The WEAVER model of word-form encoding in speech production

Ardi Roelofs*

Max Planck Institute for Psycholinguistics, P.O. Box 310, 6500 AH Nijmegen, The Netherlands

Abstract

Lexical access in speaking consists of two major steps: lemma retrieval and word-form encoding. In Roelofs (Roelofs, A. 1992a. Cognition 42, 107–142; Roelofs, A. 1993. Cognition 47, 59–87.), I described a model of lemma retrieval. The present paper extends this work by presenting a comprehensive model of the second access step, word-form encoding. The model is called WEAVER (Word-form Encoding by Activation and VERification). Unlike other models of word-form generation, WEAVER is able to provide accounts of response time data, particularly from the picture-word interference paradigm and the implicit priming paradigm. Its key features are (1) retrieval by spreading activation, (2) verification of activated information by a production rule, (3) a rightward incremental construction of phonological representations using a principle of active syllabification, syllables are constructed on the fly rather than stored with lexical items, (4) active competitive selection of syllabic motor programs using a mathematical formalism that generates response times and (5) the association of phonological speech errors with the selection of syllabic motor programs due to the failure of verification. © 1997 Elsevier Science B.V.

Keywords: WEAVER; Word-form encoding; Lemma retrieval; Speech production

1. Introduction

The production of speech is one of our most complex psychomotor skills. An adult speaker effortlessly produces two to five word forms per second. A sound error is made only once per 2000 words, approximately (cf. Levelt, 1989). As with other

* E-mail: ardi@mpi.nl

skilled performance, speech production requires advance planning or encoding (e.g., Lashley, 1951). How is the efficient encoding of forms achieved?

The encoding of word forms is one of the central topics in language production research and has been studied intensively in the last decade (see e.g., Levelt, 1989 for a review). Existing models have been built around speech errors (e.g., Dell, 1986, 1988; Shattuck-Hufnagel, 1979; Stemberger, 1985). Errors have been very informative about the cognitive systems underlying speech production (e.g., Meyer, 1992). Below, I argue that error-based models have nevertheless neglected some issues in speech production. A new model is presented that deals with these issues.

The model has been built around chronometric data rather than slips of the tongue. Chronometric studies of cognitive processes have a long tradition (e.g., Donders, 1868; Luce, 1986; Posner, 1978). Only recently, however, researchers have started to apply such techniques to language production. New insights into form encoding have been obtained (e.g., Levelt, 1989, 1992; Levelt and Wheeldon, 1994; Meyer, 1990, 1991; Meyer and Schriefers, 1991; Wheeldon and Levelt, 1995). As chronometric data begin to accumulate in the literature, there is a growing need for formal models that can provide a general unifying account of the data.

This paper presents such a formal model and its simulation. The model is called WEAVER, which stands for Word-form Encoding by Activation and VERification. WEAVER adopts Dell's (1986) assumption of form retrieval by spreading activation and Levelt's (1992) assumption of on-line syllabification and syllabary access. The encoding of word forms is achieved by a spreading-activation based form network (cf. Dell, 1986) integrated with a parallel object-oriented production system embodying linguistic rules or constraints (cf. Levelt, 1992). Words are not planned by a central agent that overlooks the whole process but by a team of procedures that work in parallel on small parts of the word, like several spiders making a single web. WEAVER generates word forms in a rightward incremental fashion by spreading of activation in the network and by verifying whether integration of activated elements into the developing form is licensed. The focus on response-time data does not imply any prejudice against speech-error data. The ultimate model should be able to handle all data, of all types. WEAVER has limitations but is argued to be a step towards a computationally and psychologically adequate model of form encoding. I show that WEAVER explains findings about error-free production and that it is compatible with the classical observations on the occasional derailment of the encoding process.

Existing models of speech production assume that the encoding of word forms involves mapping a representation of the word as a syntactic entity, the word's 'lemma', onto an articulatory program (e.g., Dell, 1986, 1988; Levelt, 1989, 1992; Shattuck-Hufnagel, 1979). The word's morphemes and their segments are recovered from memory and used to construct an appropriate motor program. During this process, segments are assigned to production slots or positions. For example, in producing the Dutch word *hamer* (*hammer*), the morpheme <hamer> and the segments /h/, /a/, /m/, /ə/, and /r/ are retrieved from memory. Subsequently, the segments are made onset and nucleus of a first syllable (i.e., *ha*), and onset, nucleus, and coda of a second syllable (i.e., *mer*), respectively. These syllables are then used to derive the corresponding articulatory program.

In the classical model (Dell, 1986), word forms are represented in a network with nodes for morphemes (e.g., <hamer>), syllables (i.e., *ha* and *mer*), rimes, segments and segment clusters labeled for syllable position (e.g., /onset h/, /nucleus a/, /onset m/, /nucleus ə/, /coda r/), and features (e.g., anterior, coronal). Morpheme nodes are connected to header and category nodes that represent their CV structure (Dell, 1988). Nodes are associated to each other by weighted bidirectional connections. Fig. 1 illustrates the form representation of *hamer* (in order to simplify the figure, I have left out rime and feature nodes).

The encoding starts by supplying a jolt of activation to a lemma node. Activation then spreads through the network following a linear activation function with a decay factor. Morphological and phonological encoding are accomplished by successively selecting the most active morpheme and segment nodes at certain moments in time. The time interval between the successive selections has a constant duration whose value depends on the speech rate. The segments for a syllable are selected following the order of activation of the category nodes (Dell, 1988). Selected segment nodes are inserted into labeled slots of syllable frames that are generated by rule (Dell, 1986). For example, the /onset h/ is inserted into the onset slot of the first syllable frame for *hamer*. To prevent that the /onset m/ of *hamer* or the onset of another morpheme is inserted, morphemes and syllables are serially encoded. The serial encoding of a polysyllabic word such as *hamer* is achieved by invoking a procedure that verifies the order of the syllables and temporarily increases the spreading rate between <hamer> and the first target syllable *ha* and decreases it for the second syllable *mer*, and vice versa in encoding the second syllable. Although it seems clear that phonetic encoding involves mapping the selected segments onto feature nodes, the mechanism that achieves this is left unspecified.
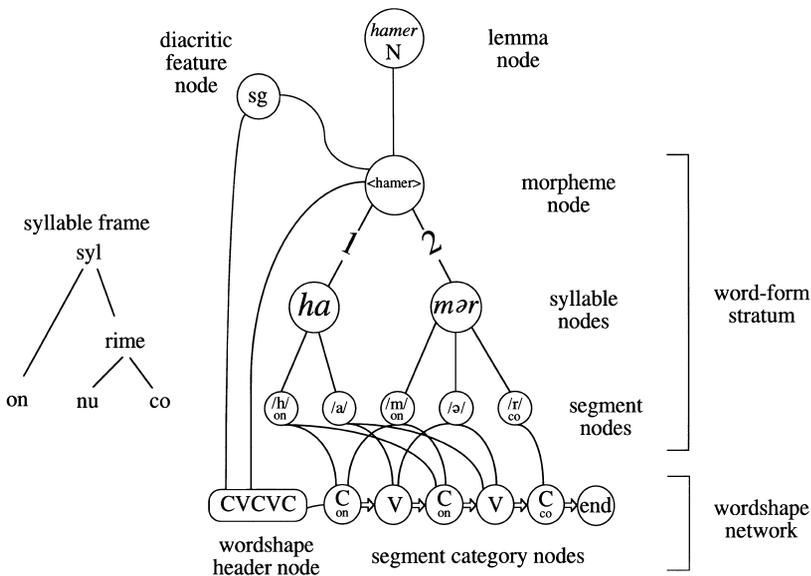

Fig. 1. Memory representation of the word form of *hamer* in the classical model (Dell, 1986, 1988).

1.1. Three issues motivating the development of WEAVER

1.1.1. Binding and latencies

Speakers plan forms while simultaneously listening to other speakers or while planning aspects of other forms (e.g., Levelt, 1989). Thus, they have to keep track of what goes with what. In existing error-based models, binding of speech segments to production slots is achieved by severe temporal constraints in order to prevent massive slips of the tongue (e.g., Dell, 1986, 1988). Consider the production of the Dutch utterance 'leg mijn hamer in de gereedschapskist' ('put my hammer in the tool box'). In producing this utterance, the segments of the morphemes <mijn>, <hamer>, and <in> have to be retrieved from memory, during which the speech production system has to keep track of what goes with what. It has to know that the /h/ is retrieved for <hamer> and the /m/ for <mijn>, otherwise the /m/ and /h/ may trade places and 'hijn mamer' might be produced instead of 'mijn hamer'. To prevent errors, existing models assume that segments are spelled out one morpheme or syllable at a time. First <mijn> is encoded, and next <hamer>. During the encoding of <mijn>, only /m/, /ɛi/, and /n/ are available, so the /h/ of <hamer> will not be selected. This may be called 'binding by timing'. Binding by timing is at the heart of the traditional account of speech errors. Errors occur during the rare event that, e.g., as a result of random noise in the system, planning another form, or hearing another speaker, a segment of another word is the most highly activated one and gets erroneously selected. For example, an error occurs when during the encoding of <mijn>, the /h/ of <hamer> has a higher level of activation than the /m/ for <mijn>, and the /h/ gets selected as the onset of <mijn>.

Chronometric studies of speech production have typically looked at facilitation and inhibition priming of production. These studies have shown that speakers do not begin to make errors when multiple word forms are available simultaneously under experimental conditions. For example, Meyer and Schriefers (1991) have shown that when Dutch speakers have to name a pictured object (e.g., a hammer) and they hear a word sharing some of its segments (e.g., /zomər/ (*summer*)), this latter word speeds up naming the object compared to an unrelated word instead of causing trouble in selecting the correct segments. The problem of selecting the correct segments is especially salient for models that assume that the same form lexicon underlies speech production as well as speech perception (e.g., MacKay, 1987) or models in which parallel activation of word forms leads to blending of these forms (e.g., Dell et al., 1993). Perceptual input can lead to errors in object naming and in spontaneous speech (e.g., Harley, 1984), but such errors are rare.

The classical model (Dell, 1986, 1988) has not been applied to multiple-input situations requiring filtering and has no competition-sensitive response time mechanism. In this model, selection takes place on an ordinal basis as the most highly activated node is selected, the absolute level of activation being irrelevant. Consequently, priming may affect levels of activation, but will not affect latencies. In the model, the target node is either the most highly activated node at the predetermined moment of selection and gets selected or it is not the most highly

activated node and an error occurs. Peterson et al. (1989) proposed a model with a competition-sensitive response time mechanism that attributed most response time effects to monitoring. Simulations showed that with high-frequency targets and distractors, the probability (P) of selection of a critical target segment in their model was 0.45. However, in picture-word interference experiments the error rate is about 5%. To account for the latency effect in the absence of errors, Peterson et al. proposed that speakers monitor their responses for errors prior to articulation and repair the errors. Error correction takes time, which may account for the latency effects. However, merely postulating a monitor postpones the binding problem. How does the monitor establish what goes with what so that it can detect errors? Furthermore, as we will see, such a monitor looks like the solution embodied in the WEAVER model. WEAVER does not blindly select an activated node but verifies whether weaving the node into the phonetic plan is licensed.

1.1.2. Cross-morpheme and cross-word syllabification

The task for a binding mechanism is more complex than keeping the segments of different morphemes and words apart because binding is context dependent. Sometimes a segment of one morpheme or word must be bound to another morpheme or word and change syllable position. This may occur in the production of polymorphemic words or connected speech (e.g., Chomsky and Halle, 1968; Kaisse, 1985; Levelt, 1989, 1992; Nespor and Vogel, 1986; Selkirk, 1984). Cross-morpheme and cross-word syllabification point to the need to deal with flexibility of syllable membership.

Consider the production of the infinitival form of the denominalized verb *hameren* (*to hammer*). The infinitive *hameren* is created by adding <en> ([ə]) to the stem <hamer>. The resulting form is syllabified as (ha)$_\sigma$(mə)$_\sigma$(rə)$_\sigma$. Thus, juxtaposing *-en* changes the syllabification of /r/. This segment occupies a coda position in *hamer*, syllabified as (ha)$_\sigma$(mər)$_\sigma$, but an onset position in *hameren*. Or consider the production of the utterance 'leg mijn hamer in de gereedschapskist'. To increase the speed and fluency of articulation, a speaker might syllabify the utterance as '… (ha)$_\sigma$ (mə)$_\sigma$(rɪn)$_\sigma$…' instead of '… (ha)$_\sigma$(mər)$_\sigma$(ɪn)$_\sigma$ …'. The coda /r/ of the second syllable of *hamer* is made the onset of the syllable *rin*. The lexical words *hamer* and *in* are combined into the new phonological word *hamerin*. The creation of a new phonological word is optional in this example, but it is obligatory for clitics. Clitics are function words such as pronouns, determiners, particles, auxiliary verbs, prepositions, and conjunctions, which unlike words of major lexical categories, 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. In producing 'geef me de hamer's' ('please give me the hammer now'), '*s* is adjoined to *hamer*. This yields the new phonological word *hamer's*, which is syllabified as (ha)$_\sigma$(mə)$_\sigma$(rəs)$_\sigma$. Models that rigidly store words as sequences of syllable nodes or models that store each consonant as an onset or coda have a difficult time dealing with the need for flexibility of syllable membership (e.g., Dell, 1986, 1988; Houghton, 1990; Shattuck-Hufnagel, 1979).

1.1.3. Phonetic encoding

Existing models do not typically have a phonetic level of encoding, and therefore do not account well for assimilation and allophonic variation of speech segments. For example, although the classical model has a level of feature nodes connected to segment nodes, it does not address phonetic encoding. The same holds for MacKay's model (MacKay, 1987), in which segment nodes are connected via features to articulatory routines. Mapping each segment onto its own set of features or routines would clearly be insufficient, because the realization of a feature depends on the context in which it is spoken. For example, in certain dialects of Dutch (as in English), plosives are aspirated in syllable onset position. Dutch devoices syllable-final obstruents, but such segments may receive the feature of voicing again in certain contexts. For example, under the influence of the /b/ of *boek*, *handboek* (*handbook*) is pronounced as [hɑnd.buk] instead of [hɑnt.buk]. Furthermore, a phonetic representation has to specify the temporal relationships between the features within and between segments (e.g., Browman and Goldstein, 1986). However, the models do not specify relationships between features. Features are linked to segments only. Of course, there are models of phonetic encoding (e.g., Jordan, 1990). However, there are no proposals that are an integral part of broader models (e.g., Dell, 1986, 1988).

1.2. New assumptions made to deal with the issues

Following the classical model, WEAVER assumes that the mental lexicon is a network of nodes and links that is accessed by spreading activation. In order to deal with the three issues, a few new assumptions are made.

WEAVER deals with binding through explicitly representing the relationship between morphemes and segments and verification of the relations (cf. Anderson, 1983; Collins and Loftus, 1975). For example, the link between /m/ and <mijn> says that /m/ is its first segment. In contrast, the classical model knows that segments belong to a morpheme or syllable because they are activated at the same time.

Explicitly representing relationships in spreading-activation models has a long tradition. Collins and Loftus (1975) and others used labeled links in network models of semantic memory, and Anderson (1983) used them in a production-system model of cognition. Also, the classical production model (Dell, 1986) marks links between morpheme and syllable nodes to indicate their serial position. In Dell's model, relations are verified in order to change the spreading rates for the syllables. In encoding *hamer*, the spreading rate is first increased for *ha* and decreased for *mer*, and later vice versa, so that the segments are activated at the appropriate time.

Verification in WEAVER differs from that in the classical model in two major respects. First, in WEAVER all links are labeled, whereas in the classical model there are no labels on the links except for polysyllabic words. Thus, WEAVER represents relations in a more principled way. Second, WEAVER directly verifies the relationship between morphemes and segments to achieve binding. In contrast, in the classical model, binding between morphemes and segments is accomplished in two steps. First, the relation between morphemes and syllables is verified in order

to change the spreading rates. Second, by changing the rates, binding between morphemes and segments is achieved. Thus, by verifying the relation between morphemes and segments directly, WEAVER avoids the extra step of changing spreading rates.

To account for latency effects, WEAVER implements active competitive selection of syllabic motor programs using a mathematical formalism that generates response times. How fast a motor program is accessed depends on how active other programs are. Motor programs for syllables are selected for constructed phonological representations, which brings us to the flexibility of syllable membership.

WEAVER deals with the issue of syllabification across morpheme and word boundaries by computing (e.g., Béland et al., 1990; Caplan, 1992; Levelt, 1992) instead of storing syllabifications (e.g., Dell, 1986, 1988; Shattuck-Hufnagel, 1979). The on-line syllabification takes neighboring morphemes and words into account in that syllable positions are computed for phonological words rather than for lexical ones (Booij, 1983, 1995; McCarthy and Prince, 1990). Syllabification proceeds in a rightward incremental fashion (Levelt, 1992; Wheeldon and Levelt, 1995).

WEAVER deals with the issue of phonetic encoding by connecting the lexical network to a phonetic syllabary, as proposed by Levelt (Levelt, 1989, 1992; Levelt and Wheeldon, 1994; see also Crompton, 1981). It is generally assumed that the ultimate motor programs are not constructed from scratch (e.g., Jordan, 1990; Levelt, 1989, 1992). Instead, when available, motor programs are retrieved from a store of learned programs. Learned motor programs must present a set of retrieval cues to higher-level processes so that the appropriate program can be accessed. Ideally, these retrieval cues constitute a reasonably small set (cf. Jordan, 1990). Levelt (Levelt 1992; Levelt and Wheeldon, 1994) proposed that the retrieval cues for articulatory programs are phonological syllables, which are constructed as part of phonological word representations. Retrieving motor programs is syllabary access. Statistical analyses by Schiller et al. (1996) revealed that 500 different syllables cover almost 85% of the Dutch syllable tokens (and, e.g., 80% of the English).

Syllabary access translates an abstract phonological representation into a context-dependent phonetic representation. The articulatory program for *hamer*, [ʹha] [mər], comprises motor programs for the syllables [ha] and [mər], where [ha] and [mər] stand for packages of scores for the articulatory movements to be made. A score specifies the gestures and their temporal relationships (e.g., Browman and Goldstein, 1986; Levelt, 1989, 1992). Scores make explicit articulatory tasks, such as lip protrusion and lowering of the jaw. The details of the movements realizing these scores are left to the articulatory system. Importantly, assimilation phenomena may result automatically from overlapping gestural scores (e.g., Browman and Goldstein, 1986). A full account of the motivation of a phonetic syllabary can be found elsewhere (i.e., Levelt, 1989, 1992; Levelt and Wheeldon, 1994).

1.3. Organization of the paper

In Section 2, I further explain WEAVER. Next, I show how the model accounts

for chronometric findings from picture-word interference (Section 3) and implicit priming (Section 4). In Section 5, I illustrate the model's approach to speech errors. Section 6 gives an overview of tests of WEAVER on new sets of data. Finally, Section 7 provides a general discussion.

2. Theoretical assumptions

2.1. Lexical access: lemma retrieval and word-form encoding

To set the stage, I discuss how word-form encoding fits into lexical access in speaking, the process by which a lexical concept is mapped onto an articulatory program. This task is carried out in two major steps: lemma retrieval and word-form encoding (e.g., Butterworth, 1989; Dell, 1986; Garrett, 1975; Kempen and Huijbers, 1983; Kempen and Hoenkamp, 1987; Levelt, 1989, 1992; Roelofs, 1992a,b, 1993; Roelofs et al., 1996). In lemma retrieval, a lexical concept is used to retrieve the syntactic properties of a word and to provide pointers to its form. A lemma is a memory representation of the syntactic properties of a word. For example, the lemma of the Dutch word *hamer* indicates that it is a noun and that its grammatical gender is nonneuter. In word-form encoding, the form pointers are used to recover the corresponding morphophonological properties (the word's 'lexeme') from memory to compute an articulatory program.

Assume that a Dutch speaker wants to name a hammer. Lexical access consists of mapping the lexical concept HAMMER(X) onto the articulatory program for *hamer*, which comprises programs for the syllables [ha] and [mər].

The lemma retriever takes HAMMER(X) and outputs the lemma of *hamer* (for a theory and computer model of lemma retrieval, see Roelofs, 1992a,b, 1993, 1996c, 1997b). To derive the singular form [hamər] instead of the plural form [hamərs], the word's number has to be specified. The lemma plus this number diacritic are then input to word-form encoding. The articulatory program is derived in three major steps (cf. Dell, 1986; Levelt, 1989, 1992): morphological, phonological, and phonetic encoding (Fig. 2). The morphological encoder takes the lemma of *hamer* plus the diacritic singular and produces the stem morpheme <hamer>. This first stage thus concerns what is traditionally called the 'syntax-morphology interface' (e.g., Spencer, 1991). The phonological encoder takes the stem morpheme and produces a phonological word representation. This representation describes the singular form of *hamer* as a phonological word consisting of a trochaic foot (i.e., the first syllable is metrically strong, s, and the second is weak, w). The first syllable has /h/ as onset and /a/ as nucleus. The second syllable has /m/ as onset, /ə/ as nucleus, and /r/ as coda. This second stage thus comprises what is traditionally called the 'morphology-phonology interface' (e.g., Goldsmith, 1990; Kenstowicz, 1994). Finally, the phonetic encoder takes this phonological word representation, accesses the corresponding syllable programs in the phonetic syllabary, and delivers the articulatory program, [′ha][mər]. This final encoding stage includes what is sometimes called the 'postlexical phonology' (e.g., Goldsmith, 1990).

The mental lexicon is assumed to be a network consisting of three strata: a conceptual stratum with lexical-concept nodes and links (e.g., HAMMER(X), TOOL(X), IS-A); a syntactic stratum with lemma nodes (e.g., *hamer*), nodes and links for syntactic properties (e.g., LEXICAL_CATEGORY: noun), and slots and fillers for diacritics (e.g., HAS-NUMBER: sg, pl); and a word-form stratum.

2.2. Declarative memory and production application

Fig. 3 illustrates the memory representation of the form of the Dutch word *hamer* in WEAVER. The form network consists of three layers of nodes: morpheme, segment, and syllable program nodes. Morpheme nodes stand for roots and affixes, connected to the lemma and its diacritics. The stem <hamer> is connected to the lemma of *hamer* and singular. A morpheme points to its segments and its metrical structure, which describes an abstract grouping of syllables (σ) into feet (Σ) and feet into phonological words (ω) (cf. Liberman and Prince, 1977). The links between morpheme and segment nodes indicate the serial position of the segments. Possible syllable positions (onset, nucleus, coda) of the segments are specified by the links between segment and syllable program nodes. For example, /r/ is the coda of [mər] and the onset of [rɪn]. Fig. 3 indicates for /m/ that phonetic information is also available from segment nodes without mediation by a syllable program node. This direct route may be used when a segment's features have to be consulted during the syllabification process or in the generation of an articulatory program when there is no stored program for a phonological syllable.
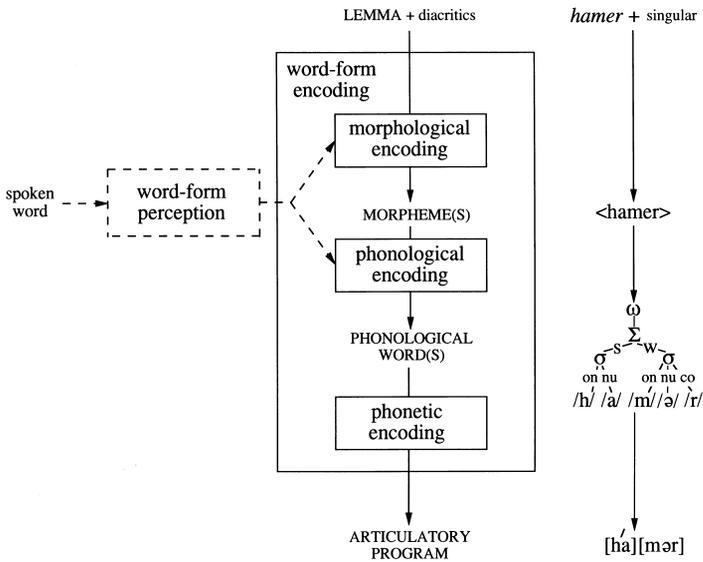

Fig. 2. Stages of word-form encoding in WEAVER. The dashed part indicates which of the levels of form encoding receive input from a spoken distractor word.

Encoding starts when a morpheme node receives activation from a lemma. Activation then spreads through the network in a forward fashion, each node sending a proportion of its activation to its direct neighbors. There is also spontaneous decay of activation. The form encoders follow simple selection rules, which are implemented in a parallel distributed manner. Attached to the nodes in the network, there are production rules (i.e., condition-action pairs) that select nodes if they are appropriately linked to the target nodes one level up. Thus, a representation scheme is proposed that is sometimes referred to as object oriented (cf. Bobrow and Winograd, 1977). The application of production rules forms verification. A production rule is triggered when the activation levels of its nodes exceed threshold. Productions may operate in parallel. Similar production systems have been proposed by Anderson (1983), Kempen and Hoenkamp (1987), and Newell and Simon (1972), among others.

2.2.1. Morphological encoding

A morphological production selects a morpheme node if it is linked to a selected lemma and its diacritics. For example, <hamer> is selected for *hamer* and singular.

2.2.2. Phonological encoding

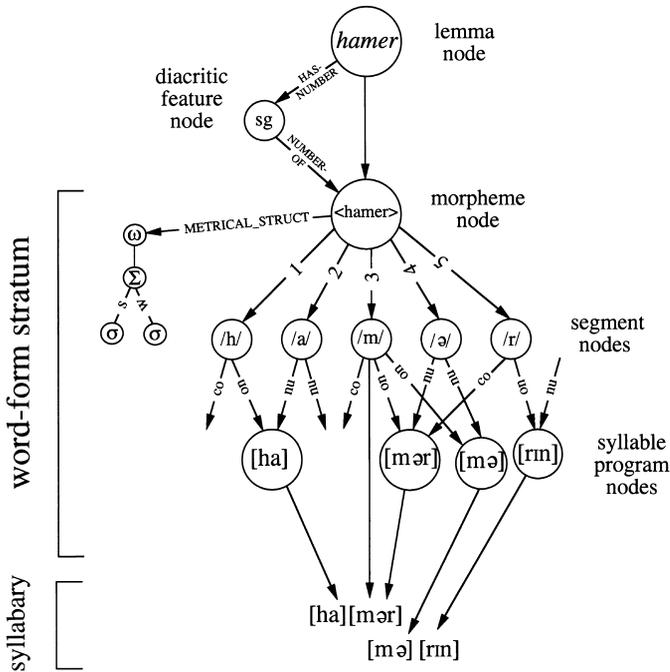A phonological production selects segments and a metrical structure if they are

Fig. 3. Memory representation of the word form of *hamer* in WEAVER.

linked to the selected morphemes. Next, the segments and metrical structure are input to a syllabification process that associates the segments to the syllable nodes within the metrical structure. Like weaving a fabric, the process has a certain direction. The syllabification proceeds from the first segment to the second, and so forth, whereby syllable positions (onset, nucleus, coda) are assigned following the syllabification rules of the language. Essentially, each vowel and diphthong is assigned to a different syllable node and consonants are treated as onsets unless phonotactically illegal onset clusters arise. In the encoding of <hamer>, the /h/ is made syllable onset and the /a/ nucleus of the first syllable, and the /m/ onset, the /ə/ nucleus, and the /r/ coda of the second syllable.

The on-line assignment of syllable positions provides for cross-morpheme and cross-word syllabification. In planning polymorphemic words or connected speech, adjacent morphemes or words may be syllabified together. For example, WEAVER may group the segments of the stem <hamer> and the suffix <en> together for the infinitive of the verb *hameren*, or may join the segments of <hamer> and <in> for the cliticization *hamerin*. Then, applying the syllabification rules, /r/ will be made onset of the third syllable instead of coda of the second syllable, yielding $(ha)_\sigma(mə)_\sigma(rə)_\sigma$ and $(ha)_\sigma(mə)_\sigma(rɪn)_\sigma$, respectively.

2.2.3. Phonetic encoding

A phonetic production selects a syllable program node if its links to the segments match the syllable positions that were on-line assigned to the segments. For example, [mər] is selected for the second phonological syllable of *hamer*, $(mər)_\sigma$, because the relation between [mər] and /m/ is onset, between [mər] and /ə/ nucleus, and between [mər] and /r/ coda. Similarly, the phonetic encoder selects [mə] and [rə] for *hameren* and [mə] and [rɪn] for *hamerin*.

When the condition of a syllable program node is satisfied, the actual moment at which syllabary access takes place is randomly distributed. How quickly access is initiated depends on how active other program nodes are. The probability of initiating access at a particular moment in time is equal to the ratio of the node's level of activation and the sum of the activation levels of all syllable program nodes in the network (cf. Roelofs, 1992a,b, 1993). Appendix A gives the equations that formalize WEAVER. After accessing the programs, the phonetic encoder uses the metrical representation to set their parameters for loudness, pitch, and duration, and makes the programs available for the control of the articulatory movements.

WEAVER also provides for a suspend/resume mechanism that supports incremental generation of word-form plans. Incremental production means that encoding processes can be triggered by a fragment of their characteristic input (cf. Levelt, 1989). For example, syllabification can start on the initial segments of a word without having all its segments. Only initial segments and the metrical structure are needed. The resulting partial representation can be buffered until the missing segments are available and syllabification can continue. Thus, computations are completed as far as possible, after which they are put on hold. When given further information, they continue from where they stopped.

3. Application to picture-word interference

3.1. Basic empirical findings

Meyer and Schriefers (1991) examined the effect of spoken distractor words on word-form encoding in object naming. The target and distractor words were either monosyllables or disyllables. The monosyllabic targets and distractors shared either the onset and nucleus (begin-related) or the nucleus and coda (end-related). For example, Dutch speakers had to name a pictured book (i.e., they had to say *boek*, [buk]; *book*), where the distractor was *boeg* ([bux], begin-related; *bow*), *doek* ([duk], end-related; *cloth*) or there was no distractor (silence condition). Unrelated control conditions for the related ones were created by re-combining pictures and distractors. The disyllabic targets and distractors shared either the first syllable (begin-related) or the second syllable (end-related). For example, the speakers had to name a pictured hammer (i.e., they had to say *hamer*, ['ha.mər]; *hammer*), where the distractor was *havik* (['ha.vɪk], begin-related; *hawk*), *zomer* (['zo.mər], end-related; *summer*) or there was no distractor.

The distractor words were presented just before (i.e. −300 or −150 ms), simultaneously with, or right after (i.e., +150 ms) picture onset, called the stimulus onset asynchrony (SOA). The SOAs in the begin and end conditions were either aligned or non-aligned. In a non-aligned experiment, the SOA was defined as the interval between the onset of the spoken distractor word and picture onset. In an aligned experiment, the SOA was defined as the interval between the onset of the critical part of the distractor and picture onset. In begin-related trials, therefore, the SOA was between the onset of both the word and the critical segment (e.g., the [b] of [bux]) and picture onset in both aligned and non-aligned experiments, but in end-related trials, the SOA was between the onset of the critical segment (e.g., the [u] of [duk]) and picture onset in the aligned experiment, but between word onset (e.g., the [d] of [duk]) and picture onset in the non-aligned experiment. The same held for the unrelated controls.

The presentation of spoken distractors yielded longer object naming latencies compared to the situation without a distractor (cf. Glaser, 1992). The naming latencies were prolonged less with related distractors than with unrelated ones (cf. Schriefers et al., 1990). Thus, a facilitatory effect was obtained from word-form overlap relative to the non-overlap situation. (Levelt et al. (1991) obtained inhibition from form overlap, but this was with a different task, namely lexical decision on auditory probes presented during picture naming).

Fig. 4 shows the SOA curves of the facilitatory effects in the experiments. For each SOA, the mean naming time in the related condition minus the mean naming time in the unrelated condition is plotted. Thus, a negative difference indicates facilitation from word-form overlap relative to the situation without overlap.

In the aligned experiment, the difference between begin and end overlap for both the monosyllables and the disyllables was in the SOA at which the facilitation was first detected, hereafter referred to as the 'onset' of the facilitatory effect. The onset of the effect in the begin condition was at SOA = −150, whereas the onset of the

effect in the end condition was at SOA = 0. With both begin and end overlap the facilitatory effect was still present at the SOA of +150.

In contrast, in the non-aligned experiment, the difference between begin and end overlap for the monosyllables was in the SOA at which the facilitation disappeared, hereafter referred to as the 'offset' of the facilitatory effect. Facilitation began at SOA = −150 in both the begin and end condition. However, with begin overlap the
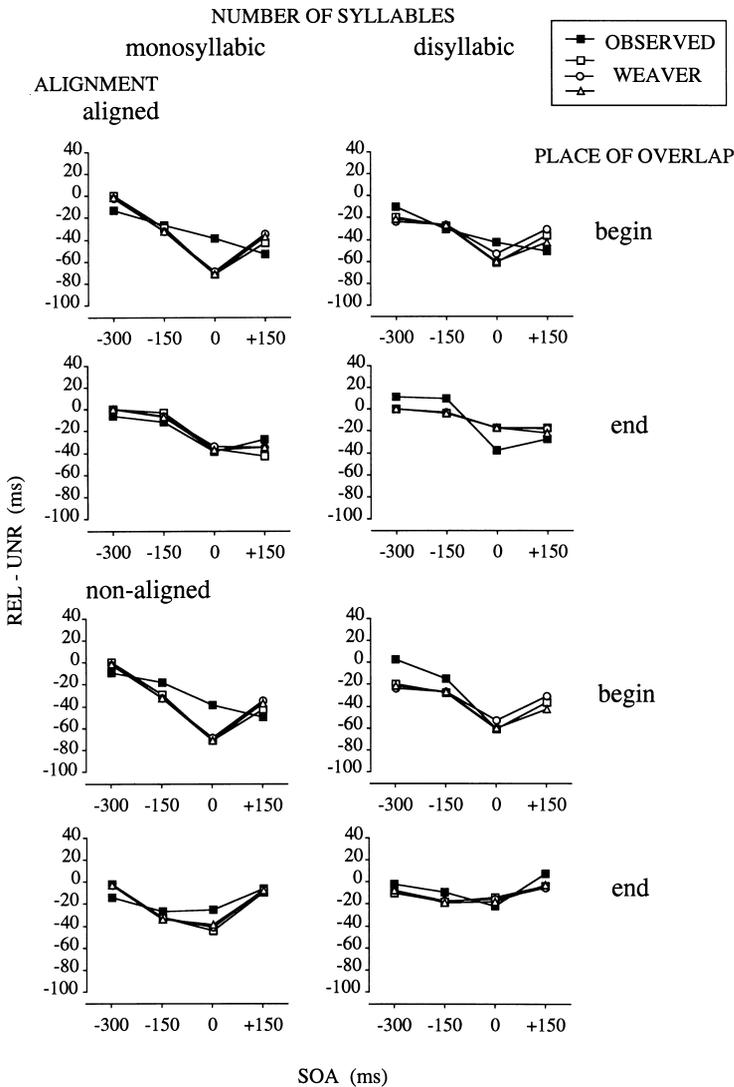

Fig. 4. The form-relatedness effects for Alignment × Number of syllables × Place of overlap × SOA in milliseconds. Real data and data simulated by WEAVER. The real data are from Experiments 2 and 3 of Meyer and Schriefers (1991). Open squares stand for the simulations with the minimal lexicon, open circles for the large highest-frequency lexicon, and open triangles for the large random lexicon.

facilitatory effect was still present at the SOA of +150, whereas with end overlap, the facilitation disappeared at SOA = +150. For the disyllables, there was a facilitatory effect at the SOAs of 0 and +150 ms in the begin condition, but there was no facilitatory effect at any of the SOAs in the end condition.

3.2. Computer simulations

I now discuss how WEAVER accounts for the patterns of effects. I first describe a simple implementation of the model for the picture-word interference paradigm. Next, I present the results of computer simulations.

The simulations were run using both small and large networks. The small network contained the words minimally needed to simulate the experiments. There were morpheme, segment, and syllable program nodes for abstract CVCs and CV.CVCs. These items represented, for example, *boek* (target), *boeg* (begin-related), *doek* (end-related), *meeuw* (name of another picture), *meer* (begin-unrelated for *boek*), *leeuw* (end-unrelated for *boek*), and all the disyllables needed. There was no segmental overlap between the forms other than that required for the simulation of the various distractor conditions.

To examine whether the size and the scope of the network influenced the outcomes, the simulations were run using larger networks. These networks contained the critical monosyllables and disyllables plus several other words. From the Dutch part of the CELEX lexical database (Baayen et al., 1995), I obtained (1) the forms of 50 nouns of highest frequency, and (2) the forms of 50 randomly selected nouns. Appendices B and C list the targets, distractors, and the other words for the larger networks. The outcomes for the small and the two large networks were the same (the Pearson product-moment correlations between the SOA curves obtained with the small and the two large simulations were, respectively, $r(48) = 0.98$ and $r(48) = 0.99$, both Ps < 0.001).

The encoding of the target word-form was simulated following the encoding algorithm described above. For the distractor word, the following extremely simplifying assumptions were made. The spoken distractor word activated a cohort of compatible elements in the output form-lexicon, analogous to what Marslen-Wilson and Welsh (1978) proposed for the input lexicon. Two levels of representation were involved: the morphological and the segmental layer (see Fig. 2). The processing of the speech signal from the beginning of the word to its end had a direct counterpart in the activation of the network. For example, for the first λ ms, the speech segment [b] of distractor [bux] activated the segment node /b/, and somewhat less the morpheme nodes <boek> and <boeg>; during the next λ ms, the [u] part activated the segment node /u/, and somewhat less the morpheme nodes <boek> and <boeg>; and during the final λ ms, the [x] part activated the segment node /x/ and somewhat less the morpheme node <boeg>, but not <boek> anymore. Current models of word recognition do not take such an all-or-none view of cohort membership. Note, however, that the cohort assumption in the simulations concerns the production lexicon, not the perception lexicon. Moreover, crucial for the account of the data is not that membership of the output cohort is all-or-none but that activation is less

for end-related than for begin-related cohort members. This assumption concerning the production lexicon corresponds to what most current models assume for the perception lexicon.

Further details of the simulations are given in Appendix A. All simulations reported in this paper were run using a single, small set of parameters whose values were held fixed. Treating some of the parameters as free ones would seem acceptable, given that, for example, for each SOA in the experiments of Meyer and Schriefers (1991) a different group of speakers was used. By allowing some parameters to vary, the fit of the model to the data could be improved. However, in holding all parameters fixed it is best illustrated that the computer model can account for the key empirical findings without free parameters doing some of the work.

3.3. Accounting for the facilitatory and inhibitory effects

Fig. 4 shows the SOA curves of the facilitatory effects as obtained by computer simulation. First, Fig. 4 shows the results for the aligned conditions. The key empirical finding here was the difference in the onset of the facilitatory effect between begin and end overlap for the monosyllables and the disyllables. With both begin and end overlap the facilitatory effect was present up to the SOA of +150. However, the onset of the effect in the begin overlap condition was at SOA = −150, whereas the onset of the effect in the end-overlap condition was at SOA = 0. The model captures this characteristic of the data. With begin overlap, the model (small-scale simulation) predicts for SOA = −150 ms a facilitatory effect of −29 ms for the monosyllables (the observed effect was −27 ms) and a facilitatory effect of −28 ms for the disyllables (real, −31 ms). In contrast, with end overlap, the effect for SOA = −150 ms was −3 ms for the monosyllables (real, −12 ms) and −4 ms for the disyllables (real, +10 ms). At the SOAs of 0 and +150, facilitation was present in both the begin and end condition. Thus, the model captures the onset difference.

Second, Fig. 4 shows the results for the non-aligned conditions. One of the key empirical findings here was the difference in the offset of the facilitatory effect between the begin and end condition for the monosyllables. Facilitation began at SOA = −150 in both the begin and the end condition. However, with begin overlap the facilitatory effect was still present at the SOA of +150, whereas with end overlap, the facilitation disappeared at SOA = +150. For the disyllables, there was a facilitatory effect at the SOAs of 0 and +150 ms in the begin condition, but there was no facilitatory effect at any of the SOAs in the end condition. Again, the model captures the basic findings, although in a somewhat exaggerated fashion. For example, the model exaggerates the difference between the monosyllables and the disyllables in the end condition, and the difference between begin and end for the disyllables. Still, the model captures the major features of the patterns of facilitation.

To assess the fit between model and real data, I computed coefficients for the correlation between the model SOA curves and the real SOA curves (shown in Figs. 4 and 5). The Pearson product-moment correlations between model and data for the small-scale and larger-scale simulations with the highest-frequency and random

lexicons were, respectively, $r(48) = 0.93$, $r(48) = 0.94$, and $r(48) = 0.95$ (all $Ps <$ 0.001). I want to emphasize that the primary aim of the simulations is to capture the presence or absence of priming effects at the different SOAs, not to account for the shape of the empirical SOA curves quantitatively. For example, in the model, the SOA curves in the monosyllabic begin condition peak at the wrong SOA. One may believe that the model should therefore be rejected. However, I have kept as many parameters constant in the simulations rather than varying them between item and speaker groups. For example, setting the correction for mental SOA to 150 ms for the monosyllables shifts the peak to a later SOA while preserving the onset effect. So, by allowing one parameter to vary, the model gets this particular aspect of the data right, and rejection of the model would be premature.

How does WEAVER account for the findings? For the purpose of exposition, I discuss the model's explanation for the monosyllables. The explanation of the effects for the disyllables goes along the same lines.

3.3.1. The basic effect of a distractor: inhibition

Consider what happens in WEAVER when the form of the target *boek* is encoded. After being activated by the lemma of *boek*, the morpheme node <boek> activates the segment nodes /b/, /u/, and /k/, which in their turn activate the syllable program node [buk]. The /b/ and /u/ also activate the competitor syllable program node [bux], and /u/ and /k/ activate [duk]. Presenting an unrelated distractor word such as *meer* will slow down the encoding of the target. The distractor *meer* activates directly the
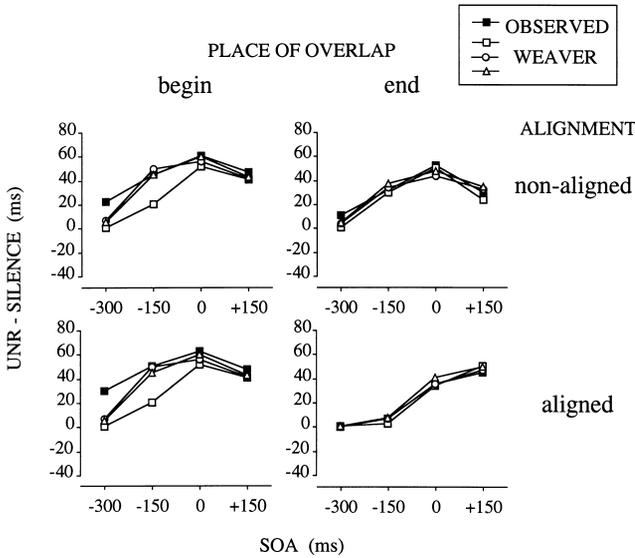

Fig. 5. Illustration of the effect of alignment on the unrelated distractors relative to the silence condition in milliseconds. Real data and data simulated by WEAVER. The real data are from the monosyllables of Experiments 2 and 3 of Meyer and Schriefers (1991). Open squares stand for the simulations with the minimal lexicon, open circles for the large highest-frequency lexicon, and open triangles for the large random lexicon.

nodes <meer>, <meeuw>, /m/, /e/, and /r/, and via these segment nodes indirectly the syllable program node [mer]. The nodes /m/ and /e/ also activate the competitor syllable program node [mew], and /e/ activates [lew]. This increases the denominator of the access ratio for the syllable program node [buk] relative to the situation without a distractor, which reduces the syllabary access probability, and thus prolongs the duration of the process of the encoding of the target.

3.3.2. The effect of relatedness: facilitation

Similar to an unrelated distractor word such as *meer*, a begin-related distractor such as *boeg* will have an inhibitory component. The distractor directly activates the morpheme node <boeg>, the segment nodes /b/, /u/, and /x/, and via these segment nodes, indirectly the syllable program node [bux]. The segment node /u/ also activates the program node [duk]. Importantly, however, the morpheme <boek> will be in the cohort of *boeg*, and the segments /b/ and /u/ activate the target syllable program node [buk]. This increases the numerator of the access ratio for [buk] relative to the non-overlap situation (i.e., *meer*, begin-unrelated), and thus speeds up the encoding of the target. This results in a facilitatory effect relative to the non-overlap situation, and may even lead to a facilitatory effect relative to the situation without a distractor (cf. Schriefers et al., 1990).

3.3.3. The effect of the place of overlap: modification of the facilitation

Consider the situation where the SOAs between distractors and pictures in both the begin and the end conditions are aligned. The later onset of the facilitatory effect from end-related distractors such as *doek* compared to begin-related ones such as *boeg* is explained as follows. According to the model, the crucial factor is that the target morpheme node <boek> is not activated by *doek*, because *boek* is not in the output cohort of *doek*. This also holds for the segmental level. Initially (i.e., during the first λ ms) *doek* primes only a competitor syllable program node (i.e., [duk] is primed by the speech signal [d] via /d/), whereas *boeg* starts by priming a competitor syllable (i.e., [b] primes [bux] via /b/) as well as the target (i.e., [b] primes [buk] via /b/). Thus, initially, *doek* acts like the unrelated distractor *leeuw*. Only later on, the target syllable program node will benefit from the form overlap through activation of segment nodes. Although the segment [u] of *doek* is presented with the same SOA as the [b] of *boeg*, the facilitatory effect of *doek* will be less as a result of the cohort factor. This surfaces as a shift of the onset of the facilitatory effect to a later moment in time (i.e., a more positive SOA). This explains the difference in the onset of the facilitation between the begin and end condition.

3.3.4. The effect of alignment: timing of facilitation

The non-alignment of the SOAs between distractors and pictures in the end condition amounts to the presentation of the distractor with a more positive SOA. For example, in case of alignment, the critical speech segment [u] of *doek* is presented at SOAs of −300, −150, 0, and +150. In contrast, in case of non-alignment, the [d] of *doek* is presented with SOAs of −300, −150, 0, and +150 ms, and thus the critical speech segment [u] will be presented with SOAs of $(-300 + \delta)$, $(-150 + \delta)$,

$(0 + \delta)$, and $(+150 + \delta)$, where δ is the duration of the [d] in the speech signal. If the duration of [d] is 150 ms, -300 becomes -150, -150 becomes 0, 0 becomes $+150$, and $+150$ becomes $+300$. Then, in case of non-alignment we obtain at SOA $= -300$ what we obtained at SOA $= -150$ in case of alignment, at SOA $= -150$ what we obtained at SOA $= 0$, and so forth. Empirically, the δ was on average 118 ms for the monosyllables, and 268 ms for the disyllables. Thus, non-alignment shifts the effect of a distractor to an earlier SOA. (This is a consequence of the fact that the distractor begins later in the non-alignment case). This explains why *doek* produced a facilitatory effect at SOA $= 0$ and SOA $= +150$ in the aligned condition and at SOA $= -150$ and SOA $= 0$ in the non-aligned condition. Furthermore, it explains the absence of any effect from the disyllables in the non-aligned experiments. The δ is greater for the disyllables than for the monosyllables. Thus, the SOA in the non-aligned case becomes even more positive (e.g., SOA $= 0$ becomes a large positive SOA, namely, empirically $+268$ ms).

3.3.5. The effect of alignment: timing of inhibition

According to the model, the alignment manipulation also affects the basic inhibitory effect from a distractor. The key finding is shown in Fig. 5. Fig. 5 shows the patterns of effect of begin-unrelated (e.g., *meer*) and end-unrelated monosyllabic distractors (e.g., *leeuw*) compared to the silence condition (the patterns for the disyllables are similar).

The shape of the SOA curve for the end-unrelated distractors in the aligned experiment differed from those in the three other conditions, which were all the same. At all SOAs except the SOA of $+150$ ms, the latencies were longer in the begin-unrelated than in the end-unrelated conditions in the aligned experiment, but not in the non-aligned one. This may come as a surprise, since both begin-unrelated and end-unrelated distractors are unrelated in form to the target. For example, *meer* (begin-unrelated) and *leeuw* (end-unrelated) do not share any part of their form in common with the target *boek*. Furthermore, in the end conditions the same words were used in all the experiments. The same held for the begin conditions. So, then, why does the pattern for the end-unrelated distractors differ between the aligned and the non-aligned experiments?

According to WEAVER, the phenomenon is a result of the alignment manipulation. Relative to non-alignment, alignment shifts the effect of a distractor to a later SOA, because it increases the time between presentation of the critical information in the distractor and the naming response. In the non-aligned experiment, the SOA for the unrelated distractors was determined by word onset in both the begin and end condition. The same held for the begin-unrelated distractors in the aligned experiment. However, the SOA in the aligned end-condition was determined by the onset of the critical part of the distractors. Thus, for *leeuw*, the SOA was between the onset of the [e] and picture onset. In the non-aligned experiment, the peaks of the inhibitory effect were at SOA $= 0$. Alignment shifts the effect of a distractor to a later SOA. Thus, this explains why the peak of the inhibition from end-unrelated distractors in the aligned experiment does not occur at SOA $= 0$ but at SOA $= +150$. Also, the onset of the inhibitory effect is shifted to a later SOA. This explains why in

the aligned experiment the inhibitory effect in the end-unrelated conditions was smaller than that in the begin-unrelated conditions, while this was not the case in the non-aligned experiment.

The model underestimates the amount of inhibition from begin-unrelated distractors at the earliest SOA in both the aligned and non-aligned experiments. However, this does not undermine the argument. Empirically, the amount of inhibition did not differ between begin and end in the non-aligned conditions, whereas it did in the aligned ones. This interaction between Place of overlap (begin, end) and Alignment (aligned, non-aligned) is attributed by the model to the alignment manipulation, not a possible main effect of Place of overlap.

For the related conditions, a similar story can be told as for the unrelated conditions. However, the account is more complex, because the effects from the related distractors involve both facilitatory and inhibitory components.

Above, I indicated that the alignment manipulation causes the facilitatory effect from end overlap to shift from the SOAs of −150 and 0 to the SOAs of 0 and +150. Now we see that this shift should be conceived of as the result of a complex interplay of factors. The alignment manipulation affects the shape of the SOA curves of the basic inhibitory effect from distractors in both the related and unrelated condition.

3.4. Accounting for binding

The key empirical finding relevant for the binding issue is that word-form encoding for production remains accurate in the context of the activation of another word form. The presentation of a distractor word does not massively increase the number of errors relative to the situation without a distractor. Furthermore, in contrast to the latency effect, the error rate in the different distractor conditions is not differentially affected by the timing of the presentation of the picture and the distractor word (Meyer and Schriefers, 1991).

Fig. 6 illustrates what is achieved by verification in WEAVER. Fig. 6 shows the activation curves for the segment nodes for /h/ and /m/ over a period of half a second. The target word in the example is *hamer* (i.e., [ˈha.mər]) and the distractor is the end-related word *zomer* (i.e., [ˈzo.mər]) presented with an SOA of +150 ms. The example is from the simulation of the aligned experiment, thus the SOA is between the onset of the [mər] of *zomer* in the speech signal and picture onset. The activation level of the /h/ node gradually increases. In contrast, there is a rapid increase in the activation level of /m/ as a result of perceiving the [m] in the distractor. The activation of the /h/ node does not exceed that of the /m/ node at any moment in time. In the encoding of *hamer*, however, first the /h/ has to be selected. Consequently, the selection of /h/ cannot be based on the node's level of activation. WEAVER correctly produces the appropriate segments. The activation of the /h/ and /m/ nodes triggers a production that selects the nodes if they are linked to the target morpheme <hamer>. This is the case. Thus, both nodes are made available to the syllabification process. The /h/ is syllabified before the /m/, because the /h/ is represented as the first segment of <hamer> and /m/ as its third.
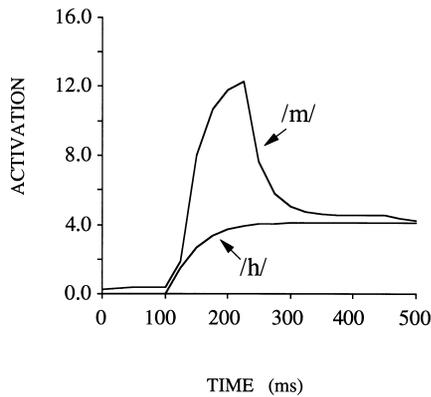
Fig. 6. Activation curves for the /h/ and /m/ nodes in WEAVER during the encoding of *hamer*. Depicted is the aligned condition with the end-related distractor word *zomer* presented at SOA = +150 ms.

4. Application to implicit priming

The patterns of inhibition and facilitation obtained in picture-word interference experiments differ in an interesting way from those obtained with another important chronometric paradigm, the so-called implicit priming paradigm. In a picture-word interference experiment, both first-syllable (*ha*) and second-syllable (*mer*) spoken primes yield facilitation in producing a disyllabic word such as *hamer* (Meyer and Schriefers, 1991). Only the onset of the facilitation differs between the types of prime. By contrast, in implicit priming experiments, the production of a disyllabic word like *hamer* is speeded up by advance knowledge about the first syllable (*ha*) but not by advance knowledge about the second syllable (*mer*), as shown by Meyer (Meyer 1990, 1991). Below, I will show that WEAVER resolves this discrepancy (cf. Roelofs, 1994).

4.1. Basic empirical findings

Meyer (Meyer 1990, 1991) examined the planning of word forms using a paradigm that she coined the implicit priming paradigm. This paradigm involves producing words from learned paired-associates. The big advantage of this paradigm compared to the more widely used picture-word interference paradigm is that the responses do not need to be names of depictable entities, which gives more freedom in the selection of materials. In Meyer's experiments, speakers first learned small sets of prompt-response word pairs such as {*spijker – hamer*, *vogel – havik*, etc.} ({*nail – hammer*, *bird – hawk*, etc.}). After learning a set of pairs, the speaker was presented with one of the prompts at random on each trial. The prompts were visually presented on a computer screen. The task for a speaker was to produce the second word of a pair (e.g., *hamer*) upon the visual presentation of the first word (*spijker*). The instruction was to respond as quickly as possible without making mistakes. The production latency (i.e., the interval between prompt onset and speech

onset) was the main dependent variable. An experiment comprised homogeneous and heterogeneous response sets. In a homogeneous set, the response words shared part of their form and in a heterogeneous set they did not. For example, the responses shared the first syllable (*HAmer*, *HAvik*, etc.) or the second syllable (*haMER*, *zoMER*, etc.) or they were unrelated (*havik*, *zomer*, etc.). The same prompt-response pairs were tested in the homogeneous and heterogeneous condition; only their combinations into sets differed. Each speaker was tested on all sets. Meyer found shorter production latencies for the words in homogeneous sets than in heterogeneous sets. Importantly, facilitation was obtained only when the overlap was from the beginning of the response words onward. Thus, a facilitatory effect was obtained for the set that included *HAmer* and *HAvik* but not for the set that included *haMER* and *zoMER*. Furthermore, facilitation increased with the number of shared segments.

According to WEAVER, this seriality phenomenon reflects the suspend-resume mechanism that underlies the incremental planning of an utterance. Assume the response set consists of *hamer*, *havik*, and so forth (i.e., the first syllable is shared). Before the beginning of a trial, the morphological encoder can do nothing, the phonological encoder can construct the first phonological syllable $(ha)_\sigma$, and the phonetic encoder can recover the first motor program [ha]. When the prompt *spijker* is given, the morphological encoder will retrieve <hamer>. Segmental spellout makes available the segments of this morpheme, which includes the segments of the second syllable. The phonological and phonetic encoders can start working on the second syllable. In the heterogeneous condition (*havik*, *zomer*, etc.), nothing can be prepared. There will be no morphological encoding, no phonological encoding, and no phonetic encoding. In the end-homogeneous condition (*hamer*, *zomer*, etc.), nothing can be done either. Although the segments of the second syllable are known, the phonological word cannot be computed because the remaining segments are to the left of the suspension point (the point at which the encoding stops and waits for new input). In WEAVER, this means that the syllabification process has to go to the initial segments of the word, which amounts to restarting the whole process, like unraveling a woven fabric. Thus, facilitation will be obtained for the homogeneous sets relative to the heterogeneous sets for the begin condition only.

4.2. Computer simulations

Computer simulations of the experiments of Meyer (1990) supported the above theoretical analysis (see Fig. 7). In the simulations, the begin-related response set consisted of *hamer* [ˈha.mər], *havik* [ˈha.vɪk], and so forth, the end-related set consisted of *hamer* [ˈha.mər], *zomer* [ˈzo.mər], and so forth, and the unrelated set was formed by *havik* [ˈha.vɪk], *zomer* [ˈzo.mər], and so forth. Advance knowledge about a syllable was simulated by completing the phonological and phonetic encoding of the syllable before the beginning of a trial. For the begin condition, the model yielded a facilitatory effect of −43 ms (real, −49 ms, collapsed across trochaic feet and iambs), whereas for the end condition it predicted an effect of 0 ms (real, +5 ms). Thus, WEAVER captures the empirical phenomenon.
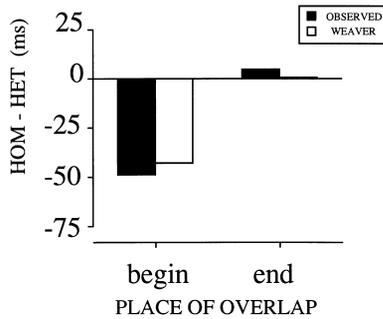
Fig. 7. Mean difference in milliseconds between the production latencies in the homogeneous and the heterogeneous condition for Place of overlap (begin, end). Real data and data simulated by WEAVER. The real data are from Meyer (1990).

5. Speech errors

In this section, I explain WEAVER's approach to errors in word-form encoding, along lines similar to proposals by Crompton (1981) and Levelt (1989). In particular, I address segment exchanges, anticipations, and perseverations, which are the most frequently occurring form errors in spontaneous speech (making up 70–90% of the sound errors; for overviews, see e.g., Cutler, 1982; Dell, 1986; Garrett, 1975; Levelt, 1989; Shattuck-Hufnagel, 1979; for Dutch, see Nooteboom, 1969).

The discussion below is only meant to illustrate WEAVER's approach to speech errors, not to account for the rich patterns found in speech error corpora. Nevertheless, it should be clear how WEAVER can, at least in principle, account for speech errors. In WEAVER, speech errors occur when verification fails. In earlier models, errors occur when, as a result of noise in the system, the target node is not the most highly activated one and another node gets erroneously selected instead (e.g., Dell, 1986; Shattuck-Hufnagel, 1979). In these models, segmental errors are explained as segment selection failures in the construction of a phonological representation. Segments and slots of frames are labeled for syllable position (e.g., /onset h/), which explains why in errors syllable onsets interact with onsets but not with codas, and vice versa. In case of on-line syllabification, selected segments themselves are not marked for syllable position. So, given the position constraint, speech errors cannot occur during the construction of a phonological representation. Below, WEAVER demonstrates that segmental errors and the position constraint may be a result of indexing failures of the device that maps syllabified phonological representations onto a mental syllabary (cf. Crompton, 1981; Levelt, 1989).

5.1. Segmental errors as failures of syllabary access

Consider the encoding of the Dutch phrase 'goed weer' [xut wer] ('good weather'). Assume phonological encoding has been completed for *goed* and *weer*. Thus, the segments /x/, /u/, and /t/ have been selected and syllabified for *goed*, and /w/, /e/, and /r/ for *weer*. Assume phonetic encoding starts for both items.

The segments /x/, /u/, and /t/ will activate the syllable node [xut], but also [wut], and so forth. Simultaneously, the segments /w/, /e/, and /r/ will activate the syllable node [wer], but also [xer], and so forth. Note that the /x/ of *goed* is not only the onset of the syllable [xut], but also the onset of a competitor of the syllable of *weer*, [xer]. The same holds for the onset of *weer*, /w/, which will activate the competitor [wut] of [xut]. On occasion, the phonetic encoder might make the following indexing error. In testing for the relation between [wut] and the selected, syllabified segments, the production rule of [wut] might detect that its /w/ indeed occupies an onset position, overlooking that it is in the wrong syllable within the phonological representation. However, when the production rule of [wut] takes its condition to be satisfied, and [xut] does its job well, then there will be a race between [xut] and [wut] for the first position within the articulatory program. When [wut] wins the race, the speaker will make an anticipation, 'woet weer', instead of 'goed weer'. This may precipitate the same mistake for the second syllable, so that an exchange occurs ('woed geer' instead of 'goed weer'). If the selection error only occurs for the second syllable, a perseveration will be made ('goed geer' instead of 'goed weer').

5.2. Similarity effects

In WEAVER, errors are a result of indexing failures during the scanning of a phonological representation by the phonetic encoder. Thus, error elements will occupy similar positions within phonological structures, as empirically observed. Slips of the tongue involving segments typically respect relationships between three different levels of phonological structure, namely segments, syllables, and stress patterns (e.g., Dell, 1986; Levelt, 1989; Shattuck-Hufnagel, 1992). For example, the segments occupy the same syllable positions (or word position, see Shattuck-Hufnagel, 1992), they tend to come from syllables with the same level of stress, they have similar segmental neighbors, and so forth.

Another general finding concerns feature similarity effects. In substitution errors, the target and the replacing segment usually share most of their features. In WEAVER, features are part of the phonetic syllabary. Both segment nodes and syllable program nodes point to their features (i.e., articulatory gestures). Indexing failures may be more likely when segments and syllables point to the same features.

6. Overview of empirical tests

After its development, WEAVER has been critically tested on new sets of data. Verification times in WEAVER vary as a function of the frequency of an item due to production rule strengthening (cf. Anderson, 1983), so that different items take different periods of encoding time within the network. Applications of WEAVER to findings concerning word, morpheme, and syllable frequency, as well as new experimental tests can be found in Roelofs (1996b, 1997c). WEAVER deals with the flexibility of syllable membership by computing syllable positions on-line rather than storing them with lexical items. Applications of WEAVER to findings on cross-

morpheme syllabification, cross-word syllabification, and cliticization, as well as new experimental tests can be found in Roelofs (1997a). Below, I review the outcomes of empirical tests concerning the procedural-declarative distinction, rightward incrementality, morphological decomposition, and stored metrical structure.

6.1. Procedural-declarative distinction

The production of a disyllabic word is speeded up by advance knowledge about the first syllable but not by knowledge about the second syllable (Meyer, 1990). In contrast, when first-syllable or second-syllable primes are auditorily presented, both primes yield facilitation (Meyer and Schriefers, 1991). According to WEAVER, both first-syllable and second-syllable spoken primes yield facilitation, because they will spread activation to segments of the target in declarative memory and therefore speed up retrieval. Implicit priming reflects production application in syllabification. Later segments cannot be syllabified before earlier ones. In this view, implicit priming reflects production application and explicit priming the spread of activation.

New experiments tested for independent effects of implicit and explicit priming (Roelofs, submitted), thereby testing the procedural-declarative distinction. In the experiments, there were homogeneous and heterogeneous response sets (the implicit primes) as well as related and unrelated spoken distractors (the explicit primes). Speakers had to produce single words such as *hamer*, simple imperative sentences such as 'zoek op!' ('look up!'), or imperative sentences with cliticizations such as 'zoek's op!' ('look up now!'). In homogeneous sets, the responses shared the first syllable, (e.g., *ha* in *hamer*), the verb (e.g., *zoek* (*look*) in 'zoek op!'), or the verb plus clitic (e.g., *zoek 's* in 'zoek's op!'), and in the heterogeneous condition there was no overlap. The spoken distractors consisted of the final syllables of the utterance, either a target syllable (e.g., *mer* for *hamer* or *op* for 'zoek op!', the related condition), a syllable of another word or sentence in the response set (the unrelated condition), or there was no distractor (the silence condition). The homogeneity variable (Context) and the distractor variable (Distractor) yielded main effects and the effects were additive. Furthermore, the effects were the same for the production of single words and simple imperative sentences, both with and without cliticizations.

Computer simulations demonstrated that WEAVER accounts for the additive effects of Context and Distractor. Fig. 8 shows the outcomes of simulations concerning the production of simple sentential forms such as 'zoek op!'.

6.2. Rightward incrementality and morphological decomposition

WEAVER provides for a suspend-resume mechanism that supports incremental generation of phonetic plans. The encoding algorithm requires morphologically decomposed form entries. In languages such as Dutch, morphemes such as prefixes, particles, and some suffixes constitute independent domains of syllabification (Booij, 1983, 1995). In Dutch, this holds for prefixes such as *ver-* and *ont-*, particles
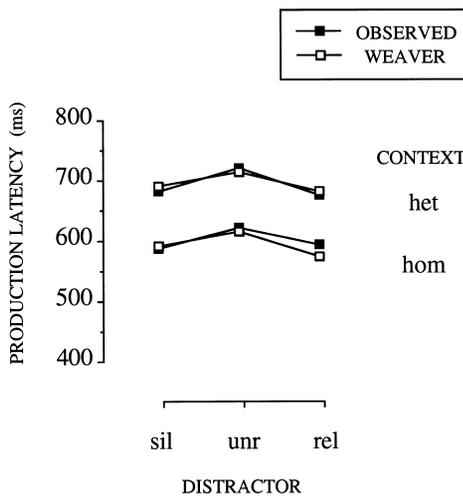
Fig. 8. Mean production latencies in milliseconds of the imperative sentences for Context × Distractor. Real data and data simulated by WEAVER. The absolute production latencies for WEAVER were obtained by adding a constant of 300 ms to the word-form encoding latencies in the simulations. The real data are from Roelofs (submitted).

such as *op*, *af*, *aan*, *uit*, and so forth, and suffixes such as *-achtig*, but not for suffixes such as *-in*, *-er*, and *-ing*. For example, the segment /r/ of the prefix *ver-* of the Dutch verb *vereren* (*honour*) is not syllabified with the base verb *eren*, as the maximal onset principle in syllabification would predict, but is syllabified as the coda of *ver-*. This does not hold for a morphologically simple verb such as *verifiëren* (*verify*), where the /r/ is the onset of the second syllable *ri* instead of the coda of *ver*.

New implicit priming experiments tested for rightward incrementality and morphological decomposition (Roelofs, 1996a,b, 1997c). WEAVER predicts that a larger facilitatory effect should be obtained when shared initial segments constitute a morpheme than when they do not. For example, the effect should be larger for sharing the syllable *bij* in response sets including Dutch compounds such as *bijrol* (morphemes <bij> and <rol>, *supporting role*) than for sharing the syllable *bij* in sets including simple words such as *bijbel* (morpheme <bijbel>, *bible*). For monomorphemic words like *bijbel* consisting of the morpheme <bijbel>, sharing the first syllable *bij* allows phonological preparation only. In contrast, for polymorphemic words like *bijrol* consisting of the morphemes <bij> and <rol>, additional morphological preparation is possible.

The outcomes of the experiments confirmed the predictions by WEAVER. In producing disyllabic simple and compound nouns, more facilitation was obtained when a shared initial syllable constituted a morpheme than when it did not. For example, the preparation effect was larger for *bij* in *bijrol* (<bij> and <rol>) than for *bij* in *bijbel* (<bijbel>).

Further experiments supported WEAVER's claim that successive morphemes are planned in a rightward fashion. In producing nominal compounds, no facilitation was obtained for shared noninitial morphemes. For example, no effect was obtained
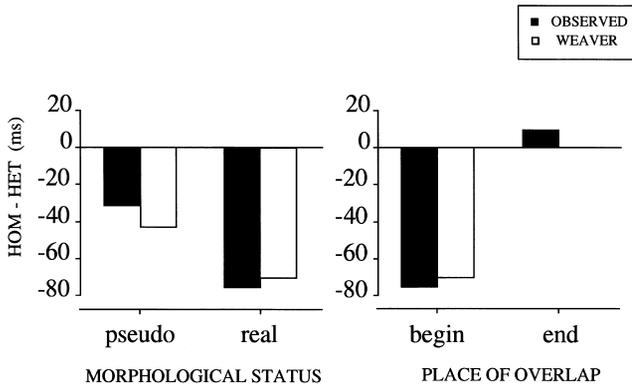
Fig. 9. Mean difference in milliseconds between the production latencies in the homogeneous and the heterogeneous condition for Morphological status (pseudo, real) and Place of overlap (begin, end). Real data and data simulated by WEAVER. The real data are from Roelofs (1996a).

for <rol> in *bijrol*. In producing prefixed verbs, a facilitatory effect was obtained for the prefix but not for the noninitial base. The same held for particle verbs.

Fig. 9 presents the results of simulations of these experiments (Roelofs, 1996a). The left-hand graph shows the effect of morphological status. Sharing *bij* in *bijrol* and *bijbel* yields a facilitatory effect. The facilitatory effect for *bij* in *bijrol* (consisting of the morphemes <bij> and <rol>, the real condition) is larger than the facilitatory effect for *bij* in *bijbel* (<bijbel>, the pseudo condition). This corresponds to the empirical findings. The right-hand graph shows the effect of linear order. A facilitatory effect is obtained when the responses share the first morpheme (*bijrol*, *bijnier*, the begin condition), but not when they share the second morpheme (*bijrol*, *koprol*, the end condition). This corresponds to the experimental findings.

6.3. Metrical structure

WEAVER assigns a specific role to metrical structures in syllabification. WEAVER's metrical structures specify the stress pattern across syllables but not the precise CV structure of the syllables as the CV headers of Dell (1988) do. The syllabification process in WEAVER associates retrieved segments to the syllable nodes within retrieved metrical structures. Roelofs and Meyer (Roelofs and Meyer, 1997; Roelofs, 1997a) conducted a number of implicit priming experiments designed to test WEAVER's view on metrical structure. On each trial, speakers had to produce one Dutch word out of a set of three, or four, as quickly as possible. In homogeneous sets, the responses shared a number of word-initial segments, whereas in heterogeneous sets they did not. Earlier research has shown that sharing initial segments reduces production latencies (Meyer, 1990, 1991; Roelofs, 1996a,b). The responses shared their metrical structure (the constant sets) or they did not (the variable sets).

WEAVER's view of syllabification implies that preparation for word-initial segments should only be possible if response words have an identical metrical structure.

If the responses in a set have different metrical structures, segment-to-frame association cannot take place before the beginning of a trial. This prediction was tested by comparing the effect of segmental overlap for response sets with a constant number of syllables such as {*manier* (*manner*), *matras* (*mattress*), *makreel* (*mackerel*)} (all 2 syllables) to that for sets having a variable number of syllables such as {*majoor* (*major*), *materie* (*matter*), *malaria* (*malaria*)} (respectively, 2, 3, and 4 syllables). In the example, the responses share the first syllable *ma*. Word stress was always on the second syllable. As predicted, facilitation was obtained for the metrically constant sets but not for the variables sets. The same predictions were also tested by comparing the effect of segmental overlap for response sets with a constant stress pattern such as {*marine* (*navy*), *materie* (*matter*), *malaise* (*depression*), *madonna* (*madonna*)} (all responses having stress on the second syllable) to that for sets having a variable stress pattern such as {*marine* (*navy*), *materie* (*matter*), *manuscript* (*manuscript*), *madelief* (*daisy*)} (first two responses having stress on the second syllable and the last two responses having stress on the third syllable). All response words were trisyllabic. Again, as predicted, facilitation was obtained for the constant but not for the variable sets. Thus, speakers can only prepare for a shared initial syllable when they also know the number of syllables that follow and which of these syllables carries the word accent.

It was also tested whether the preparation effect is affected by the constancy versus variability of the CV structure of the response words. This was done by comparing the effect of segmental overlap for response sets having a constant CV structure such as {*bres* (*breach*), *bril* (*glasses*), *brok* (*piece*), *brug* (*bridge*)} (responses all CCVC) to that for sets having a variable CV structure such as {*brij* (*porridge*), *brief* (*letter*), *bron* (*source*), *brand* (*fire*)} (responses respectively, CCVV, CCVVC, CCVC, CCVCC). In the example, the responses share the onset cluster *br*. Facilitation from segmental overlap was obtained for both the constant and the variable sets. The size of the effect was the same for both types of set.

WEAVER explains why preparation for word-initial segments is only possible for response words with an identical number of syllables and stress pattern, and why identical CV is not needed. Fig. 10 gives the results of simulations comparing the
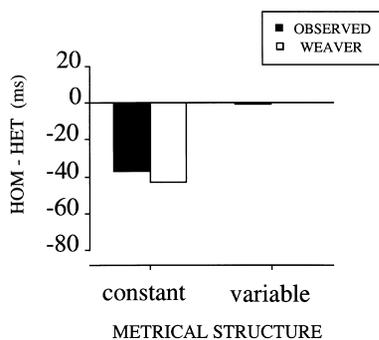

Fig. 10. Mean difference in milliseconds between the production latencies in the homogeneous and the heterogeneous condition for Metrical structure (constant, variable). Real data and data simulated by WEAVER. The real data are from Roelofs and Meyer (1997).

effect of segmental overlap for response sets with a constant number of syllables such as {*manier*, *matras*, *makreel*} to that for sets having a variable number of syllables such as {*majoor*, *materie*, *malaria*}. Varying the place of stress while keeping the number of syllables fixed gives the same results. As can be seen, WEAVER accounts for the key empirical finding. In contrast, if metrical structures are not involved in advance planning or if metrical structures are computed on-line on the basis of segments for these words, sharing metrical structure should not be necessary for preparation.

7. General discussion

WEAVER is well grounded within cognitive science: a spreading-activation based declarative network integrated with a parallel object-oriented production system. The type of system is a mixture of traditional artificial intelligence, connectionism, and traditional cognitive modeling (cf. Anderson, 1983).

WEAVER provides a comprehensive account of form planning using spreading activation and verification. The latter is not such a powerful mechanism as the term perhaps might suggest. Verification/labeling in WEAVER simply means that there is a production rule that fires if its condition and action nodes are active and appropriately connected. There are many production-system models in the literature (e.g., Anderson, 1983). Furthermore, at several places it has been shown how production systems might be realized connectionistically (e.g., Touretzky and Hinton, 1988) and in the brain (e.g., Shastri and Ajjanagadde, 1993). In addition, existing comprehensive models of speech production such as those of Dell (1986, 1988), Levelt (1989), and Shattuck-Hufnagel (1979) have similar declarative and procedural aspects. In short, there is sufficient empirical and computational ground for a model that proposes a simple mechanism, a production rule, to achieve two complex tasks: to account for latencies and binding in multiple-input situations and to account for the flexibility of syllable membership as evident from cross-morpheme and cross-word syllabification. Any model needs computational primitives that get the job done. For none of the existing models it has been demonstrated that the same tasks can be achieved with similar or simpler means.

Production rules in WEAVER account for the computational problem of selection. To account for latency effects, an access ratio was proposed, which is one of the simplest competition-sensitive response time mechanisms possible. Any model of speech production that is applied to picture-word interference seems to require a competition sensitive latency mechanism. Some models propose inhibitory links, other models such as WEAVER propose a ratio, and some models propose both (see e.g., Dell and O'Seaghdha, 1994 for discussion).

As concerns speech errors, any model needs some form of randomness, either in the spread of activation (e.g., Dell) or in production application (WEAVER). Assuming randomness in the spread of activation is not necessarily better than assuming stochastic aspects in production application. Rather, we have two alternatives here, neither of which has received more empirical support.

To conclude, WEAVER is an attempt to address the challenge posed by the morphophonological and phonetic encoding of word forms in context and by empirical findings about encoding times. A major challenge for future research is to see if the rich patterns in error corpora can be accounted for by further developing WEAVER and to see if existing error-based models can be further developed such that they can account for the rich patterns obtained in chronometric experiments.

Acknowledgements

Appendix A

Below, details of the computer simulations are given. The simulations involved word-form encoding up to the access of the phonetic syllabary. The mathematical equations for the spreading of activation, the access ratio, and the expectation of the word-form encoding latency are as follows (cf. Roelofs, 1992a, 1993). Activation spreads according to:

$$a(k, t+\Delta t) = a(k, t)(1-d) + \sum_n r\, a(n, t)$$

where $a(k,t)$ is the activation level of node k at point in time t, d is a decay rate ($0 < d < 1$), and Δt is the duration of a time step (in ms). The rightmost term denotes the amount of activation that k receives between t and $t + \Delta t$, where $a(n,t)$ is the output of neighbor n (equal to its level of activation). The factor r indicates the spreading rate.

The probability that syllabary access for a target program m will be initiated at $t < T \leq t + \Delta t$ given that it has not been initiated at $T \leq t$, and provided that the access condition for a program is met, is given by the ratio:

$$P(access\ m\ at\ t < T \leq t+\Delta t \mid \neg\ access\ m\ at\ T \leq t) = \frac{a(m, t)}{\sum_i a(i, t)}$$

The index i ranges over the syllable program nodes in the word-form network of a speaker. The access ratio equals the hazard rate $h_m(s)$ of the process of the encod-

ing of syllable m up to the access of the syllabary at time step s (cf. Luce, 1986), where $t = (s - 1)\Delta t$, and $s = 1, 2, \ldots$ The hazard rate gives the probability that the encoding is complete at a particular moment in time given that the encoding was not complete yet. The expected latencies of word-form encoding up to the access of the syllabary, $E(T)$, for the monosyllables and the disyllables are as follows. Let

$$f_m(s) = P(\text{access syllable } m \text{ at } s)$$

$$h_m(s) = P(\text{access syllable } m \text{ at } s \mid \neg \exists u : (u < s \land \text{access syllable } m \text{ at } u))$$

$$V_m(s) = P(\neg \exists u : (u \leq s \land \text{access syllable } m \text{ at } u))$$

These are, respectively, the probability mass, hazard rate, and cumulative survivor function of the encoding of syllable m. For the latter holds:

$$V_m(s) = \prod_{j=0}^{s} [1 - h_m(j)].$$

It can be shown (see Roelofs, 1992a, Roelofs, 1992b) that

$$f_m(s) = h_m(s) \left\{ \prod_{j=0}^{s-1} [1 - h_m(j)] \right\}.$$

In the encoding of a monosyllabic utterance, there is a single target syllable program m. The probability $P(\text{word-form encoding completes at } s)$, $f(s)$, equals $P(\text{access syllable } m \text{ at } s)$, $f_m(s)$. Then, the expected word-form encoding latency for monosyllabic words equals

$$E_{mono}(T) = \sum_{s=1}^{\infty} f_m(s) \, s \, \Delta t$$

$$= \sum_{s=1}^{\infty} h_m(s) \left\{ \prod_{j=0}^{s-1} [1 - h_m(j)] \right\} s \, \Delta t.$$

In the encoding of a disyllabic word, there are two target syllable programs, syllable 1 and syllable 2. The expected word-form encoding latency of a disyllable can be derived as follows. The probability $P(\text{word-form encoding completes at } s)$, $f(s)$, for a disyllabic word equals

$$(P(\text{access syllable } 1 \text{ at } s) \land P(\exists u : (u < s \land \text{access syllable } 2 \text{ at } u)) \lor$$

$$\left(P(\text{access syllable } 2 \text{ at } s) \land P(\exists u : (u < s \land \text{access syllable } 1 \text{ at } u)) \lor \right.$$

$$\left(P(\text{access syllable } 1 \text{ at } s) \land P(\text{access syllable } 2 \text{ at } s)) \right.$$

$$= f_1(s) \sum_{j=0}^{s-1} f_2(s) + f_2(s) \sum_{j=0}^{s-1} f_1(s) + f_1(s)f_2(s)$$

$$= [h_1(s)V_1(s-1)] \sum_{j=0}^{s-1} [h_2(j)V_2(j-1)] +$$

$$\left[h_2(s)V_2(s-1)\right] \sum_{j=0}^{s-1}\left[h_1(j)V_1(j-1)\right] +$$

$$\left[h_1(s)V_1(s-1)\right]\left[h_2(s)V_2(s-1)\right],$$

where $h_1(s)$ and $h_2(s)$ are the hazard rates of the encoding of the first syllable and second syllable, respectively, and $V_1(s)$ and $V_2(s)$ are the corresponding cumulative survivor functions. For the expectation of T holds

$$E_{di}(T) = \sum_{s=1}^{\infty}\left\{f_1(s)\sum_{j=0}^{s-1}f_2(s) + f_2(s)\sum_{j=0}^{s-1}f_1(s) + f_1(s)f_2(s)\right\}s\,\Delta t.$$

The parameter estimates were as follows. The spreading rate r within the word-form stratum was 0.0120 [ms^{-1}], the decay rate d was 0.0240 [ms^{-1}], and the size of the external input to the network *extin* was 0.1965 [ms^{-1}]. The latter held for the activation of a morpheme node from a lemma, and for the activation of a segment node from the perception of the corresponding speech segment in the spoken distractor. Morpheme nodes, however, got only 10% of the external input *extin* from the perception of a speech segment. This captures the idea that a speech segment (e.g., [b]) provides much more evidence for the segment (i.e., /b/) than for a morpheme that comprises the segment (e.g., <boek>). A segment perceived in the monosyllabic and disyllabic spoken distractors provided input to the network for 125 and 100 ms, respectively. The activation threshold for the triggering of a production rule was 1.5. The duration of basic events such as the time for the activation to cross a link, the latency of a selection test, and the syllabification time per syllable equaled $\Delta t = 25$ ms. The correction for the deviation of the mental SOA from the experimental one was +100 ms.

The parameter estimates were obtained by fitting the model to the data of Meyer and Schriefers (1991). This may look odd when one of the aims is to give a detailed account of the results of Meyer and Schriefers. When parameters are estimated from the same data used to test a model, the estimated parameters may allow the model to fit the data better. However, there is no serious problem with this standard procedure (for an extensive discussion, see e.g., Wickens, 1982). First, the number of data points here (48) is much larger than the number of estimated parameter values (only 6; the values of d, Δt and *extin* were held the same as in the lemma retrieval model of Roelofs, 1992a and Roelofs, 1993). Second, the parameters have been held fixed in all other applications that tested the model in this paper and elsewhere. Third, a statistic for the fit may be adjusted such that the fit of the model to the data is required to be better with than without estimated parameters. For each additional parameter that is estimated, the degrees of freedom of the statistic may be reduced by one, which yields a more stringent test. Applying this to the present fit, the correlation between model and the data of Meyer and Schriefers remains highly significant. The statistics for the correlation between data and model for the small-scale and larger-scale simulations with the highest-frequency and random lexicons become, respectively, $r(42) = 0.93$, $r(42) = 0.94$, and $r(42) = 0.95$ (all Ps < 0.001).

WEAVER was programmed from scratch in the C programming language. The simulations were run on a Hewlett-Packard workstation under the UNIX operating system.

Appendix B

The targets, distractors, and 50 highest-frequency lexical items from CELEX used in one of the simulations that examined the influence of the size and scope of the form network on the simulation outcomes.

---

boek [buk] (target)
boeg [bux] (REL, BEGIN distractor)
doek [duk] (REL, END distractor)
meeuw [mew] (other target)
meer [mer] (UNR, BEGIN distractor)
leeuw [lew] (UNR, END distractor)
hamer ['ha.mər] (target)
havik ['ha.vɪk] (REL, BEGIN distractor)
zomer ['zo.mər] (REL, END distractor)
python ['pi.tɔn] (other target)
pias ['pi.jɑs] (UNR, BEGIN distractor)
foton ['fo.tɔn] (UNR, END distractor)

tijd [tɛit]	huis [hʌys]
water ['wa.tər]	naam [nam]
groep [xrup]	man [mɑn]
hoofd [hoft]	vraag [vrax]
bed [bɛt]	stad [stɑt]
jaar [jar]	werk [wɛrk]
manier [ma.'nir]	woord [wort]
vorm [vɔrm]	school [sxol]
paar [par]	deur [dØr]
geld [xɛlt]	kant [kɑnt]
hand [hɑnt]	kind [kɪnt]
grond [xrɔnt]	staat [stat]
dood [dot]	dag [dɑx]
wereld ['we.rəlt]	kamer ['ka.mər]
moment [mo.'mɛnt]	mond [mɔnt]
vrouw [vrɑuw]	mens [mɛns]
zin [zɪn]	licht [lɪxt]
recht [rɛxt]	plaats [plats]
keer [ker]	deel [del]
wijze ['wɛi.zə]	zaak [zak]
moeder ['mu.dər]	weg [wɛx]
soort [sort]	lichaam ['lɪ.xam]
meneer [mə.'ner]	vader ['va.dər]
uur [yr]	land [lɑnt]
stem [stɛm]	feit [fɛit]

Appendix C

The targets, distractors, and 50 randomly selected lexical items from CELEX used in one of the simulations that examined the influence of the size and scope of the form network on the simulation outcomes.

---

boek [buk] (target)
boeg [bux] (REL, BEGIN distractor)
doek [duk] (REL, END distractor)
meeuw [mew] (other target)
meer [mer] (UNR, BEGIN distractor)
leeuw [lew] (UNR, END distractor)
hamer [ˈha.mər] (target)
havik [ˈha.vɪk] (REL, BEGIN distractor)
zomer [ˈzo.mər] (REL, END distractor)
python [ˈpi.tɔn] (other target)
pias [ˈpi.jɑs] (UNR, BEGIN distractor)
foton [ˈfo.tɔn] (UNR, END distractor)

schnapps [sjnɑps]
piraat [pi.ˈrat]
glitter [ˈxlɪ.tər]
niveau [ni.ˈvo]
spruit [sprʌyt]
schoft [sxɔft]
vuist [vʌyst]
input [ˈɪm.put]
spek [spɛk]
regressie [rə.ˈxrɛ.si]
teut [tØt]
cohesie [ko.ˈhe.zi]
panda [ˈpɑn.da]
kraal [kral]
maniak [ma.ni.ˈjɑk]
lantaren [lɑn.ˈta.rə]
marihuana [ma.ri.ju.ˈwa.na]
bankroet [bɑŋk.ˈrut]
monoloog [mo.no.ˈlox]
artillerie [ɑr.ti.lə.ˈri]
ingenieur [ɪn.zjən.ˈjØr]
schild [sxɪlt]
kampioen [kɑm.pi.ˈjun]
rente [ˈrɛn.tə]
piccolo [ˈpi.ko.lo]

syncope [sɪŋ.ko.ˈpe]
extase [ɛk.ˈsta.zə]
ampex [ˈɑm.pɛks]
symfonie [sɪm.fo.ˈni]
expansie [ɛks.ˈpɑn.zi]
kop [kɔp]
consult [kɔn.ˈzœlt]
faun [fɑun]
gymnast [xɪm.ˈnɑst]
modem [ˈmo.dəm]
roemer [ˈru.mər]
convent [ˈkɔn.vɛnt]
stam [stɑm]
neglige [ˈne.gli.zje]
worp [wɔrp]
robot [ˈro.bɔt]
baret [ba.ˈrɛt]
glas [xlɑs]
pik [pɪk]
travestie [tra.vɛs.ˈti]
crux [krœks]
wiskunde [ˈwɪs.kœn.də]
schaar [sxar]
salmiak [ˈsɑl.mi.jɑk]
distinctie [dis.ˈtɪŋk.si]

References

Anderson, J.R. 1983. The Architecture of Cognition. Harvard University Press, Cambridge, MA.

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

Béland, R., Caplan, D., Nespoulous, J.-L., 1990. The role of abstract phonological representations in word production: evidence from phonemic paraphasias. Journal of Neurolinguistics, 5, 125–164.

Bobrow, D.G., Winograd, T., 1977. An overview of KRL, a knowledge representation language. Cognitive Science, 1, 3–46.

Booij, G.E., 1983. Principles and parameters in prosodic phonology. Linguistics, 21, 249–280.

Booij, G.E. 1995. The Phonology of Dutch. Clarendon Press, Oxford.

Browman, C.P., Goldstein, L., 1986. Towards an articulatory phonology. Phonology Yearbook, 3, 219–252.

Butterworth, B. 1989. Lexical access in speech production. In: Marslen-Wilson, W. (Ed.), Lexical Representation and Process. MIT Press, Cambridge, MA.

Caplan, D. 1992. Language: Structure, Processing, and Disorders. MIT Press, Cambridge, MA.

Chomsky, N., Halle, M. 1968. The Sound Pattern of English. Harper and Row, New York.

Collins, A.M., Loftus, E.F., 1975. A spreading-activation theory of semantic processing. Psychological Review, 82, 407–428.

Crompton, A., 1981. Syllables and segments in speech production. Linguistics, 19, 663–716.

Cutler, A. 1982. Slips of the Tongue and Language Production. Mouton, Berlin.

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., O'Seaghdha, P.G. 1994. Inhibition in interactive activation models of linguistic selection and sequencing. In: Dagenbach, D., Carr, T.H. (Eds.), Inhibitory Processes in Attention, Memory, and Language. Academic Press, San Diego, CA, pp. 409–453.

Donders, F.C. 1868. Die Schnelligkeit psychischer prozesse. Archiv Anatomie und Physiologie, 657–681.

Garrett, M.F. 1975. The analysis of sentence production. In: Bower, G.H. (Ed.), The Psychology of Learning and Motivation. Academic Press, New York.

Glaser, W.R., 1992. Picture naming. Cognition, 42, 61–105.

Goldsmith, J. 1990. Autosegmental and Metrical Phonology. Basil Blackwell, Cambridge, MA.

Harley, T.A., 1984. A critique of top-down independent levels of speech production: evidence from non-plan-internal speech errors. Cognitive Science, 8, 191–219.

Houghton, G. 1990. The problem of serial order: a neural network model of sequence learning and recall. In: Dale, R., Mellish, C., Zock, M. (Eds.), Current Research in Natural Language Generation. Academic Press, London.

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

Kaisse, E.M. 1985. Connected Speech: The Interaction of Syntax and Phonology. Academic Press, San Diego, CA.

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

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

Kenstowicz, M. 1994. Phonology in Generative Grammar. Blackwell, Oxford, UK.

Lashley, K.S. 1951. The problem of serial order in behavior. In: Jeffress, L.A. (Ed.), Cerebral Mechanisms in Behavior. Wiley, New York.

Levelt, W.J.M. 1989. Speaking: From Intention to Articulation. MIT Press, Cambridge, MA.

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

Levelt, W.J.M., Schriefers, H., Vorberg, D., Meyer, A.S., Pechmann, T., Havinga, J., 1991. The time

course of lexical access in speech production: a study of picture naming. Psychological Review, 98, 615–618.

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

Liberman, M., Prince, A., 1977. On stress and linguistic rhythm. Linguistic Inquiry, 8, 249–336.

Luce, R.D. 1986. Response Times: Their Role in Inferring Elementary Mental Organization. Oxford University Press, New York.

MacKay, D. 1987. The Organization Of Perception and Action: A Theory for Language and Other Cognitive Skills. Springer, New York.

Marslen-Wilson, W.D., Welsh, A., 1978. Processing interactions and lexical access during word recognition in continuous speech. Cognitive Psychology, 10, 29–63.

McCarthy, J., Prince, A., 1990. Foot and word in prosodic morphology: the Arabic broken plurals. Natural Language and Linguistic Theory, 8, 209–282.

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., 1992. Investigation of phonological encoding through speech error analyses: achievements, limitations, and alternatives. Cognition, 42, 181–211.

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.

Nespor, M., Vogel, I. 1986. Prosodic Phonology. Floris, Dordrecht, Holland.

Newell, A., Simon, H.A. 1972. Human Problem Solving. Prentice-Hall, Englewood Cliffs, NJ.

Nooteboom, S.G. 1969. The tongue slips into patterns. In: Sciarone, A.G., van Essen, A.J., van Raad, A.A. (Eds.), Nomen: Leyden Studies in Linguistics and Phonetics. Mouton, The Hague.

Peterson, R.R., Dell, G.S., O'Seaghdha, P.G. 1989. A connectionist model of form-related priming effects. Proceedings of the 11th Annual Conference of the Cognitive Science Society. Erlbaum, Hillsdale, NJ, pp. 196–203.

Posner, M.I. 1978. Chronometric Explorations of Mind. LEA, Hillsdale, NJ.

Roelofs, A., 1992. A spreading-activation theory of lemma retrieval in speaking. Cognition, 42, 107–142.

Roelofs, A. 1992b. Lemma retrieval in speaking: a theory, computer simulations, and empirical data. Doctoral dissertation, NICI Technical Report 92-08, University of Nijmegen.

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: Ram, A., Eiselt, K. (Eds.), Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society. LEA, Hillsdale, NJ, pp. 772–777.

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. 1996b. Morpheme frequency in speech production: Testing WEAVER. In: Booij, G.E., van Marle, J. (Eds.), Yearbook of Morphology. Kluwer Academic Publishers, Dordrecht, pp. 135–154.

Roelofs, A. 1996c. Computational models of lemma retrieval. In: Dijkstra, T., De Smedt, K. (Eds.), Computational Psycholinguistics: AI and Connectionist Models of Human Language Processing. Taylor and Francis, London, pp. 311–330.

Roelofs, A. 1997. Syllabification in speech production: evaluation of WEAVER. Language and Cognitive Processes 12, in press.

Roelofs, A., 1997b. A case for nondecomposition in conceptually driven word retrieval. Journal of Psycholinguistic Research, 26, 33–67.

Roelofs, A. 1997c. Rightward incrementality in encoding simple phrasal forms in speech production: verb-particle combinations. Journal of Experimental Psychology: Learning, Memory, and Cognition, in press.

Roelofs, A. submitted. Implicit and explicit priming of speech production: testing WEAVER.

Roelofs, A., Meyer, A.S. 1997. Metrical structure in planning the production of spoken words. Journal of Experimental Psychology: Learning, Memory, and Cognition, in press.

Roelofs, A., Meyer, A.S., Levelt, W.J.M., 1996. Interaction between semantic and orthographic factors in conceptually driven naming: comment on Starreveld and La Heij (1995). Journal of Experimental Psychology: Learning, Memory, and Cognition, 22, 246–251.

Schiller, N.O., Meyer, A.S., Baayen, R.H., 1996. A comparison of lexical and speech syllables in Dutch. Journal of Quantitative Linguistics, 3, 8–28.

Schriefers, H., Meyer, A.S., Levelt, W.J.M., 1990. Exploring the time course of lexical access in language production: picture-word interference studies. Journal of Memory and Language, 29, 86–102.

Selkirk, E. 1984. Phonology and Syntax: The Relation Between Sound and Structure. MIT Press, Cambridge, MA.

Shastri, L., Ajjanagadde, V., 1993. From simple associations to systematic reasoning: a connectionist representation of rules, variables and dynamic bindings using temporal synchrony. Behavioral and Brain Sciences, 16, 417–494.

Shattuck-Hufnagel, S. 1979. Speech errors as evidence for a serial-order mechanism in sentence production. In: Cooper, W.E., Walker, E.C.T. (Eds.), Sentence Processing: Psycholinguistic Studies Presented to Merrill Garrett. Erlbaum, Hillsdale, NJ.

Shattuck-Hufnagel, S., 1992. The role of word structure in segmental serial ordering. Cognition, 42, 213–259.

Spencer, A. 1991. Morphological Theory. Blackwell, Cambridge, MA.

Stemberger, J.P. 1985. An interactive activation model of language production. In: Ellis, A.W. (Ed.), Progress in the Psychology of Language. LEA, London, Vol. 1.

Touretzky, D.S., Hinton, G.E., 1988. A distributed connectionist production system. Cognitive Science, 12, 423–466.

Wheeldon, L.R., Levelt, W.J.M., 1995. Monitoring the time course of phonological encoding. Journal of Memory and Language, 34, 311–334.

Wickens, T.D. 1982. Models for Behavior: Stochastic Processes in Psychology. Freeman, San Francisco.