de non-mots présentés visuellement et doivent produire les mots ou non-mots résultants aussi vite que possible. Carter et Bradshaw ont montré que les latences reflètent les fréquences des erreurs d'inversion dans la parole naturelle. La première expérience de la présente étude montre que les échanges de la consonne initiale sont favorisés par la similarité phonétique des consonnes échangées et reflètent un biais pour la production de mots réels. En contrôlant ces influences, l'Expérience 2 réduplique et étend le résultat de Carter et Bradshaw selon lequel les échanges de la consonne initiale sont réalisés plus rapidement que ceux de la consonne finale. La discussion relie la différence de latence entre ces conditions à la différence de "cohésion" entre consonnes initiale et finale avec leur voyelle. En particulier, dans l'Expérience 2, plus d'un tiers des erreurs commises sur les échanges de consonne ou de voyelle finales sont des échanges de la syllabe rime entière (VC) alors que 10% seulement des erreurs commises sur la consonne ou voyelle initiales sont des échanges de la syllabe initiale (CV) du mot. Dans les Analyses menées a posteriori, diverses explications quant à la différence de cohésion sont examinées.

Keywords. Speech production, speech errors, exchange errors, syllable structure.

Spontaneous errors of speech production (for example, "morage in the fountains" for "forage in the mountains" or "even the best team losts" for "even the best teams lost") have proven quite revealing sources of information concerning the structures and processes involved in language production (see, for example, Garrett, 1976, 1980; Shattuck-Hufnagel, 1979, 1983; Dell, 1986). The errors are strikingly systematic in the units of speech that participate in errors (for example, phonemes and words, but rarely syllables and features) and in the conditions that promote errors. Recently, however, researchers have recognized that converging evidence is needed from laboratory studies of speech (e.g. MacKay, 1971; Baars and Motley, 1976; Cutler, 1981) because speech error collections are subject to bias and because they do not provide all the information needed to evaluate error patterns. One source of bias in corpora of spontaneous errors is contamination by mishearings. This is less likely to be characteristic of speech transcribed from recordings made in a laboratory setting. As for missing information in error collections, questions relating to the frequency with which certain error types occur in comparison with their opportunities to occur cannot easily be answered using spontaneous-error collections because opportunities for error are not available except by estimate from other sources.

A variety of experimental techniques has been developed (e.g. MacKay, 1971; Baars and Motley, 1971; Kupin, 1979; Levitt and Healy, 1985; Dell, 1986; Carter and Bradshaw, 1984) to study language production and speech errors in the laboratory. These procedures have verified many of the characteristics of errors previously observed in corpora of spontaneous speech errors. But in addition, they have uncovered new ones that are difficult to notice in spontaneous-error corpora.

For example, many researchers now have identified a "lexical bias" in experimentally-elicited phoneme errors such that errors are more likely to occur if the erroneously produced string is a real word of the language than if they are not (Baars, Motley and MacKay, 1975; Dell, 1980). In contrast, based on the high proportion of non-words that they had observed in their in error corpora, Fromkin (1973) and Garrett (1975, 1976)

had concluded that phoneme errors do not tend to create real words. Uncovering a statistical tendency such as a lexical bias is made easier if experimental conditions can be established in which opportunities for word- and nonword-creating phoneme errors are controlled. (But see Dell (1980), Dell and Reich (1981) who demonstrate that the same bias does occur in spontaneous phonological-segment errors as well.)

Another area in which laboratory studies of speech errors have made a special contribution is in our understanding of the conditions that affect phoneme-substitution patterns. Examination of substituting and substituted phonemes in an error corpus had led Shattuck-Hufnagel and Klatt (1979) to conclude that, in general, errors are symmetrical such that the frequency with which, for example, /z/ substitutes for /d/ is nearly the same as the frequency with which /d/ substitutes for /z/. In turn, the symmetry suggested that substituting phonemes are not, in any sense (such as their relative frequency in the language or their relative markedness), "stronger" than substituted phonemes. However, as Levitt and Healy (1985) point out, this conclusion involves the implicit assumption that the opportunities for /d/ and /z/ to be targets of substitutions is the same. The opportunities are the same in cases of errors in which both the substituting and substituted segments are present in the intended utterance, but they are not when the substituting segment is not in the utterance as planned. Instead, /d/ occurs more frequently than /z/ in speech (Roberts, 1965). Thus, in order for the *number* of /z/ → /d/ and /d/ → /z/ substitutions to be the same, the *probability* of /d/ substituting for /z/ must in fact be higher than the probability of /z/ substituting for /d/. In experiments controlling opportunities for error, Levitt and Healy were able to support hypotheses that some segments are stronger than others in that they participate as substituting phonemes relatively frequently and that strength is in part a function of the segments' frequencies of usage in the language. Motley and Baars (1975) show, in addition, that strength is related to the transition frequencies of phoneme sequences created by an error.

The present research adopts an experimental approach to the study of the role of syllable struc-

ture in speech errors and in speech production more generally (see also MacKay, 1978). It follows a recent study by Carter and Bradshaw (1984).

In spontaneous errors, pairs of segments involved in errors (e.g. "morage in the fountains") almost always preserve their intended position in the syllable. That is, with few exceptions syllable-initial consonants involved in movement errors exchange with other syllable-initial consonants; vowels exchange only with vowels, and final consonants with final consonants. Moreover, by far the most frequent single-segment errors involve initial consonants of syllables and words; final consonant errors are reported with a much lower frequency.

Spontaneous-error corpora allow at last three interpretations of this imbalance in error frequencies. In part it may be due to differential likelihoods that error collectors will hear initial- as compared to final-consonant errors. Listeners are known to "restore" incorrectly-produced phonemes later in a word with greater frequency than they restore initial phonemes (Marslen-Wilson and Welsh, 1978). A second possibility is that whereas errors in initial and final position of a syllable or word may occur under exactly the same conditions (e.g. the interacting consonants tend to be phonetically similar, occur in similar environments (Shattuck-Hufnagel, 1979) and create real words of the language), other things equal, those conditions may arise with greater frequency in syllable-initial than in syllable-final positions. A third possibility is that, as other theorists and researchers have suggested for independent reasons, syllables may have an internal structure in which vowels and final consonants cluster into a constituent (called the "rime" or "rhyme"), while initial consonants serve as their own constituent, the "onset". One reflection of this constituent structure may be a greater cohesion of final consonants with their vowels than of initial consonants and vowels and therefore, a lesser tendency for final consonants to break off from vowels in errors. Other reflections are that poetic devices (such as alliteration on the one hand and rhyming on the other) respect the organization of syllable into onsets and rhymes, popular language games (such as Pig Latin) do, stress rules of En-

glish and many languages refer to syllable rhymes (see e.g. Prince, 1980) but not to onsets, and both children (Trieman, 1985) and adults (Trieman, 1986) learn word games more easily that involve segmenting or blending syllables along constituent-structure lines than elsewhere.

Carter and Bradshaw (1984) have developed an experimental procedure to exploit the possibility that the relative frequency with which segments participate in errors relates to their degree of detachability from a syllable or word. They explicitly asked subjects to exchange two phonemic segments in corresponding positions in a word pair and measured the latency and accuracy with which subjects were able to do so. The segments to be exchanged were, in different conditions, the initial consonants, the medial vowels and the final consonants of pairs of monosyllables presented visually. If both the frequency with which segments participate in spontaneous errors and the latency with which they can be exchanged on purpose redundantly reflect the detachability of a segment from a syllable or word, then latencies should pattern as errors do. In support of this idea, Carter and Bradshaw found two latency patterns characteristic of speech error patterns. First, they found a lexical bias in intended segment exchanges such that exchanges creating real words were made more rapidly than those creating non-words. Second, they found that initial consonants were exchanged more rapidly than medial vowels or final consonants. Because the dominance of initial consonant exchanges appeared in the latency measure, the possibility suggested earlier that initial consonant errors predominate in spontaneous errors only because they are more likely to be detected by collectors can be ruled out.

The present experiments used a modification of Carter and Bradshaw's procedure to pursue the findings on syllable structure. The experiments of Carter and Bradshaw leave open the possibility also suggested earlier that, other things equal, contextual conditions may tend to favor initial over final consonant exchanges. Although Carter and Bradshaw controlled the number of opportunities for initial- and final-consonant errors, they did not explicitly control for other variables that may generally differ between consonants in those positions in a syllable or word. In particular,

consonant exchanges in spontaneous errors appear to be promoted by phonetic (featural) similarity between the exchanging phonemes and by phonetic similarities in the environments of the exchanging segments (Shattuck-Hufnagel, 1979). In Experiment 1, using a slight modification of the paradigm of Carter and Bradshaw, I confirm that phonetic similarity of exchanging consonants promotes exchanges. In Experiment 2, I control the similarity of initial and final consonants and the environments in which they occur and ask whether their exchange latencies differ.

If differential phonetic similarity does not explain the tendency for initial consonants to dominate in single phoneme exchanges, a different possibility is that the tendency indeed reflects an onset-rhyme constituent structure of the syllable. This interpretation presumably is favored by Carter and Bradshaw and promoted also by the independent evidence on syllable structure cited earlier. Direct evidence for this interpretation might be provided by the pattern of errors in the exchange-latency paradigm. If final consonants are, in some sense, more cohesive with the vowel than are initial consonants, then on occasions in which subjects are instructed to exchange vowels, for example, mistakes in which they exchange the whole rhyme should be relatively more frequent than mistakes in which they exchange the initial consonant-vowel (CV) sequence. Likewise, instructed to exchange initial consonants, subjects should inadvertently exchange CVs rarely; however, instructed to exchange final consonants, they should exchange VCs relatively frequently. In Experiment 2, I look for such error patterns in subjects' exchange errors.

An explanation of speech errors and latency in terms of syllable structure or relative "cohesiveness" is, of course, incomplete without some indication of what work, exactly, syllable structure does in language production. Dell (1986) has proposed, for example, that the constituent structure of syllables is reflected explicitly in associative connections between phonemic segments and abstract onset and rhyme nodes in a spreading-activation network representing the lexicon. These explicit connections foster cohesiveness among the subunits sharing a common higher node. However, it would be useful to discover *why* syl-

lables structure in that way, particularly in view of suggestions that many languages, indeed the majority that have been examined (Prince, 1980) have an onset-rhyme structure. Two post-hoc analyses of the findings of Experiment 2 are performed to examine possible sources or correlates of the effects of syllable structure on speech errors.

Experiment 1

The first experiment was conducted in part to look for effects of phonetic similarity of consonants segments on time to make an initial-consonant exchange and in part to test a small modification to the procedure of Carter and Bradshaw. Behind these aims of the experiment was an attempt, in addition to those of Carter and Bradshaw, to verify that the exchange-latency paradigm yields data that pattern similarly to frequency patterns of natural speech errors.

In the research by Carter and Bradshaw, subjects were timed as they performed exchanges in a list of letter-string pairs. The dependent measure was the average time per stimulus pair to make an exchange, obtained by dividing the time to make all of the exchanges in the list by the number of pairs in the list. Because the procedure required subjects to read a pair of items, cover it with a card and then report the exchange, latencies included the times to execute all of these activities. Latencies were slow and presumably more variable than they might be if some of these components were excluded from the latency measure. In an effort to remove the interval to utter the response pair from the latency measure (that is, to remove time from articulation onset to offset), Carter and Bradshaw measured articulation time separately and subtracted it from the latency measure. One consequence of this correction was to eliminate the previously significant effects of response-pair lexicality on response time. If this procedure indeed successfully removes articulation time from the latency measure and removes nothing else, then it shows that response pair lexicality affects only articulation time, not time to exchange consonants or vowels. Too, however, it would indicate a finding using the ex-

change-latency procedure that differs from findings on speech errors. As already noted, in both error corpora and experimentally-elicited errors, there is a lexical bias in the frequency of phoneme exchanges.

The modification to the procedure adopted here was to obtain a measure from which articulation time was excluded from the outset. The procedure was to obtain a measure of time from visual presentation of a single stimulus pair to the onset of the acoustic signal for the subject's exchange response. This was accomplished by presenting stimulus pair on a computer-terminal screen and obtaining response times using a voice key. Another benefit of this modification may be to reduce some of the variability of the average latency measure of Carter and Bradshaw due to the subject's variable times to cover each stimulus pair after reading it.

Using this modified procedure, then, I asked whether effects of lexicality of the stimulus and response strings could be detected in the latency measure and whether phonetic similarity of the exchanged consonants reduces latency to exchange the pair.

Method

Subjects. Subjects were 18 students at Dartmouth College who participated in the experiment for course credit. They were native speakers of English who reported normal speech and hearing.

Materials. Materials consisted of 72 pairs of letter strings, 12 each in the six cells representing the crossing of three levels of lexicality of stimulus and response strings described below and two levels of phonetic similarity of the consonants involved in the exchange. Stimulus pairs are listed in Appendix 1.

In one lexicality condition, *W-W*, both the stimulus and response strings were real words of English (e.g. PASTE TOLL → TASTE POLL); in a second condition, *W-NW*, the stimulus strings were words, but response strings were pseudowords (e.g. PAIN TOAD → TAIN POAD); in the third condition, stimulus strings were pseudowords and response strings were words (e.g. PAME TOPE → TAME POPE).

Across conditions having real-word stimulus strings, words were matched as closely as possible in median frequency according to the tables of Francis and Kučera (1982). (The range of median frequencies across the six conditions was 12–26 occurrences in the Brown corpus.) Likewise, across conditions having real-word response strings, response words were matched in median frequency (range: 8–16 occurrences).

Twelve pairs of consonants were selected to serve as initial consonants in the condition in which consonants of a stimulus pair were phonetically similar. The same 12 consonant pairs were used at all three levels of lexicality. The consonants differed by one phonetic feature according to the feature system of van den Broecke and Goldstein (1980). This feature system was used because Levitt and Healy (1985) had found that in comparison with the feature system of Chomsky and Halle (1968), it accounted well for phoneme-error frequency patterns.

To create the condition in which consonants were phonetically dissimilar, I re-paired the consonants in the phonetically-similar condition so that new pairs differed by at least two phonetic features. There was just one exception to this in which a /t/ in the similar condition was changed to /s/ in the dissimilar condition to create a two-feature difference in a new stimulus pair of initial consonants.

An attempt was also made to use the same medial vowels across the six experimental conditions, or, in three instances, where that was not possible, to use vowels with about the same degree of opening. This precaution and the fact that the set of consonants triggering the voice key was matched across conditions obviate any need to measure and correct for voice key effects.

A set of 24 practice trials was also devised, consisting of words and pseudowords similar to those used in the test trials.

Procedure. Subjects were run individually. They were instructed that they would see pairs of letter strings presented on a computer-terminal screen. They were to exchange the initial consonant sounds of the pair members as quickly and as accurately as possible and to say the response string into the microphone in front of them. The distinc-

tion between exchanging initial *sounds* of words rather than *spellings* was illustrated by example, using words having initial consonants spelled with two letters (such as "th") and using words having initial consonant letters (e.g. "c") whose pronunciation might change were subjects to exchange letters rather than sounds.

Subjects were instructed not to begin to say the exchanged pair until they had determined the pronunciation of both members of the pair. To encourage them to do this (and, therefore, to obtain latencies that consistently reflected time to make the whole exchange), the stimulus string was line fed off the screen as soon as the voice-key was triggered by the subject's initial response.

Letter strings were printed in capital letters on the terminal screen out of sight of the subject and then were line fed into view as the response-time clock began to measure latency to respond. The experimenter sat opposite the subject, viewing a monitor that showed both the stimulus items and the correct response items. During the block of 24 practice trials, the experimenter provided corrections to the subject as needed. Following the 24 practice trials, subjects were given feedback on the terminal screen consisting of their average response time for the block. Next they received three blocks of 24 test trials with response-time feedback after each block, but with no experimenter-provided corrective feedback. Experimental conditions were randomized within blocks and were differently randomized for each subject.

Subjects' responses were recorded on audio for later checking. Failures of the voice key to trigger were eliminated from the data. In addition, erroneous responses were eliminated from analysis of latencies.

Design. The experiment had two independent variables, lexicality of stimulus and response strings, with three levels, *W-W*, *NW-W* and *W-NW*, and phonetic similarity of initial consonants, with two levels. Dependent measures were latency and percent errors.

Results and discussion

Latencies and error percentages for the three lexicality conditions and the two levels of pho-

netic similarity are given in Table 1. Errors averaged just over 10% in all conditions and showed no significant differences across conditions in an analysis of variance. They will not be considered further.

In an analysis of variance with factors, lexicality and phonetic similarity, both independent variables were significant (lexicality: $F [2,34] = 7.64$, $p = 0.002$; phonetic similarity: $F [1,34] = 4.18$, $p = 0.05$). The interaction between the variables was not significant ($F [2,34] = 1.70$, $p = 0.2$). Post hoc analyses showed that the effect of lexicality was due to a difference between the two conditions in which the response items were real words on the one hand and the condition in which they were pseudowords on the other. Conditions in which response items were real words of English led to faster exchanges than conditions in which response items were nonwords. The 68 ms difference between the *W-W* and *NW-W* conditions was not significant. Therefore, lexicality of the response did affect latency significantly, while lexicality of the stimulus items did not.¹

In large part, the effect of lexicality on latency is consistent with findings of Carter and Bradshaw before they attempted to eliminate articulation time from their response-time measure. In addition, it is consistent with the lexical bias characteristic of error corpora (Dell, 1980) and of experimentally elicited errors (Baars, Motley and MacKay, 1975).

Phonetic similarity shortened exchange latency as predicted. This finding is consistent with observations from spontaneous error corpora that exchanges occur disproportionately frequently among phonetically similar consonants. Although the effect was reliable, it was weak. Accordingly, it is unlikely to have been a major source of confounding in the research by Carter and Bradshaw even though it was not explicitly controlled in that research.

The present experiment suggests, following the work of Carter and Bradshaw, that the exchange-

¹ The conclusion that there is a lexical bias here is weakened by the presence of a kind of bias in the stimuli. Two thirds of the response strings form words; accordingly, subjects may have come to expect word outcomes. However, a similar imbalance in stimulus strings did not lead to a comparable advantage for stimulus strings that form words.

Table 1
Response times and error percentages for the experimental conditions of Experiment 1

Lexicality	W-W	NW-W	W-NW
	RT	1561	1629
error	10.6	10.7	10.3
Phonological similarity			
	Similar pairs	Dissimilar pairs	
RT	1628	1701	
error	10.4	10.7	

latency paradigm does lead to latencies that pattern in ways similar to the patterning of relative frequencies of errors on phoneme-error corpora. In the present experiment, this generalization holds true in respect to effects of lexicality of the response strings and the phonetic similarity of segments that interact in exchanges. In Experiment 2, the paradigm is used to examine effects of syllable structure on response latency and errors.

Experiment 2

In Experiment 2, phonetic similarity of consonants to be exchanged and of their vocalic environment is controlled across conditions in which subjects exchange initial and final consonants in a syllable. In the experiment, I ask first whether the initial/final consonant difference in favor of initial exchanges is preserved when phonetic similarity of the consonants and their environments is controlled, and second whether the latency difference between the conditions can be characterized as one of differential cohesion of initial and final consonants with a syllable's vowel.

I examine the second issue by looking at errors that subjects make under instructions to exchange initial consonants, vowels or final consonants. If initial consonants cohere with vowels less than final consonants do, as evidence reviewed in the introduction might predict, then, for example, when subjects are asked to exchange vowels, they should relatively more frequently erroneously move the final consonant with the vowel than

move the initial consonant with the vowel. Similarly, when initial consonants are to be exchanged, concurrent vowel exchanges should be rare among errors; when final consonants are to be exchanged concurrent vowels exchanges should be relatively more common.

Method

Subjects. Subjects were 12 undergraduates at Dartmouth College who participated in the experiment for course credit. They were native speakers of English who reported normal speech and hearing.

Materials. Stimuli were 48 pairs of words; they are listed in Appendix 2. The same stimulus pairs were used in conditions in which ICs, Vs and FCs were to be exchanged. In all conditions, both stimulus and response items were real words of English (or, in two cases, they were names). Moreover, the stimulus items were selected so that across conditions, not only were the stimulus items the same, so were the response items. (For example, "PUN TICK" was the response for "TON PICK" in the IC-exchange condition, it was the response for "PIN TUCK" in the V-exchange condition and it was the response for "PUCK TIN" in the FC-exchange condition.) This was accomplished by identifying 12 sets of word pairs (for example, TUCK PIN) such that permutations of initial consonants, vowels and final consonants all created real words of the language (PUCK TIN, TICK PUN, TON PICK). Because all four permutations were used as stimulus strings in all three exchange conditions, stimulus and response items were identical across conditions; only the pairing of stimulus and response strings differed across conditions. Table 2 illustrates how this worked for the set of response strings for the family of items including "PUN TICK". By creating stimulus items in this way, the frequency of stimulus items and response items as well as the effects of response-item initial consonants on voice key were equated across experimental conditions. In addition, the similarity of the vocalic environments of the exchanging consonants was the same for initial and final consonant exchanges.

Across the items, initial and final consonants were matched in featural similarity according to the feature system of van den Broecke and Goldstein (1980). Initial consonants within a pair differed by 1.5 features on average; final consonants differed by 1.67 features.

Procedure. The procedure was essentially the same as that in Experiment 1. Exceptions were that subjects participated in three sets of trials, each consisting of one practice block and four test blocks of 12 trials each. In one set of trials, subjects exchanged initial consonants, in another, vowels and in a third, final consonants. The order in which these sets of trials was presented was counterbalanced across the 12 subjects. As in Experiment 1, subjects' responses were recorded on audio tape. However, due to experimenter error, one subject's responses were not recorded.

Design. The experiment had one independent variable, exchanged segment, with three levels, *IC*, *V*, *FC*. Dependent measures were latency and percentage of errors.

Results and discussion

Latencies and errors. Response times and error percentages are presented in Table 3. The table shows that latencies were fastest for initial consonant exchanges and slowest for vowel exchanges. For purposes of analysis, latencies were transformed into their reciprocals to correct for inhomogeneity of variance across the *IC*, *V* and *FC* conditions. In an analysis of variance on the reciprocals, the main effect of exchanged segment (*IC*, *V*, *FC*) was highly significant ($F [2,22] = 35.27, p < 0.0001$). Post hoc analyses showed that

all pairwise differences were significant.

The analysis of errors also revealed a main effect of exchanged segment ($F [2,22] = 12.87, p = 0.0002$) with the significant effect in this case due to a smaller error rate on initial-consonant exchanges than on vowel and final-consonant exchanges which did not differ.

Analysis of errors. A major purpose of the experiment was to examine errors in which the subject moved more than just the designated segment. The question was whether whole-rhyme (henceforth, *VFC*) exchanges occurred disproportionately among errors of this type. To answer this question, errors from the 11 subjects whose responses had been recorded were classified into categories. In the *IC* exchange condition, errors were classified as *ICV* errors, in which the vowel was exchanged with the initial consonant, as *ICFC* errors in which the final consonant was moved with the initial consonant, or as "other" errors including everything else. To enable comparison with errors in the *V* and *FC* conditions, which had much higher error rates, the total errors in the *ICV* and *ICFC* conditions were each expressed as proportions of the total errors on initial consonants (that is, as conditional probabilities of an *ICV* or *ICFC* error given that an initial-consonant exchange occurred). Similarly, errors in the *V* condition were classified as *ICV*, *VFC* or as other errors and were expressed as proportions of the total errors on vowels. Finally, errors in the *FC* condition were classified as *ICFC* or *VFC* or as other errors and were converted to proportions of total *FC* errors. Finally, error percentages were collapsed across conditions so that, for example, the *ICV* percentage contributed by the *IC* exchange condition and that contributed by the *V* exchange condition were averaged. This was done because error percentages were very

Table 2
Sample stimulus and response pair used in Experiment 2

Stimulus strings	Response strings		
	<i>IC</i>	<i>V</i>	<i>FC</i>
TUCK PIN	PUCK TIN	TICK PUN	TON PICK
PUCK TIN	TUCK PIN	PICK TON	PUN TICK
TICK PUN	PICK TON	TUCK PIN	TIN PUCK
PICK TON	TICK PUN	PUCK TIN	PIN TUCK

Table 3
Response times and percentages of error for the conditions of Experiment 2

Type of exchange	Response times and percentages of error		
	<i>IC</i>	<i>V</i>	<i>FC</i>
RT	1606	3133	2448
error	9.3	22.5	23.2

closely matched across these symmetrical conditions.

Error percentages were 10.8, 36.2 and 5.5 in the categories *ICV*, *VFC* and *ICFC* respectively. As expected if segments in the rhyme are more cohesive than segments in different syllable constituents, errors occurred disproportionately in the *VFC* category. Indeed, over one-third of all *V* and *FC* exchange errors were whole-rhyme exchanges. An analysis of variance on the three percentages showed a main effect of error category ($F [2,20] = 20.25, p < 0.001$). Post hoc tests showed that the difference between the *ICV* and *VFC* categories was significant ($F [2,20] = 12.09, p < 0.004$) while that between the two cross-constituent categories (*ICV* and *ICFC*) was not.

Post hoc analyses. The foregoing analysis shows that phoneme errors exhibit the same evidence for an internal syllabic constituent structure that other evidence from linguistic theory, linguistic games, and poetic devices suggests. In itself, however, it does not reveal any basis for the constituent structure or any hints as to why syllables partition as they do. Here I look at two possible correlates of the onset-rhyme structure that may offer deeper insight into its origins or functions if any.

One correlate relates to the relative frequencies of initial and final consonants. For example, *CV*s are popular syllable types both across languages (e.g. Clements and Keyser, 1983) and in English, where, in general words syllabify so as to maximize syllable onsets (e.g. Hoard, 1971). Possibly, syllable-initial consonants have connections to many more words in the lexicon on average than do final consonants and, thereby, are more detachable from any given word. This account might be tested by correlating consonant frequency in initial and final positions of a word with latencies to make initial and final exchanges.

For purposes of this analysis, two frequency counts were used, those of Hultzen, Allen and Miron (1964) and of Roberts (1965). These tables provide word-initial and -final phoneme frequencies based on corpora of spoken English. Both counts are based on multi-syllabic words as well as monosyllables, so the frequency counts may not be entirely accurate for present purposes. If

they are accurate, however, they show that the initial/final difference in exchange latency does not derive from a frequency difference favoring initial consonants for the stimuli in Experiment 2. Instead, for those stimuli, word-final consonants have a higher overall frequency than word-initial consonants. Therefore correlations between consonant frequency and exchange latency are overall positive ($r = 0.33, p < 0.05$ for the relevant consonant in the first word of a stimulus pair and $r = 0.27, p < 0.05$ for the relevant consonant in the second word) for the initial- and final-consonant exchange conditions of Experiment 2². However, with effects of the dichotomous variable, initial versus final consonant exchange, partialled out, the correlations are both nonsignificant. The dichotomous variable by itself, with effects of frequency partialled out, correlates strongly with latency ($r = 0.75$). Although the analysis does show a relationship between phoneme frequency and latency, the direction of the correlation does not support the hypothesis that high frequency of association to words renders a consonant easily detachable from a given word.

A different way of looking at frequency as a basis for cohesion is suggested by these positive correlations. Possibly, frequent cooccurrences of particular consonant and vowel pairs in words promotes cohesiveness between them. Accordingly, perhaps *CV* and *VC* diphone frequencies will be found to predict exchange latency. For these frequencies to predict the response time difference in latency between initial and final consonant exchanges, *CV* diphones in the stimulus pairs of Experiment 2 would have to be lower in frequency than *VC* diphones, and indeed there is a small difference in their frequency in the expected direction. Correlations between exchange latency and diphone frequency in *stimulus* pairs should be positive such that high diphone frequency retards detachment of an *IC* or an *FC* from its vowel. By the same reasoning, if diphone

² These correlations are based on the analysis using tables from Hultzen et al. Those based on Roberts' tables were similar in magnitude and pattern. Substituting log frequencies for raw frequencies did not improve the correlations, and so correlations involving raw frequencies are reported throughout.

frequencies of *response* pairs predict exchange latency, correlations between the variables should be negative such that high frequencies of the newly-created CV or CV pairs promote the exchange.

For purposes of examining the effects of diphone frequency on exchange latency, Roberts' frequency counts are not appropriate, because they do not provide a fine enough transcription of the vowels in the corpus. (Roberts transcribes the vowels in the corpus as any of eight monophthongs.) Accordingly, correlations are based on word-initial CV diphones and word-final VC diphones from the tables of Hultzen et al. These were checked using the frequency tables of Carterette and Jones (1974). The two sets of correlations were very similar, and so just those from Hultzen et al are reported here.

Interestingly, the correlations patterned as predicted only for the first stimulus word and the first response word of a pair—that is, for the word whose utterance by the subject triggered the voice key. For the first stimulus word, the correlation with latency is weak but positive ($r = 0.19$, $p = 0.05$); for the second stimulus word, the correlation is negative ($r = -0.35$, $p < 0.05$). These correlations remain significant with effects of the dichotomous variable, initial- or final-consonant exchange, partialled out. For the first response word, the correlation with latency is negative ($r = -0.24$, $p < 0.05$); for the second, it is near zero ($r = 0.02$). Both of the correlations approach zero with the effect of consonant position partialled out. In a multiple regression analysis with the diphone frequencies of the first stimulus word and the first response word as predictors of latency, a multiple R of 0.3 is obtained. This is much smaller than the correlation of 0.79 between the dichotomous variable, consonant position, and latency with effects of the four diphone frequency variables (first and second stimulus and response words) partialled out.

Thus, the present analysis gives only weak support to the idea that frequency affects detachability of a consonant from its vowel. For the word whose utterance triggers the voice key, high frequency of occurrence of the to-be-detached consonant and its vowel lengthens exchange latency, while high frequency of occurrence of the to-be-

attached consonant and the same vowel shortens latency. However, the variable of consonant position itself explains much more of the variance in response latencies than do the diphone frequency variables. This suggests at least that different frequencies of occurrence of initial and final consonants with their vowel do not exhaust the reasons why final consonants are more difficult to detach from a syllable than are initial consonants. However, for this analysis as for the first, these conclusions must be qualified, because the frequencies used in the analysis may not be sufficiently close to those in the language user's "mental lexicon".

A different approach to examination of the effect of consonant position in the syllable on exchange latency focuses specifically on the finding that initial and final consonants appear to cohere differentially with the vowel. In particular, it relates to the possibility that the gestural relationship between vowels and consonants is different for syllable-initial and -final consonants.

In English (Fowler, 1983), as in some other languages including Swedish (Lindblom and Rapp, 1973), a vowel in a syllable is durationally shorter the more consonants surround it in the syllable. Moreover, the shortening is asymmetrical so that syllable-final consonants shorten the vowel more than syllable-initial consonants. One account of this shortening asymmetry (Fowler, 1983) is that it is a reflection of coarticulatory overlap between consonants and vowels (e.g. Ohman, 1966; Carney and Moll, 1971). Indeed, Lindblom, Lubker, Gay, Lyberg, Branderud and Holmgren (in press) have evidence that when closure duration of a consonant is shortened by a bite block which enforces an unusual and constant amount of jaw opening, measured vowel duration increases by the amount of closure shortening. This would be expected if the shortening is due to a coarticulatory overlaying of the vowel by the consonant. The asymmetry in shortening presumably signifies that gestural overlap of vowels and consonants is more extensive for vowels and final consonants than for vowels and initial consonants. Of course, this difference in itself may be yet another reflection of a more fundamental onset-rhyme constituent structure to the syllable. However it may indicate at least that the reasons for the VC cohesion noticeable in errors and ex-

change latencies may be the same reasons, whatever they may be, for the asymmetry in gestural overlap of initial and final consonants with vowels.

A spatial or articulatory, rather than purely temporal, manifestation of the same consonant-vowel overlap might be greater cohesiveness between a vowel and a consonant on either side of it to the extent that the consonant shares articulatory or featural properties with the vowel. That is, a vowel may cohere more with consonants that share gestural or featural properties with it than with less similar consonants. If so, then the difference in the cohesiveness of initial and final consonants with the vowel might have as its basis a difference in the extent to which the production of initial and final consonants can merge with production of the vowel.

To investigate this idea, I looked at the difference in the latencies of initial and final consonant exchanges as a function of three levels of "sonority" of the consonant pairs being exchanged. "Sonority" refers to the "loudness" [of a sound] relative to other sounds with the same length, stress and pitch" (Ladefoged, 1982, p. 221). Vowels are the most sonorous segments, followed in order by /l/ and /r/, nasals, fricatives and stops (e.g. Ewan, 1982). Among consonants, then, degree of sonority relates roughly to degree of articulatory vowel-likeness.

In many languages, syllable structure respects a sonority hierarchy such that consonantal segments in a cluster increase in sonority nearer the vowel. Accordingly, for example, in a prevocalic cluster containing /t/ and /r/, the order must be /tr/; but the ordering is reversed for clusters after the vowel. One way to conceptualize the difference in detachability of initial and final consonants from a vowel, compatible with the observation that segments more cohesive with the vowel occupy slots closer to it, would be to think of final consonants as occupying slots relatively closer to the vowel than do initial consonants.

Although in Experiment 2, I controlled overall phonetic similarity of initial and final consonants, I did not control degree of sonority. To examine any role that sonority might play in the degree of cohesiveness between vowel and consonant, I partitioned stimulus pairs into three categories ac-

ording to the sonority of the consonants that were exchanged. In one category exchanged consonants were stops; in a second category one member of an exchanged pair was a stop and one was some more sonorous consonant; in the third category, neither segment of a pair was a stop. Table 4 provides the average latencies of exchanges for stimuli in these categories, distinguished by the initial or final position of the consonant pair in the word.

The results show an increase in latency with consonantal sonority among both initial and final consonants. This is consistent with the idea that more vowel-like consonants are less detachable from the vowel than others. However, with initial and final consonants essentially matched in sonority (for example, in trials where consonants being exchanged are uniformly stops), there remains a very large difference in latency between initial and final consonant exchanges, suggesting strongly that a tendency for initial and final consonants to differ in sonority does not go far to explain the difference between them in exchange latency and, presumably, also in cohesiveness with the vowel.

An analysis of variance performed on reciprocals of the latency data summarized in Table 4 revealed significant effects of consonant position ($F [1,90] = 214.24, p < 0.0001$) and of sonority ($F [2,90] = 3.50, p = 0.03$), but no interaction between the variables ($F < 1$).

In summary, then, the analysis does suggest that more vowel-like consonants are harder to detach from the vowel than less vowel-like consonants. However, this difference by itself cannot explain the effects of consonant position on exchange latency, because with sonority of ICs and FCs essentially matched, large differences in latency remain. Of course, it is possible, even likely, that final consonants and vowels do over-

Table 4
Latencies of initial and final exchanges distinguished by the sonority of exchanged consonants

Sonority	Sonority		
	Low (Stops)	Middle	High
Initial	1534	1590	1692
Final	2333	2431	2615

lap more in production than do initial consonants and vowels, perhaps because as just suggested, in some sense, final consonants occupy slots closer to the vowel than do initial consonants.

A speculative reason why final consonants may nestle closer to the vowel derives once again from considerations of temporal coarticulatory overlap in a *CVC* syllable. The articulation of a consonant may be characterized as including three phases, a closing phase in which its characteristic constriction is approached, a closure phase in which the constriction is maintained, and a release into a following segment, if any. A vowel need have only an opening phase; in addition, in slow speech, it may have a phase in which the vocal-tract shape that is the target of the opening gesture is maintained.

Coarticulatory overlap between the initial consonant and the primary articulations for a vowel (that is, articulations involving the tongue body and jaw that produce the characteristically open tract shape for the vowel), occurs when vowel production is initiated during consonant closure or release. The early phases of vowel production will consist of positioning movements of the tongue body and of jaw opening. The opening gesture cannot occur too early within the initial consonant, else it will cause premature release of the consonant and an acoustic signal that does not specify the intended consonantal phoneme. Possibly, then, temporal overlap between the initial consonant and the jaw-opening gestures for the vowel is constrained by the gestural requirements of the consonant. Of course, as other theorists have noted (e.g. Daniloff and Hammarberg, 1973), some overlap between consonant and vowel is essential to prevent production of unintended vocalic sounds on release of the consonant.

On the other side of the vowel, overlap with the final consonant occurs when the jaw and other primary articulators for the consonant begin the consonant's closing phase. On this side, more overlap between vowel and consonant may be possible, because the final consonant has only to permit enough opening for the vowel so that the vowel's identity is specified in the acoustic signal.

Accordingly, perhaps in the planning of a syllable or monosyllabic word where phoneme errors

appear to arise (e.g. Dell, 1980), the possibilities for greater *V-FC* than *IC-V* overlap are reflected in a way that gives rise not only to asymmetrical coarticulatory overlap and shortening, but also to differences in initial- and final-consonant exchange frequencies and to the other manifestations ascribed to an onset-rhyme structure of the syllable. That "way" is, indeed, what the claim that syllables have an onset-rhyme constituency may represent.

Appendix 1

Phonetically similar

W-W	W-NW	NW-W
barn door	bars dorm	bart dord
lake room	lame rule	lail roop
shale sun	shade sup	shafe sutt
gin choice	gist choate	jick choe
paste toll	pain toad	pame tope
cat pane	cab paid	cal pake
zip sue	zig soup	zill soom
ball gun	bought gull	bolf guck
wield yell	week yes	weast yed
tall cone	toss cope	toff coes
mutt nail	muss nape	mub nake
feel poke	feast pope	feep pome

Phonetically Dissimilar

barn your	bars yawn	baht yack
lake cone	laze coast	lape cobe
shale pone	shame pour	shane pode
gin sam	gym sad	jick soe
pest gain	pen gaze	pepp gare
cat sum	calf sun	cack suff
zip new	zig newt	zill soom
balk chaste	bought choir	baw chail
wield poke	wean pole	weel pore
toll route	tall rum	toach rick
mat soul	man soak	mand soat
feel done	feast dub	fean duss

Appendix 2

root shod	rude shot	rot shooed	rod shoot
meal note	meat knoll	mole neat	moat kneel
feet made	feed mate	fate mead	fade meet
bait well	bail wet	bet wail	bell wait
sip hole	sill hope	soap hill	soul hip
moan bad	mode ban	man bode	mad bone
feel put	feet pull	full peat	foot peel
top pick	tock pip	tip pock	tick pop
rang sum	ram sung	rung Sam	rum sang
date roll	dale rote	dote rail	dole rate
tuck pin	ton pick	tick pun	tip puck
duel tine	dune tile	dial tune	dine tool

Acknowledgments

The research reported here was supported in part by NIH Grant HD 01994 to Haskins Laboratories. I thank Kristen Snow for helping with data collection and analysis and George Wolford for commenting on a draft of the manuscript.

References

- B.J. Baars and M. Motley (1976), "Spoonerisms as sequencer conflicts: Evidence from artificially elicited errors", *Amer. Psychology*, Vol. 89, pp. 467-474.
- B.J. Baars, M. Motley and D. MacKay (1975), "Output editing for lexical status from artificially elicited slips of the tongue", *J. Verbal Learning and Verbal Behavior*, Vol. 14, pp. 382-391.
- P. Carney and K. Moll (1971), "A cinefluorographic investigation of fricative-consonant vowel coarticulation", *Phonetica*, Vol. 23, pp. 193-201.
- R. Carter and J. Bradshaw (1984), "Producing 'Spoonerisms' on demand: Lexical, phonological and orthographic factors in a new experimental paradigm", *Speech Communication*, Vol. 3, pp. 347-360.
- E. Carterette and M. Jones (1974), *Informal speech* (University of California Press, Berkeley).
- N. Chomsky and M. Halle (1968), *The Sound Pattern of English* (Harper and Row, New York).
- G.N. Clements and S.J. Keyser (1983), *CV Phonology: A Generative Theory of the Syllable* (MIT Press, Cambridge, MA).
- A. Cutler, "The reliability of speech error data", *Linguistics*, Vol. 9, pp. 561-582.
- R. Daniloff and R. Hammarberg (1973), "On defining coarticulation", *Phonetica*, Vol. 1, pp. 239-248.
- G. Dell (1980), Phonological and lexical encoding in speech production: An analysis of naturally occurring and experimentally-elicited speech errors, PhD thesis, University of Toronto.
- G. Dell (1986), "A spreading-activation theory of retrieval in speech production", *Psychological Review*, Vol. 93, pp. 283-321.
- G. Dell and P. Reich (1981), "Stages in sentence production: An analysis of speech error data", *J. Verbal Learning and Verbal Behavior*, Vol. 20, pp. 611-629.
- C. Ewan (1982), "The internal structure of complex segments", in *The Structure of Phonological Representations*, ed. by H. van der Hulst and N. Smith (Foris, Dordrecht, The Netherlands) Vol. 2, pp. 27-68.
- C.A. Fowler (1983), "Converging sources of evidence for spoken and perceived rhythms of speech: Cyclic production of vowels in sequences of monosyllabic stress feet", *J. Experimental Psychology: General*, Vol. 112, pp. 386-412.
- W.N. Francis and H. Kučera (1982), *Frequency Analysis of English Usage: Lexicon and Grammar* (Houghton-Mifflin, Boston).
- V. Fromkin (1973), *Speech Errors as Linguistic Evidence* (Mouton, The Hague).
- M. Garrett (1975), "The analysis of speech production", in *The Psychology of Learning and Motivation* ed. by G. Bower (Academic Press, New York) Vol. 9, pp. 133-177.
- M. Garrett (1976), "Syntactic processes in sentence production", in *New Approaches to Language Mechanisms*, ed. by R.J. Wales and E.C. Walker (North-Holland, Amsterdam) pp. 231-255.
- M. Garrett (1980), "Levels of processing in speech production", in *Language Production*, ed. by B. Butterworth (Academic Press, London) Vol. 1, pp. 177-220.
- J. Hoard (1971), "Aspiration, tenseness and syllabication in English", *Language*, Vol. 47, pp. 133-140.
- L. Hultzen, J. Allen and M. Miron (1964), *Tables of Transitional Frequencies of English Phonemes* (University of Illinois Press, Urbana).
- J. Kupin (1979), Tongue twisters as a source of information about speech production, PhD Thesis, University of Connecticut.
- P. Ladefoged (1982), *A Course in Phonetics*, second ed. (Harcourt Brace Jovanovich, New York).
- A. Levitt and A. Healy (1985), "The roles of phoneme frequency, similarity and availability in the experimental elicitation of speech errors", *J. Memory and Language*, Vol. 24, pp. 717-733.
- B. Lindblom, J. Lubker, T. Gay, B. Lyberg, P. Branderud and K. Holmgren, The concept of target and speech timing, in: *In honor of Ilse Lehiste*, ed. by R. Channon and L. Shockey (Foris, Providence, RI), in press.
- B. Lindblom and K. Rapp (1973), "Some regularities of spoken Swedish", *Papers in Linguistics from the University of Stockholm*, Vol. 21, pp. 1-59.
- W. Marslen-Wilson and A. Welsh (1978), "Processing interactions and lexical access during word recognition in continuous speech", *Cognitive Psychology*, Vol. 10, pp. 29-67.
- D.G. MacKay (1971), "Stress pre-entry in motor systems", *Amer. J. Psychology*, Vol. 84, pp. 35-51.

- D.G. MacKay (1978), "Speech errors inside the syllable", in *Syllables and Segments*, ed. by A. Bell and J. Hooper (North-Holland, Amsterdam) pp. 201-213.
- M. Motley and B. Baars (1975), "Encoding sensitivity to phonemic markedness and transitional probability: Evidence from Spoonerisms", *Human Communication Res.*, Vol. 2, pp. 351-361.
- S. Ohman (1966), "Coarticulation in VCV syllables; Spectrographic measurements", *J. Acoust. Soc. Am.*, Vol. 39, 151-168.
- A. Prince (1980), "A metrical theory for Estonian quantity", *Linguistic Inquiry*, Vol. 11, 511-562.
- A.H. Roberts (1965), *A Statistical Analysis of American English*, (Mouton, The Hague).
- S. Shattuck-Hufnagel (1979), "Speech errors as evidence for a serial ordering mechanism in speech production", in *Sentence Production: Psycholinguistic Studies Presented to Merrill Garrett*, ed. by W.E. Cooper and E.C. Walker (Lawrence Erlbaum Associates, Hillsdale, New Jersey) pp. 295-342.
- S. Shattuck-Hufnagel (1983) "Sublexical units and suprasegmental structure in speech production planning", in *The Production of Speech*, ed. by P.F. MacNeilage (Springer, New York) pp. 109-136.
- S. Shattuck-Hufnagel and D. Klatt (1979), "The limited use of distinctive features and markedness in speech production: Evidence from speech-error data", *J. Verbal Learning and Verbal Behavior*, Vol 18, pp. 41-55.
- R. Trieman (1985), "Onsets and rimes as units in spoken syllables: Evidence from children", *J. Experimental Child Psychology*, Vol. 35, pp. 161-181.
- R. Trieman (1986), "The division between onsets and rimes in English syllables", *J. Memory and Language*, Vol. 25, pp. 476-491.
- M. van den Broecke and L. Goldstein (1980), "Consonant features in speech errors", in *Errors in Linguistic Performance: Slips of the Tongue, Ear, Pen and Hand*, ed. by V. Fromkin (Academic Press, New York) pp. 47-66.