Spoken language planning and the initiation of articulation

Ardi Roelofs

Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

Minimalist theories of spoken language planning hold that articulation starts when the first speech segment has been planned, whereas non-minimalist theories assume larger units (e.g., Levelt, Roelofs, & Meyer, 1999a). Three experiments are reported, which were designed to distinguish between these views using a new hybrid task that factorially manipulated preparation and auditory priming of spoken language production. Minimalist theories predict no effect from priming of non-initial segments when the initial segment of an utterance is already prepared; observing such a priming effect would support non-minimalist theories. In all three experiments, preparation and priming yielded main effects, and together their effects were additive. Preparation of initial segments does not eliminate priming effects for later segments. These results challenge the minimalist view. The findings are simulated by WEAVER++ (Roelofs, 1997b), which employs the phonological word as the lower limit for articulation initiation.

An important control issue in skilled behaviour such as speaking words concerns the relationship between the generation of the action plan and its execution (e.g., Lashley, 1951; Levelt, 1989). How much of the speech plan for an utterance has to be completed before articulation is initiated?[1] Minimalist theories of articulation initiation hold that articulation starts as soon as the first speech segment has been planned (e.g., Dell, Juliano, & Govindjee, 1993; Jordan, 1990; Kawamoto, 1999; MacKay, 1987; Santiago, MacKay, Palma, & Rho, 2000). For example, they hold that the production of the word *melon* is initiated when the segment /m/ is planned.

---

[1] The issue of planning and articulation initiation has also been investigated within the context of the translation of spelling to sound in oral reading (e.g., Kawamoto, Kello, Jones, & Bame, 1998), which will be taken up in the General Discussion section.

Jordan (1990) and Dell and colleagues (Dell et al., 1993) proposed parallel distributed processing (PDP) models in which there is such a minimal lag between speech planning and articulation.[2] In the models, word nodes (e.g., *melon*) are connected via a layer of hidden nodes to nodes for articulatory features (e.g., [+nasal] and [+labial] for /m/). Activation of a word node is passed through weighted connections to the hidden layer and finally to the output layer, which contains one node for each articulatory feature in the language. In producing a word, the network activates over time the features corresponding to each speech segment of the word, roughly one segment at a time. By design, articulation is initiated when the features of the first segment have become activated.

Similarly, MacKay's (1987; Santiago et al., 2000) node structure theory (NST) assumes that planning the first segment of a word determines the initiation of articulation. According to this theory, words are planned via a sequential tree traversal process that operates in a top-down and left-to-right fashion. In particular, the initial syllable of a word is planned before its non-initial syllables, and syllable onsets are planned before syllable rimes. Within syllable onsets, initial speech segments are planned before non-initial segments. At each node in the tree, a binary decision is taken (e.g., to plan *me* before *lon* for *melon*). Articulation is initiated when the left-most, bottom-most node (i.e., the first segment) is reached.

An important characteristic of these theories is the absence of articulatory buffering. In the PDP models of Jordan (1990) and Dell and colleagues (Dell et al., 1993), articulation is necessarily initiated when the features of the first segment have become activated. Moreover, whereas articulatory buffering of initial segments would be compatible with NST, Santiago et al. (2000) explicitly argued against it (p. 13). According to them "under NST, vocal production follows the heels of activating a terminal node" (p. 7).

By contrast, non-minimalist theories claim that larger units are planned and buffered before articulation is initiated. Schriefers and Teruel (1999) claim that speakers have the option to initiate articulation when the first syllable has been planned, but articulation may also be initiated later. Levelt and colleagues (Levelt, 1989; Levelt, Roelofs, & Meyer, 1999a) claim that the minimal unit for articulation initiation is the phonological word, with larger units as an option. This assumption has been implemented in the WEAVER++ model of spoken word production (e.g., Roelofs, 1997b). In WEAVER++, the speech segments of a word are activated in parallel, and they are syllabified in a sequential left-to-right manner. After syllabification, phonetic syllable programmes are accessed from memory and stored in an output buffer as a linear string. Only when the buffer contains the phonetic programmes for one phonological word may articulation begin. A phonological word is defined as the minimal prosodic unit above the metrical foot, to which clitics, such as unstressed words, can attach (e.g., Levelt, 1989; Wheeldon & Lahiri, 1997). For example, *melon, racket,* and also "she's" (consisting of *she* and the reduced form of *is*) are phonological words.

Direct empirical support for the role of the phonological word as a planning unit in speech production comes from a study by Wheeldon and Lahiri (1997), who used a modified version of the full-preparation paradigm developed by Sternberg and colleagues (see Sternberg,

---

[2]Although the PDP models of Dell, Juliano, and Govindjee (1993) and Jordan (1990) have not been explicitly applied to data from real time processing tasks, they address performance data (i.e., speech errors and co-articulation findings, respectively) and are therefore relevant in the present context. The models are mentioned to exemplify a class of model in which articulation is by design initiated when the features of the first segment are activated.

Knoll, Monsell, & Wright, 1988, for a review). In three experiments, Wheeldon and Lahiri observed that the production-onset latency for fully prepared sentences is a function of the number of phonological words that a sentence comprises when syntactic structure, number of lexical items, and number of syllables are held constant. However, although these experiments demonstrate that the number of phonological words determines the production latency when the utterance is fully prepared, they do not provide evidence on the issue of how much of the speech plan needs to be prepared before articulation is initiated for incompletely planned utterances. A fourth experiment by Wheeldon and Lahiri required on-line sentence production rather than preparation, and compared sentences with clitic and non-clitic phonological words. The results showed that production latency increases with the complexity of the first phonological word in the utterance (i.e., the latencies were larger for the clitic than for the non-clitic utterances). However, as Wheeldon and Lahiri discuss, as their clitic and non-clitic constructions differed in syntactic requirements, it is unclear exactly what aspect of the utterance types caused the increase in production latency. Specific evidence that more than the first segment or syllable needs to be planned before articulation is initiated comes from on-line *priming* and *preparation* experiments.

Meyer and Schriefers (1991) showed that auditory priming of both initial and non-initial aspects of a word is possible during the planning of its production in picture naming. When begin-related or end-related spoken primes are presented over headphones during the production of a monosyllabic word (e.g., the spoken primes /kæp/ or /hæt/ for the target word *cat*, /kæt/), both primes yield facilitation compared to unrelated primes, which yield interference relative to a control condition with no prime[3]. Moreover, when first-syllable (/metəl/) or second-syllable (/vIlən/) spoken primes are presented during the production of a disyllabic word (e.g., *melon*, /melən/), both primes yield facilitation too. The magnitude of the priming effect from initial and non-initial primes was the same. When the time between prime and target presentation is manipulated, called the stimulus-onset asynchrony (SOA), the SOA at which the facilitation is first detected differs between the two types of prime. In particular, the onset of facilitation is at a later SOA for non-initial primes than for initial ones. The priming effect for non-initial aspects of the target word and the SOA effect suggest that words are planned sequentially and that articulation is initiated after the complete word has been planned (Roelofs, 1997b). If articulation were initiated right after planning the first segment or syllable (e.g., after /m/ or /me/ for *melon*), priming of the second syllable should have no effect on the production onset latency, and no SOA effect should be obtained, contrary to the empirical findings.

Further evidence for the assumption of rightward planning and the phonological word as minimal initiation unit comes from partial-preparation experiments. Meyer (1990, 1991) asked speakers to produce words out of small sets in response to prompts. The words in the sets were different in form (the heterogeneous sets), or they shared one or more speech segments (the homogeneous sets). Meyer found shorter production latencies in homogeneous than in heterogeneous sets. However, this difference was only obtained when the response

---

[3]In the psycholinguistic literature on spoken word production, faster latencies of phonologically related primes relative to unrelated primes are often called "facilitation", even when the latencies with related primes are not faster than without a prime. To stay in line with the terminology in the literature, and my own earlier papers, I use the term facilitation here too.

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 *melon*, *metal*, and *merit*, but not for the end-related homogeneous set that included *pocket*, *ticket*, and *racket*. The magnitude of the preparation effect increased with the number of shared word-initial segments. For example, sharing the first syllable plus the first segment of the second syllable of a disyllabic word yields a larger facilitation effect than sharing the first syllable only. Also, sharing the first two syllables of a trisyllabic word yields a larger effect than sharing the first syllable only. Roelofs (1998) showed that the preparation effect increases even when the shared part crosses a phonological word boundary in small phrasal constructions. These findings suggest that words and small phrases are sequentially planned and that articulation initiation minimally requires that the first phonological word in the utterance has been planned. If articulation were initiated upon completion of the plan for the first segment or syllable, additionally sharing of a later part should not increase the preparation effect, contrary to the experimental findings.

Recently, however, researchers have provided evidence against the claim that the phonological word is the minimal unit for articulation. Left-to-right encoding with the phonological word as minimal unit for articulation initiation predicts that the production onset latencies should be larger for long (e.g., disyllabic) than for short (e.g., monosyllabic) words. However, Bachoud-Lévi, Dupoux, Cohen, and Mehler (1998) failed to find different production latencies for monosyllabic and disyllabic target words in picture-naming experiments in French and English. They concluded that either word forms are not generated sequentially or speakers start articulating before completing the planning of a phonological word. Either way, the theory of Levelt et al. (1999a) would have to be changed. The WEAVER++ model, which implements this theory, predicts a latency difference of 26 ms for the monosyllabic and disyllabic words in the experiments of Meyer (1990, 1991) and Meyer and Schriefers (1991), contrary to the findings of Bachoud-Lévi et al.

However, in a replication of the experiments of Bachoud-Lévi et al. (1998) in Dutch, Meyer, Roelofs, and Levelt (1999) did obtain the length effect (of about 30 ms), but only when the materials were blocked for length. When monosyllabic and disyllabic names were mixed, as in the study of Bachoud-Lévi et al., the latencies did not differ significantly. Therefore, Meyer et al. assume, following Lupker, Brown, and Colombo (1997), that speakers can strategically control the criterion for the initiation of articulation. In the context of the single- versus dual-route debate on oral reading, Lupker et al. found evidence suggesting that a response criterion is optimal when word types are blocked, but suboptimal when word types are mixed.[4] Van der Plaats and Van Galen (1990) observed a similar effect for handwriting between conditions that mixed or blocked word length, and made similar assumptions about a response criterion to account for their observations. Meyer et al. assume that mixing words of different length conceals length effects, which explains why Bachoud-Lévi et al. did not obtain the effect.

Schriefers and Teruel (1999) also argue for a strategic control of the response criterion by speakers. They replicated the first- and second-syllable priming effects for single-word

---

[4]Clearly, in picture-naming and preparation tasks, factors such as spelling regularity of the targets in translating from spelling to sound play no role, but production latencies may be affected by a response criterion based on other factors, such as word length, as suggested by the findings of Meyer, Roelofs, and Levelt (1999).

production that were observed by Meyer and Schriefers (1991). However, in examining the effect of spoken primes on producing German sentences like "rosa Tisch" ("pink table"), they observed priming of the first syllable of the first word (*ro* of *rosa*) but no priming for the second word (*Tisch*). For the second syllable of the first word (*sa* of *rosa*), Schriefers and Teruel obtained a weak facilitation effect. It appeared that, based on the number of restarts and hesitations, the participants could post hoc be divided into two groups, which differed in their tendency to initiate articulation quickly. The "careful" group showed facilitation for both the first and second syllable of the first word, whereas the "hasty" group (making most restarts and hesitations) showed facilitation for the first syllable only. Schriefers and Teruel concluded that this argues against the phonological word as the fixed, lower limit of planning. Instead, according to them, the findings suggest that speakers can adjust the size of the planning unit, and that planning of the first syllable may suffice. A problem with this conclusion is that it is based on only two relevant target words in the experiment, *lila* (*purple*) and *rosa* (*pink*), and it remains to be seen whether the findings can be replicated with other words. More importantly, contrary to what Schriefers and Teruel believe, their data appear to be compatible with the assumption of the phonological word as the lower boundary for articulation. Computer simulations showed that WEAVER++ yields no priming effect for second syllables when articulation is initiated upon recovery of the first-syllable motor program (simulating the hasty group) *after* having fully planned the phonological word representation. With "normal" responding (simulating the careful group), facilitation for both the first and second syllables is obtained. Thus, the difference in priming effect between the careful and hasty groups can be accounted for without assuming that articulation is initiated before completing the full phonological word representation.

This does not mean that the findings from priming (Meyer & Schriefers, 1991; Schriefers & Teruel, 1999) and preparation (Meyer, 1990, 1991; Roelofs, 1998; Wheeldon & Lahiri, 1997) should be taken as conclusive evidence against the minimalist view. Several network theories assume backward links from segments to words in the speech production lexicon (e.g., MacKay, 1987; but *not* Dell et al., 1993; Jordan, 1990). Under this assumption, the auditory second-syllable prime *lon* may facilitate the production of *melon* even when articulation is started before the planning of the second syllable. The auditory prime *lon* activates /l/, /ə/, and /n/ in the network, which may spread activation backward to the node for *melon*, which in its turn may activate forward /m/ and /e/. In this way, the planning of /m/ and /e/ may be facilitated compared to unrelated primes, yielding a priming effect, as observed by Meyer and Schriefers and Schriefers and Teruel. Furthermore, the increase of the preparation effect with the size of form overlap in a response set, observed by Meyer (1990, 1991) and Roelofs (1998), may reflect a special strategy of the participants. It may be the case that participants strategically use the shared part (e.g., *me* in *melon*, *metal*, and *merit*) as unit for the initiation of articulation, and that they stick to the same unit size (i.e., the first syllable) in the heterogeneous condition. That is, on each trial in the homogeneous condition, participants might immediately start articulating the shared part at prompt onset and plan the remainder of a word during articulation of the shared part. For example, in producing *melon* from the set *melon*, *metal*, and *merit*, they may plan *lon* during the articulation of *me*. In heterogeneous sets, this is not possible, of course. Thus, if the shared part is made larger, and the unit size of the shared part of the words determines articulation initiation, then more needs to be planned in heterogeneous sets, and an effect of the size of the shared part will be obtained, as empirically observed by Meyer

(1990, 1991) and Roelofs (1998). In summary, if one assumes (1) backward links in the production lexicon, and (2) a strategic use of the shared part in response sets as unit size for the initiation of articulation, the findings on priming and preparation are compatible with the minimalist view.

The present experiments tested between the minimalist and the non-minimalist views on the relation between planning and the initiation of articulation. To avoid differential effects of a response criterion, all responses in an experiment were of the same length in terms of number of syllables. To overcome the difficulties with the interpretation of the results from the earlier studies, the experiments tested for the combined effect of preparing early and priming later parts of an utterance. Participants produced utterances out of small sets in response to prompts. The utterances differed in form or shared the first syllable (e.g., the syllable *me* in *melon*, *metal*, and *merit*), which allowed for preparation of the syllable including the critical first segment. At prompt onset, before or during production planning, auditory syllable primes were presented that matched the second syllable of the targets or not (e.g., *lon* or *tal* for *melon*). Although preparation and priming aim at different serial loci (i.e., in the example, the first and second syllable of the target), their combined effect is predicted to be additive or interactive depending on the theoretical position.

The minimalist view holds that articulation is inevitably initiated right after having planned the first segment. Thus, on the most straightforward account (Dell et al., 1993; Jordan, 1990), preparation effects from shared initial syllables should be obtained, but auditory priming of a later syllable should not have an effect. Under the feedback hypothesis, however, priming of later syllables may have an effect, because later segments may activate earlier segments via the backward links. But note that effects of priming a later syllable may not be obtained in the homogeneous condition (with initial overlap, allowing for preparation). If speakers plan the shared segments (e.g., *me*) before the beginning of a trial (e.g., Santiago et al., 2000, p. 13) and initiate articulation on the basis of this planned syllable, priming of a later syllable (i.e., *lon*) should have no effect even if feedback exists. Activation of /m/ and /e/ by the prime *lon* via the feedback loop should have no effect because the initial segments have already been planned (this is assumed to explain the preparation effect). There will then only be an effect of priming in the heterogeneous condition (without overlap between the responses), where initial segments still need to be planned and may be affected by feedback. The same prediction holds under the additional assumption of the special initiation strategy (Assumption 2 earlier). With such a strategy and feedback, preparation effects should increase with the size of overlap, and priming effects may occur but, again, not in the homogeneous condition. Thus, if an effect of preparation is observed but no effect of priming, or an effect of priming only in the heterogeneous condition, then the minimalist view is supported. However, if a preparation effect is obtained, and a priming effect is observed in both the homogeneous and heterogeneous conditions, then the non-minimalist view is supported.

EXPERIMENT 1

Experiment 1 examined the effect of response-set homogeneity and auditory prime relatedness on the production of single words such as *melon*. It tested whether the combined use of preparation and priming replicates the key findings obtained with the priming and preparation of single-word production discussed earlier. Such a replication of results with a factorial

manipulation of priming and preparation would challenge a minimalist view on the relation between planning and articulation. In the experiment, the words had to be produced in homogeneous sets where they shared the first syllable and in heterogeneous sets where there was no overlap. As primes the second syllables of the words were used. A prime corresponded to either a target syllable (the related condition), or a syllable of another word in the response set (the unrelated condition), or no spoken prime was given (the control condition, consisting of a stretch of pink noise). Note that the control condition coincides with the situation in the experiments of Meyer (1990, 1991) and Roelofs (1996, 1997b, 1998, 1999) in that there also no spoken primes were given.

Method

Participants

The participants in all experiments were native speakers of Dutch, who were randomly selected from the pool of participants of the Max Planck Institute. None of the participants took part in more than one experiment. They were paid for their participation. Each experiment was conducted with a different group of 18 participants.

Materials and design

The materials for all three experiments were obtained from the Dutch part of the CELEX lexical data base (Baayen, Piepenbrock, & Gulikers, 1995). The materials of Experiment 1 consisted of two practice sets and six experimental sets of three word pairs each. The response sets are listed in Table 1. All prompts and responses were nouns (except for one response, the adjective *lening* [*supple*]), because suitable items were easiest to find in this word class. All responses were disyllabic. Each set was tested in a separate block of trials. In three experimental sets (the homogeneous sets), the response words shared the first syllable, and in the remaining three sets (the heterogeneous sets) they were unrelated in form. Thus, in the homogeneous condition, each response word was tested together with other response words with the same first syllable, whereas in the heterogeneous condition, the response words tested together in a block did not share a syllable. The first independent variable—homogeneous versus heterogeneous sets—is called *preparation*. The same prompt-response pairs were tested in the homogeneous and heterogeneous condition; only their combinations into sets differed.

Each response was combined with three auditory primes. A prime was the second syllable of the target word (related) or the second syllable of another response word in the set (unrelated), or there was no spoken syllable but pink noise (control). The pairing of responses and primes is also listed in Table 1. The second independent variable—related versus unrelated versus control—is called *priming*.

Each participant was tested on all sets. The order of the sets was rotated across participants in the following way. One group of 9 participants was first tested on the three homogeneous sets and then on the three heterogeneous sets. For the group of the remaining 9 participants, the order of testing homogeneous and heterogeneous sets was reversed. A different order of the three sets was used for each participant within a group, such that each set was tested once as the first, second, and third set. The homogeneous and heterogeneous sets were tested twice, that is, in two consecutive repetitions of the experimental session.

In a block of trials, each of the three pairs was tested four times in each of the three priming conditions (related, unrelated, control). There were 36 trials within a block. In all experiments, the order of testing the word pairs and primes was random, except that immediate repetitions of pairs and primes were excluded. A different random order was used for each block and each participant.

TABLE 1
Response sets and spoken primes of Experiment 1

Preparation	Set	Target	Priming	
			Unrelated	Related
Homogeneous	1	hamer (hammer)	gel	mer
		hagel (hail)	vik	gel
		havik (hawk)	mer	vik
	2	lenig (supple)	ger	nig
		leger (army)	pel	ger
		lepel (spoon)	nig	pel
	3	polis (policy)	ging	lis
		poging (effort)	ker	ging
		poker (poker)	lis	ker
Heterogeneous	4	leger	vik	ger
		havik	lis	vik
		polis	ger	lis
	5	lenig	gel	nig
		hagel	ker	gel
		poker	nig	ker
	6	lepel	ging	pel
		poging	mer	ging
		hamer	pel	mer

Note: English translations are given in parentheses. The author can provide a listing of the materials including the prompt words. The control condition is omitted from the table.

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 (a homogeneous and a heterogeneous block with the same structure as an experimental block, but with different items) were administered followed by the 12 experimental blocks. In the learning phase before each block, the three 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 ms. Next, the screen was cleared for 500 ms, followed by the display of the written prompt for 1500 ms. Upon presentation of the written prompt, the participant produced the corresponding response. For example, when the prompt *zaag* (*saw*) appeared on the computer screen, the participant said *hamer* (*hammer*) as fast as possible while trying to make no mistake. Simultaneously with prompt onset, an auditory prime was presented over closed headphones. The onset of presentation of the prompt and the prime was the same (i.e., SOA = 0 ms). 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 ms. Thus, the total duration of a trial was 3 s. A Hermac computer controlled the experiment.

The prime syllables were spoken by a female speaker and recorded using a Sony DCT55 DAT recorder. They were digitized with a sampling frequency of 20 kHz and stored on the hard disk of a speech server. For the control condition, a stretch of pink noise was generated having the same duration as the mean duration of the prime syllables. The primes were presented to the participants using Sennheiser MD281N closed headphones.

Analyses

For all experiments, the error coding and statistical analyses were the same. 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 participant's speech and tones indicating the onset of the prompt (1 kHz) and the moment of the triggering of the voice key (2.5 kHz). The experimenter also heard these tones on each trial (via closed headphones). The experimenter did not hear the auditory primes. The recordings were consulted after the experiment when the experimenter was in doubt about whether a response was fully correct. Four types of incorrect response were distinguished: (1) A participant produced a wrong response; (2) the response exhibited a disfluency—that is, the participant stuttered, paused within the utterance, or repaired the utterance; (3) the voice key was triggered by a non-speech sound (noise in the environment or a smacking sound produced by the lips or tongue); (4) the participant failed to respond within the time-out period of 1500 ms. Incorrect responses were excluded from the statistical analysis of the production latencies.

In all experiments, the production latencies and error rates were submitted to by-participant and by-item analyses of variance with the crossed variables preparation (homogeneous, heterogeneous) and priming (related, unrelated, control) as repeated measures factors. Both factors were tested within participants and within items. For the errors, no main effect or interaction was significant either by participants or by items in any experiment ($p > .05$). Therefore, the means are reported but not the test statistics for the error rates. Items were repeated within each block of trials and in two consecutive experimental sessions. Pre-analyses showed that repetition did not interact with the experimental factors in any of the experiments (all $ps > .05$). Therefore, the results for repetition are not reported.

Results and discussion

Table 2 gives the mean production latencies, standard deviations, and error percentages as a function of preparation and priming in Experiment 1. The table shows that the words were produced faster in the homogeneous condition than in the heterogeneous condition. Also, the words were produced faster with related primes and without primes than with unrelated primes. Thus, compared to the control condition the effect of priming was one of interference by the unrelated rather than facilitation by the related primes. The effect of prime relatedness was almost the same with (i.e., $617 – 590 = 27$ ms) and without (i.e., $648 – 623 = 25$ ms) preparation. Also, the effect of preparation was almost the same with related (i.e., $623 – 590 = 33$ ms) and unrelated (i.e., $648 – 617 = 31$ ms) primes.

TABLE 2
Mean production latencies[a] (M), standard deviations (SD), and error percentages (E%) as a function of preparation and priming in Experiment 1

	Preparation								
	Homogeneous			Heterogeneous			Total		
Priming	M	SD	E%	M	SD	E%	M	SD	E%
Related	590	87	1.6	623	84	1.9	607	87	1.8
Unrelated	617	108	3.2	648	92	1.4	633	101	2.3
Control	583	74	1.6	611	74	0.9	597	75	1.3
Total	597	92	2.1	627	85	1.4	612	90	1.8

[a]In ms.

The statistical analyses yielded main effects for preparation, $F_1(1, 17) = 20.62$, $p < .001$; $F_2(1, 8) = 14.65$, $p < .005$, and priming, $F_1(2, 34) = 19.48$, $p < .001$; $F_2(2, 16) = 28.14$, $p < .001$. Preparation and priming did not interact at all, $F_1(2, 34) < 1$, $p > .87$; $F_2(2, 16) < 1$, $p > .91$. Importantly, there was a clear priming effect in the homogeneous condition, $F_1(2, 34) = 8.58$, $p < .001$; $F_2(2, 16) = 10.83$, $p < .001$.

In sum, the experiment replicated the preparation effect from shared first syllables obtained by Meyer (1990, 1991) and Roelofs (1996, 1997a, 1998, 1999) and the priming effect of spoken second syllables obtained by Meyer and Schriefers (1991) and Schriefers and Teruel (1999). Most important, the experiment showed that in the context of preparation effects, priming effects are still obtained, and that priming effects are not restricted to the heterogeneous condition (there is a clear priming effect of 27 ms in the homogeneous condition). This argues against the minimalist view.

Of course, accepting that there is no difference in priming effect between the homogeneous and heterogeneous conditions amounts to accepting a null hypothesis. However, the aim of the present study was to see whether one could replicate the priming effect for the homogeneous condition. Having obtained the effect shows that preparation of initial segments does not eliminate priming effects for later segments and involves rejecting the null hypothesis. The fact that the size of the priming effect is constant across the homogeneous and heterogeneous conditions further suggests that preparation of initial segments is entirely irrelevant for the priming effect. Of course, there may be a very small effect that has gone undetected. Empirically, the difference was 2 ms, but it is in the wrong direction for the minimalist models (i.e., the priming effect is larger for the homogeneous than for the heterogeneous condition).

EXPERIMENT 2

Experiment 2 tested the same predictions as those in the Experiment 1, but now with utterances consisting of two phonological words. It examined the effect of response-set homogeneity and prime relatedness on the production of disyllabic responses consisting of monosyllabic verbs and particles. For example, participants had to produce simple imperatives such as "zoek op" [zuk ɔp] ("look up"). In the experiment, the verb could be prepared or not and the particle could be primed or not. The responses consisted of all nine possible combinations of three verbs and three particles. This means that feedback from a particle to a verb (if it exists) would not particularly favour any of the three verbs, so priming via feedback is excluded. Moreover, as the particle and verb are separate syntactic entities and different morphemes, feedback from the particle to a verb (e.g., from *op* to *zoek*) would in any case be much reduced.

The particle and verb of a verb–particle construction each make up an independent phonological word, and together they constitute a small phonological phrase. Phonological words are independent domains of syllabification, and most phonological rules apply within phonological words (Booij, 1995). Thus, the phonological word is an ideal candidate for the initiation of articulation. If the first segment is the obligatory unit of articulation initiation, priming the second phonological word should yield no facilitation. If facilitation is nevertheless obtained, this would challenge the minimalist view. Furthermore, if the first syllable suffices for initiating articulation (Schriefers & Teruel, 1999), and a syllable makes up a phonological word, this should much increase the chance that participants initiate articulation upon completion of the speech plan for the first syllable. If effects of priming in the context of preparation are still

obtained, this would suggest that more than the first syllable is needed before the initiation of articulation. Such a result for these utterances would suggest that speakers may even plan beyond the first phonological word at the cost of delaying the initiation of articulation.

Evidence for planning beyond the first phonological word would fully agree with the phonological word as minimal unit in WEAVER++. In the model, a phonological word is a minimal but not an obligatory unit. Articulation is not initiated earlier than completion of the phonological word, but it may be initiated later in order to respect the metrical cohesion of the items in phonological phrases, such as in verb–particle constructions (Roelofs, 1998).

Method

The method was the same as that in Experiment 1. The responses consisted of the nine possible combinations of three verbs and three particles. The verbs were *draaien*, *zoeken*, and *geven* (*turn*, *search*, *give*). The particles were *op*, *af*, and *uit* (*on*, *up*, and *off*). In homogeneous sets, the responses shared the verb, and in heterogeneous sets they did not. The spoken primes corresponded with the particle of the target (the related condition) or the particle of another response in the set (the unrelated condition), or no particle was presented (the control condition). Table 3 lists the materials of the experiment.

Results and discussion

Table 4 gives the mean production latencies, standard deviations, and error percentages as a function of preparation and priming in Experiment 2. The table shows that the utterances were produced faster in the homogeneous condition than in the heterogeneous condition.

TABLE 3
Response sets and spoken primes of Experiment 2

| | | | Priming | |
| | | | Unrelated | Related |
Preparation	*Set*	*Target*	*Unrelated*	*Related*
Homogeneous	1	zoek op (look up)	af	op
		zoek af (search)	uit	af
		zoek uit (sift out)	op	uit
	2	draai op (wind up)	af	op
		draai af (show)	uit	af
		draai uit (turn out)	op	uit
	3	geef op (give up)	af	op
		geef af (deliver)	uit	af
		geef uit (spend)	op	uit
Heterogeneous	4	zoek op	af	op
		draai af	uit	af
		geef uit	op	uit
	5	geef op	af	op
		zoek af	uit	af
		draai uit	op	uit
	6	draai op	af	op
		geef af	uit	af
		zoek uit	op	uit

Note: English translations are given in parentheses. The control condition is omitted from the table.

TABLE 4
Mean production latencies[a] (M), standard deviations (SD), and error percentages (E%)
as a function of preparation and priming in Experiment 2

| | Preparation | | | | | | | | |
| | Homogeneous | | | Heterogeneous | | | Total | | |
Priming	M	SD	E%	M	SD	E%	M	SD	E%
Related	593	76	3.2	684	93	2.5	638	96	2.9
Unrelated	634	90	3.8	730	103	4.8	682	108	4.3
Control	587	71	3.2	690	94	3.6	639	98	3.4
Total	605	82	3.4	701	99	3.6	653	103	3.5

[a]In ms.

Also, the utterances were produced faster with related primes and without primes than with unrelated primes. Thus, again, compared to the control condition the effect of priming was one of interference by the unrelated rather than facilitation by the related primes. The effect of prime relatedness was almost the same with preparation (i.e., 634 – 593 = 41 ms) and without (i.e., 730 – 684 = 46 ms). Also, the effect of preparation was almost the same with related primes (i.e., 684 – 593 = 91 ms) and with unrelated primes (i.e., 730 – 634 = 96 ms).

The statistical analyses yielded main effects for preparation, $F_1(1, 17) = 42.14, p < .001$; $F_2(1, 8) = 170.24, p < .001$, and priming, $F_1(2, 34) = 69.91, p < .001$; $F_2(2, 16) = 59.53, p < .001$. Priming and preparation did not interact, $F_1(2, 34) = 1.32, p > .28$; $F_2(2, 16) < 1, p > .41$. Importantly, there was a clear priming effect in the homogeneous condition, $F_1(2, 34) = 33.25, p < .001$; $F_2(2, 16) = 39.05, p < .001$.

In sum, the experiment allowed for preparation of a monosyllabic phonological word and priming of a second phonological word. The results show, again, that preparation and priming effects do occur, and that the priming effect is not restricted to the heterogeneous condition (there is a clear priming effect of 41 ms in the homogeneous condition). This argues against the minimalist view on articulation initiation.

EXPERIMENT 3

Experiment 3 tested the same predictions as those in Experiments 1 and 2, again with different participants and materials. In the previous experiments, one syllable could be prepared, and the other was primed. Experiment 3 increased the size of the overlap in homogeneous sets in order to increase the time for the planning of the remainder during articulation. The experiment changed the format of the target utterances to constructions that include cliticizations. Clitics are reduced forms of words such as pronouns, determiners, particles, auxiliary verbs, prepositions, and conjunctions that are phonologically dependent on a host (e.g., Booij, 1995; Levelt, 1989). For example, the reduced form 's [əs] of the Dutch adverb eens (now) cannot stand alone phonologically, but has to be integrated with a host.

The experiment examined the effect of response-set homogeneity and prime relatedness on the production of trisyllabic responses consisting of a verb-particle construction including the clitic 's [əs], creating utterances such as "zoek 's op" [zu.kəs ɔp]. Thus, the first phonological word now consists of two syllables rather than one, as it did in Experiment 2. Given that the

first phonological word is now longer, there is more time to plan the remaining third syllable during the articulation of the first two syllables. If we still obtain effects of priming in the context of preparation, this would challenge the minimalist view and strongly support the non-minimalist view. Furthermore, such a result would suggest that, even when the first phonological word is disyllabic, a speaker may plan beyond the first phonological word at the cost of delaying the initiation of articulation. Note that by comparing the effects for cliticized utterances only, the experiment remedies the shortcoming of the fourth experiment by Wheeldon and Lahiri (1997), who compared clitic and non-clitic constructions.

Method

The materials of the experiment were the same as those of Experiment 2, which were listed in Table 3, except that the utterances now contained the reduced form of the adverb *eens* (*now*), *'s* [əs], which must be adjoined phonologically to the preceding word. Thus, the participants had to produce, for example, "zoek 's op" [zu.kəs ɔp] instead of "zoek op" [zuk ɔp]. The method was the same as that of the previous experiments.

Results and discussion

Table 5 gives the mean production latencies, standard deviations, and error percentages as a function of preparation and priming in Experiment 3. The table shows that the utterances were produced faster in the homogeneous condition than in the heterogeneous condition. Also, the utterances were produced faster with related primes and without primes than with unrelated primes. Again, compared to the control condition the effect of priming was one of interference by the unrelated rather than facilitation by the related primes. The effect of prime relatedness was almost the same with (i.e., 620 – 596 = 24 ms) and without (i.e., 752 – 727 = 25 ms) preparation. Also, the effect of preparation was almost the same with related (i.e., 727 – 596 = 131 ms) and unrelated (i.e., 752 – 620 = 132 ms) primes.

The statistical analyses yielded main effects for preparation, $F_1(1, 17) = 87.38$, $p < .001$; $F_2(1, 8) = 537.00$, $p < .001$, and priming, $F_1(2, 34) = 52.52$, $p < .001$; $F_2(2, 16) = 20.21$, $p < .001$. Priming and preparation did not interact at all, $F_1(2, 34) < 1$, $p > .93$; $F_2(2, 16) < 1$, $p > .98$. Importantly, there was a clear priming effect in the homogeneous condition, $F_1(2, 34) = 23.87$, $p < .001$; $F_2(2, 16) = 10.27$, $p < .001$.

TABLE 5
Mean production latencies[a] (M), standard deviations (SD), and error percentages (E%)
as a function of preparation and priming in Experiment 3

| | Preparation | | | | | | | | |
| | Homogeneous | | | Heterogeneous | | | Total | | |
Priming	M	SD	E%	M	SD	E%	M	SD	E%
Related	596	107	2.9	727	97	2.0	661	121	2.4
Unrelated	620	121	3.7	752	83	3.1	686	123	3.4
Control	582	105	1.2	716	106	2.2	649	125	1.7
Total	600	112	2.6	731	97	2.4	665	124	2.5

[a]In ms.

In sum, the experiment allowed for preparation of a disyllabic phonological word and priming of a second phonological word. The results show, again, that preparation and priming effects do occur, and that the priming effect is not restricted to the heterogeneous condition (there is a clear priming effect of 24 ms in the homogeneous condition). This argues against the minimalist view on articulation initiation.

COMPUTER SIMULATIONS

In all three experiments, preparation and priming yielded main effects, and together their effects were additive. Furthermore, compared to the control condition the effect of priming was one of interference by the unrelated rather than facilitation by the related primes, in agreement with what Meyer and Schriefers (1991) observed in their three experiments for monosyllabic and disyllabic targets and spoken distractors. This was also observed by Schriefers and Teruel (1999). Whereas the current experimental findings pose difficulty for the minimalist view on the relationship between planning and articulation, they are consistent with the non-minimalist view.

In this section, the WEAVER++ model of speech production is shown to account for the experimental findings, not only for the main effects of priming and preparation but also for their combined effect, and for their effect relative to the control condition. This is demonstrated by discussing a computer simulation of Experiment 2 concerning the production of verb–particle constructions such as "zoek op" ("look up").

WEAVER++ is described in detail in many other places (e.g., Levelt et al., 1999a; Roelofs, 1992, 1993, 1996, 1997a, b, 1998, 1999; Roelofs & Meyer, 1998). Relevant for this study, the model assumes that word forms are represented in a network that is accessed by spreading activation. The segments of words are activated in parallel but are selected from "left to right" in order to construct phonological words. For example, in planning "zoek op", the /z/, /u/, /k/, /ɔ/, and /p/ are simultaneously active, but /z/ is selected before the /u/, and so forth. Articulation is not initiated before the first phonological word is ready. Planning may even go beyond the first phonological word in the utterance in order to respect the metrical cohesion of the items in small phonological phrases, such as the verb–particle constructions in Experiments 2 and 3 (see Roelofs, 1998, for discussion and converging evidence).

In the simulation, the homogeneous response sets consisted of "zoek op", "zoek af", and so forth. Recombining the responses of different homogeneous sets created the heterogeneous sets. As primes the final syllables of the utterances (i.e., the particle syllables *af*, *op*, and *uit*) were used. The prime corresponded either to a target syllable (e.g., *op* for "zoek op", the related condition) or to a syllable of another utterance in the response set (e.g., *af* for "zoek op", the unrelated condition), or no prime was given (the control condition). The critical items were embedded in a network of 50 words randomly selected from the CELEX lexical database (no embedding produced the same simulation outcomes). Advance knowledge about the form of the response was simulated by completing the morphological, phonological, and phonetic encoding of the utterance as far as possible before the beginning of a trial. A spoken prime was simulated by sequentially providing external input to the network. For example, spoken [ɔ] in the prime *op* activated the segment node /ɔ/ in the network, and [p] activated the /p/ node. The parameters for WEAVER++ that fit Experiment 2 were identical with fits of the model to various other sets of data (see Roelofs, 1992, 1993, 1996, 1997b; Roelofs & Meyer, 1998). To fit
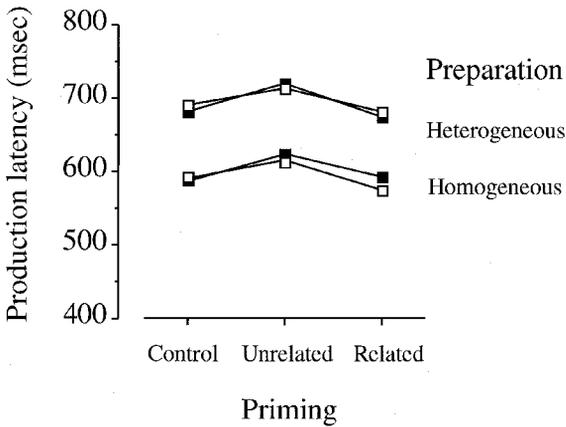
Figure 1.   Mean production latencies in milliseconds for Preparation (homogeneous, heterogeneous) × Priming (control, unrelated, related): ■ = Empirical (Experiment 2), ❑ = WEAVER++. To obtain the absolute production times, a constant of 300 ms was added to the encoding latencies of WEAVER++.

the six data points of the priming (related, unrelated, control) by preparation (homogeneous, heterogeneous) cells, only one free parameter was used, namely a constant that was added to the word-form encoding latencies of WEAVER++ to obtain the absolute production times.

Figure 1 shows the results of the simulations. As can be seen, the model yields main effects of preparation and priming, and together the effects are additive, as empirically observed. Furthermore, compared to the control condition, the effect of priming is one of interference by the unrelated rather than facilitation by the related prime, which also corresponds to the empirical results. Thus, the model captures the empirical findings.

Related primes yield facilitation compared to unrelated primes but not compared to the control condition with no spoken primes, both empirically and in the model. (For an extensive discussion of why this is the case in the model, see Roelofs, 1997b.) In short, a prime like *op* for "zoek op" not only facilitates the second syllable of the target (*op*) but also competes with the first syllable (*zoek*). The net result is the same as that obtained with presenting no prime. However, an unrelated prime like *af* competes with both the first (*zoek*) and the second syllable (*op*) of the target. Thus, relative to related primes and no prime, unrelated primes will yield interference, and related primes will yield "facilitation" compared to unrelated primes.

GENERAL DISCUSSION

In summary, the three experiments have shown that preparation and priming yield main effects and that together these effects are additive.[5] Most important, a priming effect is obtained even in the homogeneous condition. In contrast to Schriefers and Teruel (1999), who

---

[5]The number of items in each experiment is relatively small. However, despite the small number, the results from the by-item analyses matched those from the by-participant analyses. Furthermore, the same results were obtained in three experiments using different types of material. This suggests that the results can be generalized to other items and that they are not due to the particular items used in the experiments.

found a weak priming effect for second syllables and no effect for third syllables (for their two critical items), the present experiments did observe priming effects for second and third syllables (with nine items and more participants, and hence more power), even when they were part of the second phonological word. In the current experiments, priming and preparation effects were obtained when the primed and prepared syllables came from the same phonological word (Experiment 1) and when they came from different phonological words (Experiments 2 and 3). Also, priming and preparation effects were obtained when the preparable part was increased so that there was more time for the planning of the remainder during articulation of the prepared part (Experiment 3).

In all three experiments, the effect of priming was, compared to the control condition, one of interference by the unrelated rather than facilitation by the related primes (exactly as observed by Meyer & Schriefers, 1991, and Schriefers & Teruel, 1999). Given that the primes hampered responding, or at best, made no difference compared to no prime, it would have been a good strategy for participants to initiate articulation before planning the primed syllable. However, the results from all three experiments suggest that they could not adopt this avoidance strategy, but that initiation of articulation depended on planning the primed syllable, even when this was detrimental to the speed of performance.

Furthermore, the finding of interference by the unrelated primes rather than facilitation by the related primes, compared to no prime, suggests that the priming effect is not due to a conflict between a semantic cue provided by the prompt (pointing to one response) and the prime (pointing to another response). If priming were due to such a conflict, one would expect facilitation of a related prime relative to no prime, because a related prime points to the correct response, whereas there is no such help in the absence of a prime. If an unrelated prime is effective in indicating the wrong direction (as the conflict hypothesis assumes), a related prime should be effective in indicating the right direction. However, facilitation of related primes compared to no prime is not observed, which suggests that the priming effect is not due to a conflict between "cues" provided by the prompt and the prime.

Alternatively, auditory priming effects might perhaps reflect priming of "lemma retrieval" for production rather than word-form encoding. Verb–particle constructions such as "zoek op" ("look up") comprise two lemma nodes at the syntactic level, corresponding to the verb and the particle (Roelofs, 1998). Thus, for example, the spoken particle *op* may activate the lemma of *op*, thereby speeding up its selection. Although it cannot be excluded that priming of lemma retrieval has contributed to the effects in Experiments 2 and 3, it cannot account for the full effect because it leaves the effect in Experiment 1 unexplained, where the targets were simple words. The effect sizes of priming in Experiments 1 and 3 were similar (i.e., about 25 ms). However, the targets in Experiment 1 were single monomorphemic words such as *melon*, whereas the targets in Experiment 3 were particle verbs. Given the importance of word onsets in the spoken-word recognition process (e.g., Marslen-Wilson & Zwitserlood, 1989), it is unlikely that the lemma of the target *melon* will be activated to a high degree as part of the spoken cohort established by the prime *lon*. Studies of spoken-word recognition have shown that when the initial segments of a spoken prime and a target differ in more than two phonological features (such as voicing), no priming is observed (e.g., Connine, Blasko, & Titone, 1993; Marslen-Wilson, Moss, & van Halen, 1996; Marslen-Wilson & Zwitserlood, 1989). For example, Allopenna, Magnuson, and Tanenhaus (1998) observed in an eye-tracking study that hearing *collar* had little effect compared to hearing *dolphin* on the probability of fixating a

dollar presented as a visual target object. Given these findings in the literature, one expects much less activation (if any at all) of *melon* by the spoken prime *lon* than of *up* by the prime *up*, but the magnitude of the effect was the same for both types of prime. This suggests that the observed priming effects reflect priming of form encoding for production.

The observation of a priming effect in the homogeneous condition in all three experiments poses a challenge to minimalist theories of the relation between speech planning and the initiation of articulation. For example, MacKay's (1987; Santiago et al., 2000) NST assumes that words are planned via a tree traversal process that operates in a sequential top-down and left-to-right fashion. The initial syllable of a word is planned before its non-initial syllables, and within a syllable initial segments are planned before non-initial segments. Planning the left-most, bottom-most node (i.e., the first segment) determines the initiation of articulation. This theory may account for preparation effects by assuming that "sequential decisions can be executed in advance, during the preparation time" (Santiago et al., 2000, p. 13). However, because in the homogeneous condition, the critical first segment has already been planned and would trigger articulation, auditory priming of later syllables should not affect production onset latencies. Also, articulation initiation immediately after having planned the first segment fails to explain why the auditory priming effect has the same magnitude in the condition with prepared (homogeneous) and unprepared (heterogeneous) initial segments, as observed in the current experiments. For the same reason, the current findings pose difficulty for the minimalist PDP models of Jordan (1990) and Dell and colleagues (Dell et al., 1993).

Whereas in the PDP models of Jordan (1990) and Dell and colleagues (Dell et al., 1993), articulation is by design initiated when the features of the first segment have become activated, adopting articulatory buffering would seem to be compatible with NST, although Santiago et al. (2000) explicitly argued against it. Articulation may then be initiated when the buffer contains a phonological word or more rather than the initial segment or syllable. However, adopting these modifications for NST amounts to adopting a non-minimalist position, like that advocated in the current study.

Whereas the current findings pose a challenge to minimalist theories, they are fully compatible with theories that assume that the initial syllable is merely an optional unit (e.g., Schriefers & Teruel, 1999) and theories that assume that the lower limit for articulation is the phonological word (e.g., Levelt et al., 1999a). This latter view is implemented in the WEAVER++ model of speech production (e.g., Levelt et al., 1999a; Roelofs, 1992, 1993, 1996, 1997a, 1999). The model simulated the experimental findings, not only the main effects of priming and preparation but also their combined effect. In WEAVER++, the initiation of articulation occurs not earlier than completion of the construction of a phonological word, but the encoder may occasionally opt for a larger unit, for example, to respect the metrical cohesion of the items in a phonological phrase (Roelofs, 1998). After all, the task of the form encoder is to generate a pronounceable articulatory programme for the whole utterance and not to maximize the speed of responding per se (see Levelt, 1989, for an extensive discussion of this point).

The current study concentrated on spoken-word production in picture naming and related tasks, such as the preparation task used in Experiments 1–3. The issue of planning and articulation initiation has also been investigated within the context of translating from spelling to sound in oral reading. Kawamoto, Kello, Jones, and Bame (1998) tested between initial segment and whole-word criteria for articulation initiation in reading words. Participants had to

read aloud words with regular versus irregular vowel pronunciations (e.g., *pump* vs. *pint*). The results were equivocal: Response onset latencies were longer for irregularly than for regularly spelled words (as predicted by the whole-word criterion), but initial-segment pronunciation durations were *also* longer for the words with irregular spellings (as predicted by the initial-segment criterion), although this duration effect was not always reliable by items. According to Kawamoto et al., these results suggest that some readers might be using the whole-word criterion, whereas others use the initial-segment criterion, or that the criteria vary between trials.

Kawamoto (1999) conjectured that an initial-segment criterion might explain the results from the preparation paradigm. As discussed by Levelt, Roelofs, and Meyer (1999b), there are several problems with this proposal. Most relevant in the current context is that Kawamoto's proposal fails to explain the finding that the preparation effect increases when more segments are shared (Meyer, 1990, 1991), even when the shared part crosses a phonological word boundary (Roelofs, 1998). Furthermore, Meyer (1990, 1991) observed no difference in pronunciation durations between preparation conditions, but only an effect on articulation onset latency. Finally, Kawamoto's proposal fails to explain the current findings.

To conclude, the findings from Experiments 1–3 suggest that articulation is not inevitably initiated when the first segment has been planned. The findings from Experiment 1 further suggest that the unit for articulation initiation may be as large as the phonological word. In addition, the results from Experiments 2 and 3 suggest that speakers may even plan beyond the first phonological word in utterances at the cost of a delay in the initiation of articulation, showing that the phonological word is also not an obligatory unit for articulation initiation.

REFERENCES

Allopenna, P.D., Magnuson, J.S., & Tanenhaus, M.K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language*, *38*, 419–439.

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

Bachoud-Lévi, A.-C., Dupoux, E., Cohen, L., & Mehler, J. (1998). Where is the length effect? A cross-linguistic study of speech production. *Journal of Memory and Language*, *39*, 331–346.

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

Connine, C.M., Blasko, D.G., & Titone, D. (1993). Do the beginnings of spoken words have a special status in auditory word recognition? *Journal of Memory and Language*, *32*, 193–210.

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.

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: Lawrence Erlbaum Associates, Inc.

Kawamoto, A.H. (1999). Incremental encoding and incremental articulation in speech production: Evidence based on response latency and initial segment duration. *Behavioral and Brain Sciences*, *22*, 48–49.

Kawamoto, A.H., Kello, C.T., Jones, R., & Bame, K. (1998). Initial phoneme versus whole-word criterion to initiate pronunciation: Evidence based on response latency and initial phoneme duration. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 862–885.

Lashley, K.S. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior* (pp. 112–136). New York: Wiley.

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

Levelt, W.J.M, Roelofs, A., & Meyer, A.S. (1999a). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22,* 1–38.

Levelt, W.J.M., Roelofs, A., & Meyer, A.S. (1999b). Multiple perspectives on word production. *Behavioral and Brain Sciences*, *22*, 61–75.

Lupker, S.J., Brown, P., & Colombo, L. (1997). Strategic control in a naming task: Changing routes or changing deadlines? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 570–590.

MacKay, D. G. (1987). *The organization of perception and action: A theory for language and other cognitive skills*. New York: Springer–Verlag.

Marslen-Wilson, W., Moss, H.E., & van Halen, S. (1996). Perceptual distance and competition in lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 1376–1392.

Marslen-Wilson, W.D., & Zwitserlood, P. (1989). Accessing spoken words: The importance of word onsets. *Journal of Experimental Psychology: Human Perception and Performance*, *15*, 576–585.

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., Roelofs, A., & Levelt, W.J.M. (1999). *Word length effects in picture naming: The role of a response criterion*. Manuscript submitted for publication.

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.

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. (1996). Serial order in planning the production of successive morphemes of a word. *Journal of Memory and Language*, *35*, 854–876.

Roelofs, A. (1997a). Syllabification in speech production: Evaluation of WEAVER. *Language and Cognitive Processes*, *12*, 659–696.

Roelofs, A. (1997b). The WEAVER model of word-form encoding in speech production. *Cognition*, *64*, 249–284.

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. (1999). Phonological segments and features as planning units in speech production. *Language and Cognitive Processes*, *14*, 173–200.

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.

Santiago, J., MacKay, D.G., Palma, A., & Rho, C. (2000). Sequential activation processes in producing words and syllables: Evidence from picture naming. *Language and Cognitive Processes*, *15*, 1–44.

Schriefers, H., & Teruel, E. (1999). Phonological facilitation in the production of two-word utterances. *European Journal of Cognitive Psychology*, *11*, 17–50.

Sternberg, S., Knoll, R.L., Monsell, S., & Wright, C.E. (1988). Motor programs and hierarchical organization in the control of rapid speech. *Phonetica*, *45*, 175–197.

Van der Plaats, R.E., & Van Galen, G.P. (1990). Effects of spatial and motor demands in handwriting. *Journal of Motor Behavior*, *22*, 361–385.

Wheeldon, L., & Lahiri, A. (1997). Prosodic units in speech production. *Journal of Memory and Language*, *37*, 356–381.