LANGUAGE AND COGNITIVE PROCESSES, 1999, *14* (2), 173–200

Phonological Segments and Features as Planning Units in Speech Production

Ardi Roelofs
School of Psychology, University of Exeter, UK

The author reports four experiments that examined phonological processes in spoken word production. A form-preparation paradigm was applied to the question of whether phonological features can be preplanned to facilitate spoken word production. In Experiment 1, monosyllabic words were produced in sets different in form, or in sets sharing either the initial segment or initial segments differing only in voicing. Only shared initial segments yielded facilitation. A similar pattern of results was observed when the sets were matched for the following vowel (Experiment 2), when words were produced in response to pictured objects (Experiment 3), and when place of articulation rather than voicing was manipulated (Experiment 4). The special status of identity suggests that segments are planning units independent of their features. The results are explained in terms of the WEAVER model of word-form encoding, in which a serial encoding of segments is followed by a parallel activation of features. A WEAVER simulation of the experiments is presented which supports these claims.

INTRODUCTION

Theories of speech production differ in their claims about the representation and manipulation of features and segments in planning utterances. According to featural theories, phonological segments are represented by nothing but their features. For example, the segment /b/ is represented and manipulated in terms of features such as [+ voiced], [+ labial], [– nasal],

---

Requests for reprints should be addressed to Ardi Roelofs, School of Psychology, University of Exeter, Washington Singer Laboratories, Perry Road, Exeter EX4 4QG, UK. E-mail: a.roelofs@ex.ac.uk

I am indebted to Daniëlle van den Brink for her help in preparing and running the experiments, to Antje Meyer and Pim Levelt for discussion, and to the anonymous reviewers for helpful comments on a version of this article. The experiments were conducted at the Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands, when I was a staff member of the Institute.

and so forth, but there is no independent representation of the segment that is manipulated in production. In planning the form of a word, morphemes are mapped onto the features that make up the sound of the morpheme. For example, the morpheme <bear> is mapped onto the features [+ voiced], [+ labial], [– nasal], and so forth. The mapping of a morpheme onto its features is typically taken to unfold over time, such that, roughly, the features of one segment are activated at a time. In this view, a segment is an emergent property of the featural representation of a word. The featural position has been advanced by a number of researchers in the production literature (e.g. Mowrey & MacKay, 1990) and has recently been developed in some computational detail by Dell, Juliano and Govindjee (1993; Dell & Juliano, 1996).

In contrast, according to segmental theories, phonological segments have their own abstract representation in memory, which is manipulated in planning utterances independent of their features. In this view, a segment recodes a set of features into a representation that refers to the features but that does not contain the features as a proper part, which is traditionally called a *chunk* (cf. Miller, 1956; Simon, 1969). Memory contains representations that refer to segments qua segments. In planning the form of a word, morphemes are mapped via these segment representations onto features. For example, the morpheme <bear> is mapped via the segment /b/ onto the features [+ voiced], [+ labial], [– nasal], and so forth. This account of the representation of segments and their role in production is advanced by Dell (1986, 1988), Levelt (1989), Roelofs (1994, 1997a), Stemberger (1982), and Shattuck-Hufnagel (1979), among others.

One of the main motivations for assuming an explicit representation and manipulation of segments has come from speech errors. Most of the phonological speech errors concern the anticipation, perseveration, exchange, shift, addition, or deletion of a single segment. It has been estimated that about 70 to 90% of the phonological slips of the tongue concern a single segment (Dell, 1986; Stemberger, 1982). Nooteboom's (1969) analysis of Dutch phonological errors, which is typical, yielded 89% single segment errors, 7% consonant cluster errors, and 4% for the remainder. Furthermore, the segmental environment influences the likelihood of a speech error but the feature environment does not. Stemberger (1990) looked at speech errors in natural speech and observed that shared segmental environment (the segments immediately after the segment in error in one word and the source of that segment in another word) has an effect only if the neighbouring segments are exactly identical. There is no effect if even a single feature is different. These findings suggest that segments themselves are units that are manipulated during the planning of an utterance.

Assigning a role to segments does not deny a role for features in the planning process. In most phonological errors, the target and the intruding segment differ in only one or two features (Shattuck-Hufnagel, 1979, 1983; Shattuck-Hufnagel & Klatt, 1979; Stemberger, 1982). Thus, slips of the tongue are clearly sensitive to the featural make-up of the segments in error. In the classification of speech errors, a distinction is made between contextual and non-contextual errors (also called plan-internal and plan-external errors). Contextual errors concern misorderings of intended elements. In non-contextual errors, no source for the error is evident in the context. For example, in the contextual error "*g*lear *p*lue sky" for "clear blue sky" (from Fromkin, 1971), the voicing feature of the /k/ (voiceless) and the /b/ (voiced) is reversed, which creates two new segments, namely /g/ (voiced) and /p/ (voiceless). In the non-contextual error "what do you mean. I'll loo*v*e it?" for "I'll lo*s*e it?" (from Stemberger, 1985), where the /z/ of *lose* is replaced by /v/, the target and intruder also differ in one feature: /z/ has the feature [+ coronal], whereas /v/ has instead the feature [+ labial]. Both contextual and non-contextual errors exhibit a feature similarity effect, although Stemberger (1985) observed that the effect is greater in non-contextual substitutions (with about 90% of consonant errors showing a difference in just one feature) than in contextual errors (about 70% of the error segments is one feature away).

Since for both contextual and non-contextual errors, the norm is for similar segments to interact, some researchers have conjectured that perhaps most segmental errors might be construed as feature errors (e.g. Garrett, 1988). This view meets with a difficulty, however. Shattuck-Hufnagel and Klatt (1979) have argued that genuine feature slips are very rare. Unambiguous feature slips concern errors in which new segments are formed, as in "*g*lear *p*lue sky" where the /k/ has turned into /g/ and the /b/ into /p/ (the /g/ nor the /p/ were part of the intended utterance) rather than involving an exchange of the segments /k/ and /b/. The /k/ and /b/ differ in more than one feature, so that the exchange of the voicing feature creates two new segments, /g/ and /p/. By contrast, an error like "*d*arn *b*ore" for "barn door" (Dell, 1988), where the segments differ by only one feature, can readily be explained both as a single feature exchange and as a segment exchange. That is, both an exchange of the place of articulation and an exchange of full segments leads to substituting /d/ for /b/ and /b/ for /d/. Thus, only errors in which the interacting segments differ by more than one feature (so that a feature exchange creates a pair of new segments) provide the opportunity for detecting unambiguous single-feature movements. Shattuck-Hufnagel and Klatt observed that their MIT corpus contained 70 cases where the interacting segments differed by more than one feature and where a feature exchange would result in a legal English feature configuration. However, only three of these 70 cases could be classified as unambiguous feature exchanges, that is, errors in which new segments were formed. Thus, errors involving multiple features (resulting in an exchange of full segments on the surface) occurred far more frequently than those involving the movement of a single feature did. This suggests that some higher-order unit underlies these errors. The fact that most errors are the size of a segment has therefore widely been accepted as support for the idea that segments rather than features are the prime units that are manipulated, and may get misordered or replaced, during the planning of the phonological shape of an utterance (e.g. Shattuck-Hufnagel, 1979, 1983; Shattuck-Hufnagel & Klatt, 1979; Stemberger, 1982).[1]

To explain the predominance of segment-sized slips, the influence of feature similarity, and the effect of shared segmental environment on the error rate, researchers have advanced interactive-activation network models (e.g. Dell, 1986; Stemberger, 1982). These models assume that lexical memory contains segment nodes that are bidirectionally connected to morpheme nodes and to nodes for the features of the segments. In planning an utterance, segments are activated and selected. Selected segments are inserted into slots of independently generated frames that determine the serial order of the segments (Dell, 1986). Phonological errors are conceived of as failures in selecting the segments of a word as fillers for the slots, for instance, due to noisy activation levels in the network. Segments that share most of their features are most likely to be involved in a selection error, due to the backward spreading of activation from features to segments. For example, if /b/ is to be produced, a segment like /p/ receives feedback from most features of the target /b/, whereas a segment like /h/ receives feedback from only a few features of /b/. Thus, /p/ will have a higher level of activation than /h/, and /p/ will be more likely than /h/ to be substituted for the target /b/. Shared segmental environment facilitates an error (Dell, 1984; Stemberger, 1990) due to feedback from shared segments to morphemes and from secondarily activated morphemes to competing segments. For example, in planning *cat*, the /f/ may become activated due to the vowel shared between *cat* and *fan*. The fact

[1] Levelt, Roelofs, and Meyer (in press) have suggested that features are chunked into segments in the course of learning a language. This would explain Stemberger's (1989) finding that young children are more likely to make feature errors than adults are. However, Wijnen (1992) observed no difference between the errors of children and adults. Perhaps the difference between these studies is due to the fact that Wijnen collected errors of children at a later age than Stemberger. Stemberger analysed the errors of two children from age 1;0 to 5;11 and 1;0 to 3;4, whereas Wijnen examined the errors of two children from age 2;4 to 3;0 and from 3;0 to 3;10.

that non-identical but similar segmental neighbours do not have any effect on error rates (Stemberger, 1990) is attributed to the longer path. For example, the fact that the vowels in *cat* and *fin* are both [+ lax] and [+ front] would have no effect. Activation has to go from the target morpheme <cat> via the segment /æ/ to its features and via some of these features via the segment /ɪ/ to the morpheme <fin>, which should activate the competing segment /f/. This longer path supposedly attenuates any effects that might be present, although it would not categorically rule them out.

The segmental view has been the received view in language production research for the last decades (cf. Levelt, 1989). Recently, however, the featural view has gained some ground and become a more believable alternative. Dell et al. (1993) made it plausible that the segment and similarity effects may also arise in a system without an independent representation and selection of segments. In their view, there are no segments that serve as fillers for the slots of frames that determine serial order. Dell and colleagues proposed a featural model developed within the theoretical framework for speech production advanced by Jordan (1986, 1990). Their model is a PDP model in which morphemes are connected via a layer of hidden units to features. Activation of a morpheme is passed through weighed connections to the hidden layer, and finally to the output layer, which contains one unit for each feature in the language. When fully trained, the model activates the features corresponding to each segment of the morpheme, one segment at a time. In computer simulations that were run by Dell et al. (1993), errors were generated either by incomplete training or by adding noise to the connections. All errors thus obtained were non-contextual errors. Dell et al. argued that, to a first approximation, the resulting errors were of the right size, namely most of the errors could be classified as segment-sized slips. However, the errors occur because the wrong features are generated, not because one segment as a whole replaces another. Furthermore, the errors were argued to be properly sensitive to feature similarity in that most errors were one or two features off.

However, Dell et al. (1993) observed that their featural model fell short of the existing segmental models in one important respect. The model was able to produce non-contextual errors, but it could not produce contextual errors. Dell et al. discussed a number of ways in which contextual influences may be introduced in future versions of their model. For example, the input pattern for one morpheme may have an influence on the input pattern for the next morpheme. In this way, anticipation and perseveration errors may perhaps be produced, although this remains to be confirmed by simulation. Dell and colleagues doubted, however, whether their model would be able to produce exchange errors. “A perturbation in the model’s output that causes a later sound to replace an earlier one would not lead to a corresponding perturbation for the originally replaced sound to then substitute for the original replacing sound'' (Dell & Juliano, 1996, p. 355). Thus, exchange errors cannot occur in the model. However, in order to be able to explain contextual errors like ''*g*lear *p*lue sky'' in terms of an exchange of the voicing feature of the /k/ of *clear* and the /b/ of *blue*, featural models should allow that features can rather freely trade places.

To conclude, one of the principal arguments for assuming an explicit representation and manipulation of segments has come from speech errors. However, featural models might have the potential to account for the error evidence if these models are developed further such that they contain independently movable features, which is what current models (i.e. Dell et al., 1993) are lacking. Of course, it should also be demonstrated that errors involving multiple features occur more frequently in such a model than single feature errors, to account for the observation of Shattuck-Hufnagel and Klatt (1979). Also, the fact that only identical but not similar neighbouring segments have an effect on error rates (Stemberger, 1990) should be explained. Before exploring the potential of such a more advanced featural model, however, it would be worthwhile to see if one can first obtain empirical support for this featural position. Speech errors are rare events, and their properties do not necessarily reflect the norm. The aim of the present paper is to gain further, independent evidence about the role of features and segments in planning speech. The paper reports a series of chronometric experiments contrasting the featural and segmental view.

The paper is organised as follows. To provide a detailed theoretical framework for phonological encoding, I review the WEAVER model (Roelofs, 1994, 1996a, 1997a), which is a comprehensive, computational model of word-form encoding in speech production. WEAVER falls into the class of segmental models. The model is unique in the production literature in that it provides an account of encoding times, which have also been collected in the experiments in the present paper. Although WEAVER has not been specifically designed to account for speech errors, the model is compatible with the evidence for segment-sized slips, the influence of feature similarity, and the effect of the segmental environment on error rates (Levelt, Roelofs, & Meyer, in press; Roelofs, 1997a). Next, I review a number of key chronometric findings concerning segments in phonological encoding, thereby describing the form-preparation paradigm that will be used in all the experiments in the present paper, and I discuss how WEAVER accounts for these key findings. Then, I report a series of four new experiments with this paradigm that tested, in a general fashion, predictions derived from the featural and segmental views outlined above. Next, I show by computer simulations that the WEAVER model accounts for the experimental results. The paper ends with a discussion of the theoretical implications of the findings.

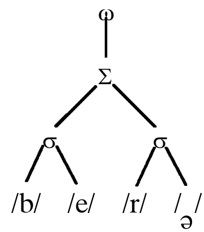
FEATURES AND SEGMENTS IN WEAVER

To set the stage, it is important to indicate where features and segments fit into the cognitive system that plans the production of speech. A widely accepted view holds that the planning of speech proceeds through conceptualisation and formulation, which is followed by articulation (e.g. Dell, 1986; Garrett, 1975; Kempen & Hoenkamp, 1987; Levelt, 1989). Conceptualisation processes map a communicative intention onto a message, which indicates the conceptual information to be verbally expressed to reach a speaker's communicative goal. Formulation processes activate and select words for the message concepts, which is called lexical access, and plan a syntactic and a morpho-phonological structure. The result is an articulation program for the utterance, which, when executed by articulation processes, yields overt speech.

It is generally assumed that lexical access to a single word consists of two major steps, namely lemma retrieval and word-form encoding (Butterworth, 1989; Dell, 1986; Garrett, 1975; Kempen & Huijbers, 1983; Kempen & Hoenkamp, 1987; Levelt, 1989, 1992; Levelt et al., in press; Meyer, 1990, 1991; Roelofs, 1992, 1993, 1996a; but see Caramazza, 1997). During lemma retrieval, a message concept is used to retrieve the lemma of a corresponding word from long-term memory. In WEAVER, lemmas represent the syntactic properties of words, crucial for their use in sentences. For example, the lemma of the Dutch word *beer* (English *bear*) specifies that it is a noun and that its grammatical gender is non-neuter. The lemma retriever makes these properties available for syntactic encoding.

Word-form encoding comprises three major steps, namely morphological, phonological, and phonetic encoding (cf. Dell, 1986; Levelt, 1989, 1992). The morphological encoder takes the lemma and abstract morphosyntactic parameters such as "singular" or "plural" as input and produces as output one or more morphemes, for instance, a noun stem (e.g. <beer>) and a plural suffix (<en>). The phonological encoder recovers the corresponding segments from memory and uses them to construct a phonological word representation. Finally, the phonetic encoder generates a more detailed and context-dependent form representation, which specifies the articulatory tasks to be achieved (Goldsmith, 1990; Kenstowicz, 1994).

WEAVER realises this global view of word-form encoding, which is shared by many models (Dell, 1986, 1988; Levelt, 1989, 1992), in a specific manner. The mental lexicon is conceived of as a network of information that is accessed by spreading activation. In form encoding, activation starts at a lemma node and then spreads out through the network. Activation of parts of the network triggers production rules that select nodes and integrate them into a phonetic plan. Production rules are condition/action pairs (Anderson, 1983; Newell, 1990). Morphological productions select the morpheme nodes that encode a selected lemma node and its abstract morpho-syntactic parameters such as number (e.g. <beer> and <en> for the lemma of *beer* plus ‘‘plural’’). Phonological products select the segment nodes linked to the morpheme nodes and syllabify the segments to construct phonological word representations (cf. Levelt, 1992). These representations specify the syllables and the stress pattern, for example,



which describes the plural form *beren* as a phonological word (ω) consisting of a trochaic foot (Σ), whose first, stressed syllable (σ) has as onset /b/ and as nucleus /e/ and whose second syllable has as onset /r/ and nucleus /ə/. Like weaving a fabric, the encoding has a certain direction in that phonological forms are computed from left to right. Finally, phonetic productions select syllable-based articulatory programs for each of the phonological syllables in the phonological word representation (cf. Levelt & Wheeldon, 1994). Selection of articulatory programs takes place as soon as the computation of a phonological syllable is completed. For example, for *beren*, the programs

[be] and [rə]

are recovered. During this final step, the featural make-up of the syllables becomes available. The features of the segments in a syllable are accessed in parallel.

WEAVER provides for a suspend-resume mechanism that supports incremental generation of phonetic plans. Incremental production means that encoding processes can be triggered by a fragment of their characteristic input (Levelt, 1989). For example, syllabification of a word can start as soon as the first few segments are available. The resulting partial representation can be buffered until the missing segments are available and syllabification can continue. Thus, when given partial information, computations are completed as far as possible, after which they are put on hold. When given further information, the encoding processes continue from where they stopped. Importantly, buffered forms in WEAVER are only expandable to the right.

THE FORM-PREPARATION PARADIGM

In investigating the issue of segments and features in phonological encoding, I employ the form-preparation paradigm that has been developed by Meyer (1990, 1991). A big advantage of this paradigm compared to a more widely used paradigm such as picture-word interference (Meyer & Schriefers, 1991), where spoken primes are presented during picture naming, is that in the form-preparation paradigm the responses do not have to be names of depictable entities. This removes a strong constraint on the choice of materials. Speakers in Meyer's experiments first had to learn small sets of word pairs such as *fruit—melon*, *iron—metal*, and *grass—meadow* (the example is in English, but the original materials were in Dutch). During the following test phase, they had to produce the second word of a pair (e.g. *metal*) upon visual presentation on a computer screen of the first word (*iron*), called the prompt. On each trial, one of the prompts was presented. The order of prompts across trials was random. The production latency, the interval between prompt onset and speech onset, was the main dependent variable. Each experiment contained two types of sets, called homogeneous and heterogeneous sets. In a homogeneous set, the response words shared part of their form, for example, the first syllable (ME*lon*, ME*tal*, MEA*dow*) or the second syllable (*po*CKET, *ti*CKET, *ra*CKET). In the heterogeneous sets, the response words were unrelated in word form. Regrouping the pairs from the homogeneous sets created the heterogeneous sets. Therefore, each word pair was tested both under the homogeneous and the heterogeneous condition, and all uncontrolled item effects were kept constant across these conditions.

Meyer found shorter production latencies in homogeneous than in heterogeneous sets. However, this difference was only obtained when the response words in homogeneous sets shared one or more word-initial segments, but not when they shared word-final segments. For example, a preparation effect was obtained for the begin-related homogeneous set that included ME*lon*, ME*tal*, and MEA*dow*, but not for the end-related homogeneous set that included *po*CKET, *ti*CKET, and *ra*CKET. The size of the preparation effect increased with the number of shared word-initial segments. Meyer's findings have been replicated not only for a variety of types of other morphologically simple words, but also for complex words and for phrasal constructions (Roelofs, 1996a,b, 1997a, 1998). Also, a similar sequential effect has been obtained with other experimental paradigms such as picture-word interference (Meyer & Schriefers, 1991), where spoken primes are presented during picture naming.

According to WEAVER, facilitation arises when participants prepare and buffer partial phonological representations of the response words before prompt presentation. That is, a response is prepared by an advance reduction of the degrees of freedom (i.e. a reduction of the number of choices to be made later), which appears to be a general phenomenon in preparing skilled actions (Kelso, 1982; Rosenbaum, 1991; Smyth & Wing, 1984). The confinement of the facilitatory effect to begin-related homogeneous sets reflects the suspend-resume mechanism that underlies the incremental planning of utterances.

Assume that the set of response words consists of *melon*, *metal*, and *meadow*. Before the beginning of a trial, the morphological encoder can do nothing, but the phonological encoder can construct the first phonological syllable $(\text{mɛ})_\sigma$, and the phonetic encoder can recover the first syllable-based articulatory program [mɛ]. As soon as a prompt (e.g. *iron*) is given, the morphological encoder will retrieve the associated target morpheme (for *iron* this is <metal>). Segmental spell-out makes all segments of this morpheme available including those of the second syllable, and the phonological and phonetic encoders can start to work on the second syllable. In the heterogeneous condition (*metal*, *racket*, etc.), nothing can be prepared before prompt presentation. There will be no morphological, phonological, or phonetic encoding. In the end-related homogeneous condition (*pocket*, *ticket*, *racket*) nothing can be done either. Although the segments of the second syllable are known, the phonological form cannot be computed because the missing segments precede the suspension point. In WEAVER, this means that after prompt presentation syllabification must restart from the first segment of the word, which amounts to restarting the whole process, like unravelling a woven fabric. Thus, a facilitatory effect for the homogeneous relative to the heterogeneous condition will only be obtained for begin-related response words, as empirically observed. Computer simulations of Meyer's (1990) experiments can be found in Roelofs (1994, 1997a).

RATIONALE BEHIND THE EXPERIMENTS

The experiments in the present paper examined the role of features and segments in planning speech. The target language was Dutch. Booij (1995) gives a description of the phonology of the Dutch language. All experiments tested whether preparation in homogeneous sets requires that the responses share their initial segments fully or whether it suffices that the initial segments share most of their features. Is a preparation effect only obtained in a "segmental" set that includes, for example, B*oat*, B*ird*, and Boy, or also in a "featural" set that includes Boat, Bird, and Paint, where the initial segments are the same except for one feature, namely voicing?

In the featural view, a preparation effect such as observed by Meyer (1990, 1991) would arise from the advance planning of features. When the sound of a morpheme is encoded sequentially (cf. Dell et al., 1993), as is illustrated in (1), preparing the features of a segment would yield facilitation both in the case of feature and of segment overlap.

(1)

features of segment *n* → features of segment *n* + 1

If features of segment *n* can be prepared before the beginning of a trial, they do not have to be encoded on the trial itself. Instead, less features of segment *n* have to be encoded (in sets with only feature overlap), or the encoding can go directly to segment *n* + 1 (in sets with full segment overlap), and facilitation should be obtained in both cases.

Note that with parallel rather than sequential encoding of the features of consecutive segments, a preparation effect is not to be expected. A preparation effect is only obtained if encoding segment *n* plus segment *n* + 1 takes more time than encoding segment *n* only, which need not hold if segments *n* and *n* + 1 are encoded in parallel. If the encoding time increases with the number of features that have to be specified, the model predicts more facilitation for the segmental than for the featural sets. This is because in segmental sets more features can be prepared than in featural sets. Most important, however, if feature preparation is the cause of preparation effects such as observed by Meyer (1990, 1991), both segment and feature sets should yield facilitation.

In contrast, WEAVER predicts that feature overlap alone should yield no preparation effect at all. In the model, segments are syllabified serially but their features may be accessed in parallel (once a phonological syllable is ready), as illustrated in (2).

(2)

segment *n*	→	segment *n* + 1
↓		↓
features of segment *n*		features of segment *n* + 1

According to the model, the preparation effect observed by Meyer (1990, 1991) emerges from the seriality of syllabifying segments in phonological encoding. Sharing full segments is crucial for advance planning in the model, because preparation consists of computing and buffering part of a representation that is made up of segments. Since computing this representation proceeds from the beginning of an utterance to its end, the advance computing of part of it yields facilitation. However, the features of the segments that make up a syllable in the phonological word may be accessed in parallel. Due to this parallelism, if segments are prepared at the level of phonological words, preparing their features will not yield additional facilitation. Consequently, the possibility for preparation in a feature-homogeneous set should be equivalent to that in a segment-heterogeneous set and a feature-heterogeneous set, namely preparation should not be possible. Thus, WEAVER predicts that the responses should be faster in the homogeneous condition with the segment overlap than in the three other conditions, which should not differ.

Experiment 1 compared the preparation effect for sets whose first segment is the same, such as B*oat*, B*ird*, and B*oy*, to that of sets whose initial segments differ in one feature, such as B*oat*, B*ird*, and P*aint*. When a preparation effect is only obtained for full segment overlap but not for feature overlap, this would support the claim of an explicit representation of segments. In contrast, when preparation is obtained both for segment and feature overlap, this would support a featural theory. Experiment 2 tested the same predictions, but now with disyllabic words sharing the first syllable fully (in the segment-homogeneous condition) and with disyllabic words whose first syllable is the same except that the syllable onset segment differs in one feature (in the feature-homogeneous condition). For example, the experiment compared the preparation effect for sets such as TE*nnis*, TE*rrace*, and TE*ddy* to that of sets such as DE*vil*, TE*nnis*, and TE*ddy*. Experiment 3 tested the same predictions by using pictures instead of written prompts. Thus, rather than producing the response-members of paired associates, speakers had to name pictures in the featural and segmental conditions. In the first three experiments, the critical feature was voicing. Experiment 4 tested the same predictions as in Experiments 1 to 3, but now by looking at segments differing in a place of articulation feature, namely whether the segment is pronounced labial or coronal. The greatest number of single-feature errors in slips of the tongue involve a change in place of articulation (Shattuck-Hufnagel & Klatt, 1979), thus place of articulation features appear to have the greatest mobility. Therefore, it should be tested whether place of articulation is special in that it can be prepared while other features cannot.

EXPERIMENT 1

Experiment 1 used monosyllables as the response members of paired-associates to test whether preparation requires that the responses share initial segments or whether it suffices that they share features only. The experiment compared the preparation effect for sets whose words share initial segments such as B*oat*, B*ird*, and B*oy* to that of sets of words whose initial segments are the same except for one feature, for example B*oat*, B*ird*, and P*aint*. In the latter set, /b/ and /p/ share all features except that /b/ is voiced and /p/ is voiceless. If a substantial preparation effect were obtained for both types of set, this would imply that sharing full segments is not necessary for preparation, which would support the featural view. In contrast, if a preparation effect is obtained only for segment overlap but not for feature overlap, this would imply that sharing full segments is necessary for preparation, which would support the segmental view.

Method

Participants. All experiments were conducted with paid participants from the pool of the Max Plank Institute. All participants were young adults and native speakers of Dutch. Each participant took part in only one experiment. Experiment 1 was carried out with 16 participants.

Materials and Design. The materials for all experiments were obtained from the Dutch part of the CELEX lexical database (Baayen, Piepenbrock, & Gulikers, 1995). All prompts and response words were nouns because suitable items were easiest to find in this word class. The materials consisted of two practice sets and 24 experimental sets of three prompt-response word pairs each. Each set was tested in a separate block of trials. There were 12 homogeneous and 12 heterogeneous sets. In six homogeneous sets, the response words shared the first consonant, and in the other six homogeneous sets the first segment shared all features except one. This independent variable, which had two levels (segments, features), will be called *level of overlap*. In the heterogeneous sets there was no such overlap. Following Meyer (1990), this independent variable—homogeneous *vs* heterogeneous sets—will be called *context*. The same prompt-response word pairs were tested in the homogeneous and heterogeneous condition. Only their combinations into sets differed.

The critical segments in the experiment were the minimal pairs /d/ (voiced) and /t/ (voiceless), /b/ (voiced) and /p/ (voiceless), and /v/ (voiced) and /f/ (voiceless). All responses occurred in all conditions. In three homogeneous and the corresponding heterogeneous sets of the segment condition, the first segment was the voiced member of a minimal pair. In the remaining three homogeneous and heterogeneous sets of the segment condition, the first segment was the voiceless member of a minimal pair. In three homogeneous and the corresponding heterogeneous sets of the feature condition, two of the responses had the voiced member of a minimal pair as first segment and the third response had the voiceless member as first segment. In the remaining three homogeneous and heterogeneous sets, two of the responses started with the voiceless member and the third response started with the voiced member. Table 1 lists the materials of Experiment 1.

Each participant was tested once on each set. The order of the sets was rotated across participants in the following way. Eight participants (groups A and B) were first tested on the homogeneous sets and then on the heterogeneous sets. For the remaining eight participants (groups C and D), the order of testing the homogeneous and heterogeneous conditions was reversed. The participants of groups A and C were first tested on the sets of the segment condition, then on the sets of the feature condition. For

TABLE 1
Response Sets of Experiment 1

Level of overlap	*Context*	*Set*
Segments	Homogeneous	Set 1: been, bos, baard (leg, forest, beard)
		Set 2: damp, dolk, deur (steam, dagger, door)
		Set 3: vel, vat, vork (skin, barrel, fork)
		Set 4: pand, pet, peer (forfeit, cap, pear)
		Set 5: touw, teil, thee (rope, tub, tea)
		Set 6: fiets, fort, film (bike, fort, film)
Segments	Heterogeneous	Set 7: damp, vork, been
		Set 8: vel, bos, dolk
		Set 9: baard, deur, vat
		Set 10: thee, fiets, pand
		Set 11: film, peer, touw
		Set 12: pet, teil, fort
Features	Homogeneous	Set 13: bos, been, pet
		Set 14: dolk, damp, teil
		Set 15: vork, vel, film
		Set 16: pand, peer, baard
		Set 17: thee, touw, deur
		Set 18: fort, fiets, vat
Features	Heterogeneous	Set 19: been, dolk, film
		Set 20: vel, pet, damp
		Set 21: teil, vork, bos
		Set 22: pand, touw, vat
		Set 23: fiets, baard, thee
		Set 24: deur, fort, peer

Note. English translations are given in parentheses. The author can provide a listing of the materials including the prompt words.

participants of groups B and D, the order of testing these conditions was reversed. Each of the three prompt-response word pairs of a set was tested four times within each block of trials. In all experiments, the order of testing the word pairs was random, except that immediate repetitions of pairs were excluded. A different random order was generated for each block and each participant.

Procedure and Apparatus. In all experiments, the participants were tested individually. They were seated in a quiet room in front of a computer screen (NEC Multisync30) and a microphone (Sennheisser ME40). After the participant had read the instructions, two practice blocks (with the same structure as the experimental blocks but with different items) were administered followed by the 24 experimental blocks. In the learning phase before each block, the three word pairs of a set were presented on the screen. As soon as the participant indicated having studied the pairs sufficiently, the experimenter started the test phase. The structure of a trial was as follows. First, the participant saw a warning signal (an asterisk) for 500 msec. Next, the screen was cleared for 500 msec, followed by the display of the prompt for 1500 msec. The asterisk and prompt were presented in white on a black background. Finally, before the start of the next trial there was a blank interval of 500 msec. Thus, the total duration of a trial was 3 sec. A Hermac 166 MHz computer controlled the experiment.

Analyses. The error coding and statistical analyses were the same for all experiments. After each trial, the experimenter coded the response for errors. Experimental sessions were recorded on audio-tape by a Sony DTC55 DAT recorder. The recordings contained the participants' speech and tones indicating the onset of the prompt (1 kHz) and the moment when the voice key was triggered (2.5 kHz). The experimenter heard these tones during the experiment via closed headphones. The recordings were consulted after the experiment whenever the experimenter was uncertain about whether a response was fully correct. Five types of incorrect responses were distinguished: wrong response words, wrong pronunciation of the words, a dysfluency (stuttering, within-utterance pauses, repairs), triggering of the voice key by non-speech sounds (noise in the environment or smacking sounds participants produced with the lips or tongue), and failures to respond within 1500 msec after prompt presentation. Incorrect responses were excluded from the statistical analysis of the production latencies. In all experiments, the production latencies and errors were submitted to by-participant and by-item analyses of variance with the crossed variables context and level of overlap. Both variables were tested within participants and within items. For the errors, no main effect or interaction was significant both by participants and items in any of the experiments. Therefore, I report the error means but not the test statistics.

Results

Table 2 gives the mean production latencies, error percentages, and preparation effects for Experiment 1. Table 2 shows that the production latencies were faster for the homogeneous sets in which the segments were shared than for the other three conditions, which did not differ much. A facilitatory effect from homogeneity was obtained for the sets with the segment overlap, but there was no effect at all for the sets with the feature overlap. There was an interaction between context and level of overlap [$F_1(1, 15) = 3.97$, $MSE = 782$, $P = 0.065$; $F_2(1, 17) = 6.96$, $MSE = 502$, $P = 0.017$]. The simple effect of context was significant for the segment overlap [$F_1(1, 15) = 9.74$, $MSE = 468$, $P = 0.007$; $F_2(1, 17) = 7.99$, $MSE = 642$, $P = 0.012$], but there was no effect for the feature overlap [$F_1(1, 15) < 1$, $MSE = 1150$, $P = 0.743$; $F_2(1, 17) < 1$, $MSE = 383$, $P = 0.547$]. The simple effect of level of overlap was significant by items for the homogeneous condition [$F_1(1, 15) = 1.02$, $MSE = 2618$, $P = 0.329$; $F_2(1, 17) = 6.77$, $MSE = 444$, $P = 0.019$], but there was no effect for the heterogeneous condition [$F_1(1, 15) < 1$, $MSE = 1791$, $P = 0.530$; $F_2(1, 17) = 1.59$, $MSE = 524$, $P = 0.225$]. To conclude, a preparation effect was only obtained when the response words shared an initial segment, but not when all but one of the features of the first segment were shared.

EXPERIMENT 2

The second experiment tested the same predictions as Experiment 1 with new materials and participants. There were two main reasons for such a replication. First, WEAVER predicts that the homogeneous condition with the segment overlap should be faster than the three other conditions, which should not differ. The predicted effect of context was confirmed in Experiment 1, but the predicted effect of level of overlap was significant by

TABLE 2
Mean Production Latencies (M, in msec), Error Percentages (E%), and Preparation Effects per Context and Level of Overlap for Experiment 1

	Context					
	Homogeneous		*Heterogeneous*		*Preparation*	
Level of overlap	*M*	*E%*	*M*	*E%*	*M*	*E%*
Segments	674	1.5	698	1.5	– 24	0.0
Features	692	1.7	688	1.0	4	0.7

items only. Thus, it is important to see whether the findings can be replicated in a new experiment. Secondly, Experiment 1 showed a clear difference in preparation effect between exact identity (shared segments) and close similarity (shared features). However, the reason why the feature overlap yielded no preparation at all was perhaps that the onset consonants differed more than intended due to co-articulation or assimilation. The vowels following the onset consonants were not the same. Anticipating the pronunciation of the vowels in the initial consonants may have reduced the similarity between the consonants. To match for the following vowel, Experiment 2 used disyllabic words sharing the first syllable fully (in the segment-homogeneous condition) and disyllabic words whose first syllable differs in one feature only (in the feature-homogeneous condition). For example, the experiment compared the preparation effect for sets such as TE*nnis*, TE*rrace*, and TE*ddy* to that of sets such as DE*vil*, TE*nnis*, and TE*ddy*. Thus, Experiment 2 presents a stronger test of the segment and featural views than the previous experiment.

Method

The experiment was carried out with 16 participants. The method was the same as in Experiment 1. The critical segments were again the minimal pairs /d/ (voiced) and /t/ (voiceless), /b/ (voiced) and /p/ (voiceless), and /v/ (voiced) and /f/ (voiceless). However, now the vowels that followed these segments within the syllable were the same. The critical syllables were /de/ and /te/, /ba/ and /pa/, and /vi/ and /fi/. So, in homogeneous sets, the initial syllables were fully shared (in segmental sets) or differed by one feature (in featural sets). All responses occurred in all conditions. The Appendix lists the materials.

Results

Table 3 gives the mean production latencies, error percentages, and preparation effects for Experiment 2. Table 3 shows that the production

TABLE 3
Mean Production Latencies (M, in msec), Error Percentages (E%), and Preparation Effects per Context and Level of Overlap for Experiment 2

	Context					
	Homogeneous		*Heterogeneous*		*Preparation*	
Level of overlap	*M*	*E%*	*M*	*E%*	*M*	*E%*
Segments	684	2.9	714	2.1	– 30	0.8
Features	717	3.1	722	2.9	– 5	– 0.2

latencies were faster for the homogeneous sets in which the segments were shared than for the other three conditions, which did not differ much. A facilitatory effect from homogeneity was obtained for the sets with the segment overlap, but there was no effect at all for the sets with the feature overlap. The interaction between context and level of overlap was significant [$F_1(1, 15) = 6.97$, $MSE = 332$, $P = 0.019$; $F_2(1, 17) = 6.55$, $MSE = 397$, $P = 0.020$]. The simple effect of context was significant for the segment overlap [$F_1(1, 15) = 14.67$, $MSE = 464$, $P = 0.002$; $F_2(1, 17) = 24.33$, $MSE = 315$, $P = 0.001$], but there was no effect for the feature overlap [$F_1(1, 15) < 1$, $MSE = 593$, $P = 0.559$; $F_2(1, 17) < 1$, $MSE = 485$, $P = 0.493$]. The simple effect of level of overlap was significant for the homogeneous condition [$F_1(1, 15) = 6.21$, $MSE = 1361$, $P = 0.025$; $F_2(1, 17) = 26.16$, $MSE = 363$, $P < 0.001$], but there was no effect for the heterogeneous condition [$F_1(1, 15) < 1$, $MSE = 686$, $P = 0.376$; $F_2(1, 17) = 1.68$, $MSE = 382$, $P = 0.212$]. To conclude, a preparation effect was obtained only when the response words shared the segments of a syllable, but not when all but one of the features of the syllable were shared. The production latencies were fastest in the homogeneous condition in which the segments were shared and the latencies were slower in the other three conditions, which did not differ statistically.

EXPERIMENT 3

The retrieval of the sound of a word using paired-associates may differ from the retrieval in normal speech production, where words are generated for concepts to be expressed. In order to exclude that the results in Experiments 1 and 2 were due to some peculiar aspect of paired-associate memory retrieval, Experiment 3 tested the same predictions now by using pictures instead of written prompts. The responses were the names of the pictures. Thus, rather than producing the response-members of paired associates, speakers had to name pictures in feature and segment homogeneous and heterogeneous sets.

Method

The experiment was carried out with 16 participants. The method was the same as in Experiment 1. The critical segments in the experiment were again the minimal pairs /d/ (voiced) and /t/ (voiceless), /b/ (voiced) and /p/ (voiceless), and /v/ (voiced) and /f/ (voiceless). Now, these segments were part of picture names. Again, all responses occurred in all conditions. The Appendix lists the materials of Experiment 3.

Results

Table 4 gives the mean production latencies, error percentages, and preparation effects for Experiment 3. It shows that the production latencies were faster for the homogeneous condition in which the segments were shared than for the other three conditions, which did not differ much. A facilitatory effect from homogeneity was obtained for the sets with the segment overlap, but there was no effect for the sets with the feature overlap. The interaction between context and level of overlap was significant [$F_1(1, 15) = 6.86$, $MSE = 213$, $P = 0.019$; $F_2(1, 17) = 9.98$, $MSE = 165$, $P = 0.006$]. The simple effect of context was significant for the segment overlap [$F_1(1, 15) = 10.82$, $MSE = 209$, $P = 0.005$; $F_2(1, 17) = 10.42$, $MSE = 245$, $P = 0.005$], but there was no effect for the feature overlap [$F_1(1, 15) < 1$, $MSE = 148$, $P = 0.60$; $F_2(1, 17) < 1$, $MSE = 182$, $P = 0.614$]. The simple effect of level of overlap was significant for the homogeneous condition [$F_1(1, 15) = 9.68$, $MSE = 136$, $P = 0.007$; $F_2(1, 17) = 7.51$, $MSE = 197$, $P = 0.014$], but there was no effect for the heterogeneous condition [$F_1(1, 15) < 1$, $MSE = 489$, $P = 0.43$; $F_2(1, 17) = 1.90$, $MSE = 189$, $P = 0.186$]. To conclude, a preparation effect was only obtained when the picture names shared an initial segment, but not when all but one of the features of the first segment were shared. The production latencies were fastest in the homogeneous condition in which the segments were shared and the latencies were slower in the other three conditions, which did not differ statistically.

EXPERIMENT 4

In Experiments 1 to 3, the critical feature concerned voicing. It is always possible that some other feature would have worked differently. Obviously, there is no need to look at every single feature in order to make the case, but one wonders whether a difference on the dimension of place of articulation may have another effect than a difference in voicing. As mentioned earlier, the greatest number of segment errors in slips of the

TABLE 4
Mean Production Latencies (M, in msec), Error Percentages (E%), and Preparation Effects per Context and Level of Overlap for Experiment 3

	Context					
	Homogeneous		*Heterogeneous*		*Preparation*	
Level of overlap	*M*	*E%*	*M*	*E%*	*M*	*E%*
Segments	549	0.9	565	1.0	– 16	– 0.1
Features	561	0.6	559	0.7	2	– 0.1

tongue involve a change in place of articulation (Shattuck-Hufnagel & Klatt, 1979), thus place of articulation appears to have the greatest mobility. Given this independence, it is possible that a difference in place of articulation of a segment does not block preparation while a difference in voicing does. Experiment 4 tested the same predictions as in Experiments 1 to 3, but now by looking at segments differing in a place of articulation feature, namely whether the segment is pronounced labial or coronal.

Method

The experiment was carried out with eight participants. The method was the same as in Experiment 1. The target words were all disyllabic. The critical segments in the experiment were the minimal pairs /m/ (labial) and /n/ (coronal), /v/ (labial) and /z/ (coronal), and /b/ (labial) and /d/ (coronal). As in Experiment 2, the vowels that followed these segments within the syllable were the same. The critical syllables were /mo/ and /no/, /ve/ and /ze/, /ba/ and /da/. In homogeneous sets, the initial syllables were fully shared (in segmental sets) or differed by one feature (in featural sets). All responses occurred in all conditions. The Appendix lists the materials.

Results

Table 5 gives the mean production latencies, error percentages, and preparation effects for Experiment 4. It shows that the production latencies were faster for the homogeneous condition in which the segments were shared than for the other three conditions, which did not differ much. A facilitatory effect from homogeneity was obtained for the sets with the segment overlap, but there was no effect at all for the sets with the feature overlap. The interaction between context and level of overlap was significant [$F_1(1,7) = 16.65$, $MSE = 852$, $P = 0.005$; $F_2(1,17) = 40.99$, $MSE = 778$, $P < 0.001$]. The simple effect of context was significant for the segment overlap [$F_1(1,7) = 12.08$, $MSE = 1746$, $P = 0.010$;

TABLE 5
Mean Production Latencies (M, in msec), Error Percentages (E%), and Preparation Effects per Context and Level of Overlap for Experiment 4

	Context					
	Homogeneous		*Heterogeneous*		*Preparation*	
Level of overlap	*M*	*E%*	*M*	*E%*	*M*	*E%*
Segments	643	3.5	716	3.1	– 73	0.4
Features	709	4.0	697	3.1	12	0.9

$F_2(1, 17) = 145.11$, $MSE = 327$, $P < 0.001$], but there was no effect for the feature overlap [$F_1(1, 7) = 1.12$, $MSE = 478$, $P = 0.325$; $F_2(1, 17) = 1.0$, $MSE = 1208$, $P = 0.332$]. The simple effect of level of overlap was significant for the homogeneous condition [$F_1(1, 7) = 4.42$, $MSE = 3896$, $P = 0.074$; $F_2(1, 17) = 42.55$, $MSE = 911$, $P < 0.001$], but there was no reliable effect for the heterogeneous condition [$F_1(1, 7) = 2.14$, $MSE = 645$, $P = 0.187$; $F_2(1, 17) = 4.00$, $MSE = 776$, $P = 0.062$].

The facilitatory effect of shared first segments in the present experiment is huge (73 msec). This is more than double what is seen in the previous experiments although mean production times and error rates are comparable. It is unclear why this is. Presumably the difference in effect size is due to differences in participants and responses. Most important for now, however, is the obtained *pattern* of results: The production latencies were fastest in the homogeneous condition in which the segments were shared and the latencies were slower in the other three conditions, which did not differ statistically.

To conclude, a preparation effect was only obtained when the response words shared the segments of a syllable, but not when all but one of the features of the first syllable were shared. The place of articulation feature tested here behaved the same as the feature of voicing tested in the previous experiments.

COMPUTER SIMULATIONS

Experiments 1 to 4 showed that a preparation effect is only obtained when words share initial segments, but not when all but one of the features of their initial segments are shared. Thus, the results support the claim that segments themselves are planning units in speech production rather than their features only, in agreement with the speech error findings. Below, I show by means of computer simulation that the WEAVER model accounts for the observations. The simulations involved word-form encoding up to the access of the syllable-sized articulatory programs. The mathematical equations for the spreading of activation and the expectation of the word-form encoding latency can be found in Roelofs (1994, 1996a, 1997a). The parameters for WEAVER that simulate the current phenomenon were identical with those in applications of the model to a wide range of other findings (Roelofs, 1994, 1996a, 1997a; Roelofs & Meyer, 1998).

The simulations concerned the preparation effect for sets with either segment or feature overlap (level of overlap). All conditions were tested within items. Disyllabic words were used throughout. The simulations tested homogeneous response sets including words like English TE*nnis* and TE*ller* (the segment-overlap condition) versus TE*nnis* and DE*vil* (the feature-overlap condition, where /t/ and /d/ share all features except voicing). The heterogeneous sets combined the responses of different homogeneous sets, such that there was no form overlap. Advance knowledge about the form of the response word was simulated by completing the phonological and phonetic encoding of the word form as far as possible before the beginning of a trial. To test whether the size and content of the network influenced the simulation outcomes, both small- and large-scale simulations were run. The critical items were tested alone or embedded into networks containing 50 other words, which were randomly selected from the CELEX lexical database (Baayen et al., 1995). The small and large simulations yielded the same outcomes.

Figure 1 presents the results of the simulations. Sharing *te-* in *tennis* and *teller* (the condition with segment overlap) yielded facilitation, whereas sharing features but no initial segment in *tennis* and *devil* (the condition with feature overlap) yielded no effect. Thus, crucially, to obtain facilitation from preparation in WEAVER, segments rather than features must be shared. If only features are shared, this does not even lead to a reduction of the preparation effect, but results in no facilitation at all. This corresponds to the phenomenon that has been empirically observed.

GENERAL DISCUSSION

This paper addressed the issue of whether there is an intermediate unit in speech production (the segment) between the morpheme and the feature. Using a form-preparation paradigm, I showed that there is facilitation if all

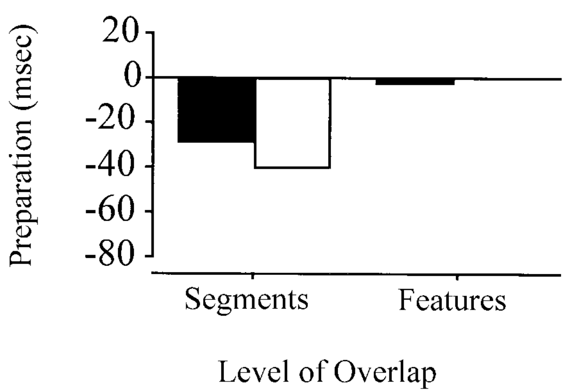


FIG. 1. Mean difference in msec between the production latencies in the homogeneous and the heterogeneous sets for Level of Overlap (segments *vs* features): Real data (■) and data simulated by WEAVER (□).

the words in the response set start with the same segment but no facilitation if the segments are minimally different from each other. On the basis of the special status of identity, I argued that segments exist as an integral part of speech planning. The present findings do not exclude that features play a role in planning a word. WEAVER explains why preparing a segment has an effect while preparing features has not in the form-preparation task. In the model, segments are encoded serially but the features of a string of segments may be accessed in parallel. Due to this parallelism, preparing features will not yield facilitation.

An issue that has come up time and again (Meyer, 1990, 1991; Roelofs, 1998) is whether the facilitation effects in the preparation paradigm are due to phonological encoding or to lower levels such as articulator preparation, that is, moving the articulators in the correct starting position before the beginning of a trial. The present results strongly suggest that the priming effects in the form-preparation paradigm cannot be assigned to articulator preparation. The critical segments in the homogeneous condition of the first three experiments always shared all their supralaryngeal features of manner and place of articulation and differed only in the laryngeal feature of voicing. In the fourth experiment, all features were shared except the place of articulation. Thus, if articulator preparation is the cause of the facilitation effects in the preparation paradigm, facilitation should have been obtained in both the segment and the feature condition. However, since a preparation effect is observed only when segments are shared but not when features are shared, the locus of the preparation effect must be the level of phonological encoding rather than a lower level such as articulator preparation.

In the introduction, I discussed how featural theories might account for segmental effects on speech errors without representing and manipulating segments explicitly. Some featural models (Dell et al., 1993) take a dynamical approach to segments. Following Dell et al. (1993), I argued that one outstanding difficulty for these featural models, as they stand, is posed by exchange errors. These errors require an account in which the features of non-adjacent segments are available simultaneously and in which features are independently movable. Both requirements pose difficulty to a dynamical approach such as that explored by Dell et al. (1993), where, roughly, the features of one segment are activated at a time and where features are strongly tied to each other. The fact that features are tied might perhaps explain the segmental effect observed in Experiments 1 to 4. However, in order to explain the exchange errors, one has to make the opposite assumption, namely that features are independently movable, thus loosely connected. As a general approach, the featural view may adopt this assumption of independence, which would allow it to explain exchange errors. However, by adopting this assumption, the segmental effect in the present experiments remains unexplained. If preparation effects were due to planning features that are independently movable, feature overlap should yield facilitation, as explained in the introduction. However, the present experiments show that full segment overlap is necessary. This suggests that memory contains codes that represent the phonological unity of the features that together form a segment. Features appear to be chunked into segments.

In the experiments, I employed a form preparation, or “implicit priming”, task. This task falls into the general class of choice-response tasks (Donders, 1868; Luce, 1986). Priming and precuing of choice responses has been widely used in studying the advance planning of skilled action (for review, see Kelso, 1982; Rosenbaum, 1991). For example, Rosenbaum (1980) used precuing to control the amount of preparation in arm movement. He manipulated the uncertainty (number of choices) in the specification of arm direction and extent, and observed that as more information was available to allow preparation, movement initiation time decreased. Implicit priming differs from precuing in that no cues are given in advance but the cue is implicit in the response set. However, the logic is the same in that both implicit priming and precuing allow for preparation of the action. The findings for arm movements and the present findings differ, however. For each increment of information about the arm movement, one finds a decrement in initiation time. This is *not* what the present experiments show: no amount of available information has an effect unless the first segments are exactly identical. This implies a special organisation different from movement in general. It strengthens the notion that models that assume the sameness of the learning of arm movements and the learning of the movements of lexical items (like most PDP connectionist models) are missing a level of organisation.

Obviously, the form-preparation task differs in a number of ways from ordinary speech production. Speakers rarely say the same three words over and over again, and they cannot normally predict how the next word to be uttered will begin. Yet, the effects in the experiments show very systematic patterns. Speakers can exploit certain types of information, whereas others are useless. A natural account of these limits is to relate them to the mechanism by which speech is normally planned as a form of skilled action (cf. Simon, 1969). Thus, I assume that the reason why preparation effects were only obtained for segment but not for feature overlap is that features are accessed from segments in normal production and that these segments are selected serially but their features in parallel.

Manuscript received April 1998
Revised manuscript received November 1998

REFERENCES

Anderson, J.R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Baayen, R.H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX Lexical Database*. (CD-ROM). Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.

Booij, G.E. (1995). *The phonology of Dutch*. Oxford, UK: Clarendon Press.

Butterworth, B. (1989). Lexical access in speech production. In W. Marslen-Wilson (Ed.), *Lexical representation and process*, pp. 108–135. Cambridge, MA: MIT Press.

Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology*, *14*, 177–208.

Dell, G.S. (1984). The representation of serial order in speech: Evidence from the repeated phoneme effect in speech errors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*, 222–233.

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

Dell, G.S. (1988). The retrieval of phonological forms in production: Tests of predictions from a connectionist model. *Journal of Memory and Language*, *27*, 124–142.

Dell, G.S., Juliano, C., & Govindjee, A. (1993). Structure and content in language production: A theory of frame constraints in phonological speech errors. *Cognitive Science*, *17*, 149–195.

Dell, G.S., & Juliano, C. (1996). Computational models of phonological encoding. In T. Dijkstra, & K. De Smedt (Eds), *Computational psycholinguistics: AI and connectionist models of human language processing*, pp. 328–359. London: Taylor & Francis.

Donders, F.C. (1868). Die Schnelligkeit psychischer Prozesse. *Archiv Anatomie und Physiologie*, 657–681.

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

Garrett, M.F. (1975). The analysis of sentence production. In G.H. Bower (Ed.), *The psychology of learning and motivation*, pp. 133–177. New York: Academic Press.

Garrett, M.F. (1988). Processes in language production. In F.J. Newmeyer (Ed.), *Linguistics: The Cambridge survey*, vol. 3, pp. 69–96. Cambridge, MA: Harvard University Press.

Goldsmith, J. (1990). *Autosegmental and metrical phonology*. Cambridge, MA: Basil Blackwell.

Jordan, M.I. (1986). *Serial order: A parallel distributed processing approach*. Technical Report No ICS-8604. La Jolla, CA: University of California at San Diego.

Jordan, M.I. (1990). Motor learning and the degrees of freedom problem. In M. Jeannerod (Ed.), *Attention and Performance, XIII*, pp. 796–836. Hillsdale, NJ: LEA.

Kelso, J.A.S. (Ed.) (1982). *Human motor behavior: An introduction*. Hillsdale, NJ: LEA.

Kempen, G., & Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. *Cognitive Science*, *11*, 201–258.

Kempen, G., & Huijbers, P. (1983). The lexicalization process in sentence production and naming: Indirect election of words. *Cognition*, *14*, 185–209.

Kenstowicz, M. (1994). *Phonology in generative grammar*. Oxford, UK: Blackwell.

Levelt, W.J.M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Levelt, W.J.M. (1992). Accessing words in speech production: Stages, processes and representations. *Cognition*, *42*, 1–22.

Levelt, W.J.M., Roelofs, A., & Meyer, A.S. (in press). A theory of lexical access in speech production. *Behavioral and Brain Sciences*.

Levelt, W.J.M., & Wheeldon, L. (1994). Do speakers have access to a mental syllabary? *Cognition*, *50*, 239–269.

Luce, R.D. (1986). *Response times: Their role in inferring elementary mental organization*. New York: Oxford University Press.
Meyer, A.S. (1990). The time course of phonological encoding in language production: The encoding of successive syllables of a word. *Journal of Memory and Language*, *29*, 524–545.
Meyer, A.S. (1991). The time course of phonological encoding in language production: The phonological encoding inside a syllable. *Journal of Memory and Language*, *30*, 69–89.
Meyer, A.S., & Schriefers, H. (1991). Phonological facilitation in picture-word interference experiments: Effects of stimulus onset asynchrony and types of interfering stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 1146–1160.
Miller, G.A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*, 81–97.
Mowrey, R.A., & MacKay, I.R.A. (1990). Phonological primitives: Electromyographic speech error evidence. *Journal of the Acoustical Society of America*, *88*, 1299–1312.
Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.
Nooteboom, S.G. (1969). The tongue slips into patterns. In A.G. Sciarone, A.J. van Essen, & A.A. van Raad (Eds), *Nomen: Leyden studies in linguistics and phonetics*, pp. 114–132. The Hague: Mouton.
Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, *42*, 107–142.
Roelofs, A. (1993). Testing a non-decompositional theory of lemma retrieval in speaking: Retrieval of verbs. *Cognition*, *47*, 59–87.
Roelofs, A. (1994). On-line versus off-line priming of word-form encoding in spoken word production. In A. Ram, & K. Eiselt (Eds), *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society*, pp. 772–777. Hillsdale, NJ: LEA.
Roelofs, A. (1996a). Serial order in planning the production of successive morphemes of a word. *Journal of Memory and Language*, *35*, 854–876.
Roelofs, A. (1996b). Morpheme frequency in speech production: Testing WEAVER. In G.E. Booij and J. van Marle (Eds), *Yearbook of Morphology 1996*, pp. 135–154. Dordrecht: Kluwer Academic Publishers.
Roelofs, A. (1997a). The WEAVER model of word-form encoding in speech production. *Cognition*, *64*, 249–284.
Roelofs, A. (1997b). Syllabification in speech production: Evaluation of WEAVER. *Language and Cognitive Processes*, *12*, 659–696.
Roelofs, A. (1998). Rightward incrementality in encoding simple phrasal forms in speech production: Verb-particle combinations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 904–921.
Roelofs, A., & Meyer, A.S. (1998). Metrical structure in planning the production of spoken words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 922–939.
Rosenbaum, D.A. (1980). Human motor initiation: Specification of arm direction and extent. *Journal of Experimental Psychology: General*, *169*, 444–474.
Rosenbaum, D.A. (1991). *Human motor control*. San Diego, CA: Academic Press.
Shattuck-Hufnagel, S. (1979). Speech errors as evidence for a serial-order mechanism in sentence production. In W.E. Cooper, & E.C.T. Walker (Eds), *Sentence processing: Psycholinguistic studies presented to Merrill Garrett*, pp. 295–342. Hillsdale, NJ: Lawrence Erlbaum.
Shattuck-Hufnagel, S. (1983). Sublexical units and suprasegmental structure in speech production planning. In P.F. MacNeilage (Ed.), *The production of speech*, pp. 109–136. New York: Springer.
Shattuck-Hufnagel, S., & Klatt, D.H. (1979). The limited use of distinctive features and markedness in speech production: Evidence from speech error data. *Journal of Verbal Learning and Verbal Behavior*, *18*, 41–55.

Simon, H.A. (1969). *The sciences of the artificial*. Cambridge, MA: MIT Press.
Smyth, M.M., & Wing, A.M. (1984). (Eds) *The psychology of human movement*. London: Academic Press.
Stemberger, J.P. (1982). The nature of segments in the lexicon: Evidence from speech errors. *Lingua*, *56*, 235–259.
Stemberger, J.P. (1985). An interactive activation model of language production. In A.W. Ellis (Ed.), *Progress in the psychology of language*, Vol. 1, pp. 143–186. London: LEA.
Stemberger, J.P. (1989). Speech errors in early child language production. *Journal of Memory and Language*, *28*, 164–188.
Stemberger, J.P. (1990). Wordshape errors in language production. *Cognition*, *35*, 123–157.
Wijnen, F. (1992). Incidental word and sound errors in young speakers. *Journal of Memory and Language*, *31*, 734–755.

APPENDIX

Response Sets of Experiments 2 to 4

Experiment 2

Level of overlap: Segments
Context: Homogeneous
Set 1: bami, bajes, balie (noodles, gaol, counter)
Set 2: degen, demo, deken (sword, demo, blanket)
Set 3: virus, visum, villa (virus, visa, villa)
Set 4: pater, paling, parel (father, eel, pearl)
Set 5: tegel, telex, teken (tile, telex, sign)
Set 6: ficus, file, fiche (ficus, file, counter)
Context: Heterogeneous
Set 7: degen, visum, bami
Set 8: virus, bajes, demo
Set 9: balie, deken, villa
Set 10: teken, ficus, pater
Set 11: fiche, parel, tegel
Set 12: paling, telex, file

Level of overlap: Features
Context: Homogeneous
Set 13: bajes, bami, paling
Set 14: demo, degen, teken
Set 15: villa, visum, fiche
Set 16: pater, parel, balie
Set 17: telex, tegel, deken
Set 18: file, ficus, virus
Context: Heterogeneous
Set 19: bami, demo, fiche
Set 20: villa, paling, degen
Set 21: teken, visum, bajes
Set 22: pater, tegel, virus
Set 23: ficus, balie, telex
Set 24: deken, file, parel

Note. English translations are given in parentheses.

Experiment 3

Level of overlap: Segments
Context: Homogeneous
Set 1: boek, bijl, beer (book, axe, bear)
Set 2: deur, das, dolk (door, tie, dagger)
Set 3: vis, vork, veer (fish, fork, feather)
Set 4: pet, pauw, pop (cap, peacock, doll)
Set 5: tas, tent, taart (bag, tent, tart)
Set 6: fiets, fluit, fee (bike, flute, fairy)
Context: Heterogeneous
Set 7: deur, vis, boek
Set 8: veer, bijl, das
Set 9: beer, dolk, vork
Set 10: taart, fee, pet
Set 11: fiets, pop, tas
Set 12: pauw, tent, fluit

Level of overlap: Features
Context: Homogeneous
Set 13: pauw, bijl, boek
Set 14: taart, das, deur
Set 15: fee, vork, vis
Set 16: beer, pet, pop
Set 17: dolk, tent, tas
Set 18: veer, fiets, fluit
Context: Heterogeneous
Set 19: boek, das, fee
Set 20: vork, pauw, deur
Set 21: taart, vis, bijl
Set 22: pop, tas, veer
Set 23: fiets, beer, tent
Set 24: dolk, fluit, pet

Experiment 4

Level of overlap: Segments
Context: Homogeneous
Set 1: vete, vezel, venus (feud, fibre, Venus)
Set 2: bami, bajes, balie (noodles, gaol, counter)
Set 3: moker, molen, mode (sledge, mill, fashion)
Set 4: zebra, zetel, zenuw (zebra, seat, nerve)
Set 5: dadel, dame, datum (date [fruit], lady, date)
Set 6: nozem, notie, noga (yob, notion, nougat)
Context: Heterogeneous
Set 7: bami, molen, vete
Set 8: moker, vezel, bajes
Set 9: venus, balie, mode
Set 10: datum, nozem, zebra
Set 11: noga, zenuw, dadel
Set 12: zetel, dame, notie

Level of overlap: Features
Context: Homogeneous
Set 13: zetel, venus, vezel
Set 14: datum, bajes, bami
Set 15: noga, mode, molen
Set 16: vete, zebra, zenuw
Set 17: balie, dame, dadel
Set 18: moker, notie, nozem
Context: Heterogeneous
Set 19: venus, bajes, mode
Set 20: noga, zetel, bami
Set 21: datum, molen, vezel
Set 22: zebra, dadel, moker
Set 23: notie, vete, dame
Set 24: balie, nozem, zenuw