LANGUAGE AND COGNITIVE PROCESSES, 2002, *17* (3), 209–236

Inflectional frames in language production

Dirk P. Janssen
University of Arizona, Tucson, USA

Ardi Roelofs and Willem J.M. Levelt
Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

The authors report six implicit priming experiments that examined the production of inflected forms. Participants produced words out of small sets in response to prompts. The words differed in form or shared word-initial segments, which allowed for preparation. In constant inflectional sets, the words had the same number of inflectional suffixes, whereas in variable sets the number of suffixes differed. In the experiments, preparation effects were obtained, which were larger in the constant than in the variable sets. Control experiments showed that this difference in effect was not due to syntactic class or phonological form per se. The results are interpreted in terms of a slot-and-filler model of word production, in which inflectional frames, on the one hand, and stems and affixes, on the other hand, are independently spelled out on the basis of an abstract morpho-syntactic specification of the word, which is followed by morpheme-to-frame association.

INTRODUCTION

One of the more elusive distinctions in linguistics is that between inflectional and derivational morphology. Derivational morphology is prototypically used to create new words from existing words. For example,

---

Requests for reprints should be addressed to Dirk Janssen, GK Universalität und Diversität, Brühl 34–50, 04109 Leipzig, Germany. Email: dirkj@rz.uni-leipzig.de
Most of the research reported on in this paper was carried out while the first author was working at the MPI Nijmegen on a grant provided by the Max Planck Gesellschaft, München. Further work was done at the University of Arizona, supported by a grant from the Niels Stensen Stichting, Amsterdam. The authors would like to thank Ram Frost, Gerard Kempen, Dominiek Sandra, and three anonymous reviewers for their comments on earlier versions of this paper.

http://www.tandf.co.uk/journals/pp/01690965.html DOI: 10.1080/01690960143000182

by adding the suffix *-y* to the noun *wave*, the noun is turned into an adjective with a similar meaning. One can go on to add *-ness* to this, showing that derivation is recursive in a restricted way (**wav+y+y* is not allowed) and that meaning is far from perfectly preserved, as *waviness* cannot be substituted for *wave* in a great many contexts.

Inflectional morphology, on the other hand, creates semantically transparent forms like *sailed*, clearly analysable as *sail* plus <past tense>, and cannot be recursively applied. Depending on the syntactic class, words can select only from a limited assortment of inflectional affixes, for example, nouns take number but not tense affixes. In many languages, a further limitation is formed by the conjugational classes, which dictate the form of the inflectional affixes. In Spanish, the third person present tense form of *navegar* is *naveg+a*, while for *vivir* it is *viv+e*. Contrary to the situation for derivations, these limitations are all systematic, even though irregular cases exist.

Inflectional morphology has a number of interesting properties: First, individual inflectional forms are transparent in meaning. Transparency appears to be one of the factors that influence whether a form will be subject to morphological processes in language comprehension (e.g., Marslen-Wilson, Tyler, Waksler, & Older, 1994; Sandra, 1990; Zwitserlood, 1994). Second, inflectional forms are both transparent and predictable and this raises the question whether they should be stored or not. For derivations, both meaning and existence of a form cannot be predicted (compare *wave–wavy* and *cave–*cavy*), making it likely that they are listed in some way. Third, a small but frequent group of inflectional forms is irregular in form (e.g., *swim–swam*) and in contrast to other inflectional forms, these forms must necessarily be listed.

Whereas inflectional morphology has been intensively studied in the fields of language comprehension (e.g., McQueen & Cutler, 1998, for a recent review) and language acquisition (see Bloom, 1994, and Tager-Flusberg, 1997, for overviews and Plunkett, 1998, for connectionist approaches), it has largely been ignored in the field of language production. The results obtained for comprehension and acquisition cannot automatically be generalised to production. For example, in studies of language acquisition of the last two decades, learning the past tense has typically been construed as a task of learning a mapping between a present tense and a past tense phonological form of a word (e.g., Plunkett, 1998). However, this construal of the problem of inflection ignores the mapping between syntax and morphology, which seems to play a crucial role in the adult generation of inflected forms. Adult speakers generate inflected forms to express or mark certain morpho-syntactic information (e.g., to express number, person, tense, etc.) rather than as a response to a given present tense form (cf. Levelt, Roelofs, & Meyer, 1999).

In the present paper, we address the issue of adult generation of inflected forms from an abstract morpho-syntactic representation of the word, often called the lemma. Most of the evidence bearing on the inflectional mechanisms underlying adult speech production comes from speech errors (see Levelt, 1989, for a review). The evidence supports a distinction between an abstract morpho-syntactic specification of the word (e.g., first person singular, present tense, *to work*) and a level of concrete morphemic forms (e.g., *work*, *-z*, *-ed*, or an irregular form like *knew*). An error such as *that I'd hear one if I knew it* for *that I'd know one if I heard it* (from Garrett, 1980) suggests that words at the abstract morpho-syntactic level (*hear*, *know*) may trade places while stranding their morpho-syntactic specification (first person singular, present tense; first person singular, past tense). Furthermore, the swapped elements are typically of the same syntactic class (e.g., verbs exchange with verbs, and nouns with nouns). If concrete morphemic forms were exchanged, the error should have been *that I'd heard one if I know it.*

Such exchanges involving full forms also do occur, though. For example, in *well you can cut rain in the trees* (from Garrett, 1982), the full forms *rain* and *trees* have traded places, without stranding their abstract morpho-syntactic specification. In the latter type of form error, the syntactic class constraint often does not hold, but a constraint on morphological class applies. In morphological form errors, there is a strong morphological class constraint in that stems exchange with stems, and affixes exchange with affixes, but not stems with affixes. For example, stem exchange errors such as *rolls windowed up* for *windows rolled up* (from Garrett, 1975) occur, but errors such as *windowroll sed up*, in which a stem and an affix is exchanged, never occur.

To account for the properties of these types of morphological exchange errors (i.e., the stranding of abstract morpho-syntactic information, the syntactic class constraint, and the morphological form class constraint), researchers have proposed slot-and-filler accounts of speech production (e.g., Dell, 1986; Garrett, 1975, 1980, 1982; Levelt, 1989; Shattuck-Hufnagel, 1979). The general idea is that in planning speech, a speaker generates abstract syntactic, morphological, and phonological frames with categorically labelled slots (e.g., noun/verb, stem/affix). Retrieval from the mental lexicon makes available the so-called fillers (stems for the stem slots and affixes for the affix slots). Speech errors occur when a retrieved item is inserted in the wrong slot of the right category.

Although there are detailed theoretical proposals for the retrieval of the fillers, how the frames are generated is typically left unspecified. Levelt (1989) made a specific suggestion. According to his proposal, a word is represented at the abstract morpho-syntactic level as lemma (e.g., work), which has diacritical slots (for the specification of tense, number, etc.). The same word is represented at a form level by concrete morphemes (stems and affixes) and morphological frames. The lemma and diacritics form the basis for generating the morphological frames as well as the retrieval of a stem and affixes from the mental lexicon. This is followed by morpheme-to-frame association.

The goal of the current paper is to test Levelt's proposal for Dutch. As concerns regular nouns and verbs, two different inflectional frames are needed, namely a stem+suffix frame for nouns and stem+suffix+suffix frame for verbs. Following Dell's (1986) case for null elements as slot fillers, we assume that null morphemes are generated when an overt suffix is absent (an issue that is taken up in the General Discussion). For Dutch, this entails that the form *werk (to work,* present tense, first person singular) is generated from a frame containing *werk*+ø+ø. The past tense of irregular verbs is not expressed by a suffix but by stem allomorphy (compare English *eat-ate*), and no regular past tense suffix is necessary or possible (**ated*). Irregular verbs therefore require a stem+suffix frame, which happens to be the same as the frame for regular nouns. Table 1 lists the inflectional frames plus slot fillers for regular nouns and verbs in Dutch.

In planning a word, the appropriate frame is generated on the basis of the syntactic class of the word, which is specified at the lemma level. Independent of that, the system activates and selects stem and affix morphemes that appropriately encode the lemma diacritics. Generating an inflectional frame costs time, and therefore it should be possible to measure the process of frame generation in appropriately designed chronometric experiments. We have chosen to use as our research tool

TABLE 1
The inflectional frames plus slot fillers for regular nouns and verbs in Dutch. The example words are *boek (book)* and *zeilen (to sail)*

	Slot fillers		
	Stem	*Suffix*	*Suffix*
Noun diacritics			
singular	boek	+ ø	
plural	boek	+ en	
Verb diacritics			
present tense, 1st person singular	zeil	+ ø	+ ø
present tense, 2nd/3rd person singular	zeil	+ ø	+ t
present tense, any person plural	zeil	+ ø	+ en
past tense, any person singular	zeil	+ de	+ ø
past tense, any person plural	zeil	+ de	+ en

the speech-preparation or implicit priming paradigm originally developed by Meyer (1990, 1991). This paradigm has successfully been applied in several studies on form planning in production (e.g., Roelofs, 1996b, 1996a, 1997b, 1998, 1999; Roelofs & Meyer, 1998).

The preparation paradigm has been described in depth in various other places, and we refer to these publications for an extensive discussion of its properties (see especially Meyer, 1990, and Roelofs, 1998). In a preparation experiment, participants have to produce words out of small sets in response to written prompts. The words in a set differ in form (the heterogeneous sets) or share word-initial segments (the homogeneous sets). Meyer (1990, 1991) observed that production onset latencies for the response words are reduced in the homogeneous sets compared to the production of the same words in the heterogeneous sets, henceforth referred to as the *preparation effect*. For example, latencies are shorter for producing words from a set that includes *melon*, *metal*, and *merit* than for producing the same words grouped in sets where they do not share the first syllable. Faster latencies are not observed for sets in which the words share non-initial segments, for example the second syllable as in the set *pocket*, *ticket*, *racket*. This suggests that the facilitation effect from homogeneity is due to the preparation of word production rather than to general memory retrieval processes. Research on paired-associate learning has shown that form overlap helps memory retrieval independently of the place of overlap (see Meyer, 1990, for a review of the memory literature). Furthermore, immediate serial recall is hampered (i.e., a lower rather than a higher level of recall is observed) when the items are similar in sound or articulatory characteristics (see Baddeley, 1997, for a review). Thus, the findings from the memory literature are opposite to the results of Meyer (1990, 1991), which rules out a general memory account of her (and our) results. It has also been shown that preparation requires sharing full segments, and that sharing most phonological features (as with /b/ and /p/, which differ in voicing) yields no preparation effect at all (Roelofs, 1999).

Furthermore, Roelofs (1996b) showed that the preparation effect is larger when the shared segments make up a morpheme compared to when they do not (the words in the experiments of Meyer, 1990, 1991). Again, the increase in the preparation effect is only observed for word-initial morphemes, but not for non-initial ones. Also, the preparation effect for initial morphemes is larger for morphemes that are low-frequency in the language than for high-frequency morphemes (Roelofs, 1996a, 1998), which shows that benefit of preparation is larger for items that normally take longer to retrieve. In summary, in Meyer's original work a segmental preparation effect is obtained for word-initial segments that is independent of shared morphemes. If word-initial morphemes are shared however, an additional preparation effect is obtained.

In previous experiments (Meyer, 1990, 1991; Roelofs, 1996a, 1996b, 1997a, 1997b, 1998, 1999; Roelofs & Meyer, 1998), all words in a set had the same inflectional make-up. That is, all words had the same inflection or were not inflected and the appropriate inflectional frame could be generated before the beginning of a trial because it was predictable. This should not be possible, however, in sets where the inflections differ. If generating an inflectional frame takes time, one expects that having variable inflectional frames in a set incurs a production cost compared to sets with constant frames (just as segments that are not morphemes yield less preparation than segments that make up morphemes). This prediction is tested in the experiments in this paper.

Overview of the experiments

Experiments 1 and 2 tested the prediction that the preparation effect is reduced in inflectionally variable sets compared to inflectionally constant sets. In the constant sets, all words were singular nouns (requiring a stem+suffix frame) whereas the variable sets included singular nouns (requiring a stem+suffix frame) and inflected verbs (requiring a stem+suffix+suffix frame). We predicted that a preparation effect is obtained in all sets, but that the effect is larger in the constant than in the variable sets. Experiment 3 served as a control experiment to exclude that the preparation reduction is due to a special property of the phonological form of inflected verbs by replicating Experiment 1 with the inflected verb replaced by a homophonous singular noun (much like replacing the English verb *shoot* by *chute*). Since the critical sets are now inflectionally constant, no reduction should be obtained. Experiment 4 tested whether the reduction is indeed due to an inflectional frame difference as compared to different types of diacritical slots at the lemma level. Irregular verbs have exactly the same lemma level slots as regular verbs (e.g., tense, person, and number), but they require a stem+suffix frame, just like nouns. Again, no difference in preparation effect should be obtained, which would suggest that the reduction is due to differing morphological frames rather than lemma diacritics. Experiment 5 tested the same prediction using the present tense of irregular verbs. Up to now, the words possessed at most one overt suffix. Experiment 6, finally, tested for an effect of a frame difference using words with one or two overt suffixes. A frame difference should reduce the preparation effect, irrespective of the number of overt suffixes.

EXPERIMENT 1

The first experiment tested for a difference in preparation effect between sets with constant and sets with variable inflectional frames. To this end, two types of words were contrasted that are associated with different inflectional frames. We chose to use Dutch simple nouns (with 2-slot frames) and Dutch regular verbs (with 3-slot frames).

In the variable condition of this experiment, two nouns are combined with one verb. All items have the same CV onset and this gives rise to the classic phonological preparation effect (Meyer, 1990). This condition is compared to the constant condition, in which three nouns are used. These nouns share their CV onsets and in this condition the inflectional frames for the words can also be prepared, because all three words have the same 2-slot frame. This is not the case in the variable condition and if construction of an inflectional frame is mandatory for speech production, the frame has to be constructed on-line in the variable condition. This should lead to smaller preparation effects in the variable condition compared to the constant condition.

Method

Participants. Twelve participants took part in each of the experiments reported here, except where noted otherwise. They were all undergraduate students at the University of Nijmegen, native speakers of Dutch, and randomly taken from the Max Planck subject pool. They were paid for their participation. Each individual took part in only one of the experiments.

Materials. All experiments were carried out in Dutch. For each experiment, we created 4 practice sets and 12 experimental sets of three word pairs each. Each pair consisted of a prompt and a response word. Three onset syllables were selected that formed the basis for constructing the experimental sets. In this experiment, these were /he/, /po/, and /wa/. This last set is shown in Table 2 as an example.

There were two frame conditions, variable and constant. For the response words in a constant set, we selected three bisyllabic simple nouns,

TABLE 2
Example materials for Experiment 1, variable word type only

Homogeneous			
rivier	–waadde	*river*	*–wade* PAST
damp	–wasem	*steam*	*–mist*
kar	–wagen	*cart*	*–wagon*
Heterogeneous			
rivier	–waadde	*river*	*–wade* PAST
paradijs	–hemel	*paradise*	*–heaven*
model	–pose	*model*	*–pose*

that shared the initial syllable. For the response words in the variable sets, two more simple nouns were taken, together with a past tense form that started with the same initial syllable as the nouns. The past tense forms were *heette, pootte, waadde* (*be called*, *planted*, *waded*).

The response words were then combined with semantically related prompts. The prompts were chosen such that they formed strong and unambiguous retrieval cues for the corresponding responses. The sets mentioned earlier were used for the homogeneous conditions. For the heterogeneous conditions, prompt-response pairs were regrouped in such a way that a heterogeneous set contained three response words with three different onset syllables. In all sets, care was taken not to introduce unwanted phonological or semantic overlap between the three response words or between a prompt and any other prompt or response word. The pairing of prompts with responses was the same for homogeneous and heterogeneous conditions.

All response words were bisyllabic simple nouns, except for the three past tense forms that were bisyllabic root-affix combinations. All response words had initial stress and were chosen to be as dissimilar in form as possible. All prompts were nouns or adjectives of approximately the same length in letters and syllables, and were chosen to maximally differ from the other prompts in the set and from the response words. When combining prompts and responses we avoided pairs forming possible lexicalised combinations. See the Appendix for the full set of materials.

Four practice sets were created that mimicked heterogeneous and homogeneous sets of the constant and variable condition. The words in the practice sets were not related to any of the words used in the main experiment.

It should be noted that the past tense *waadde* is homophonous to the plural present tense form and the infinitive (both written *waden*). The same holds for the two other verbs. There can be no confusion though, because subjects have to read the response word in the learning phase and the written form is unambiguously that of a singular past tense.

Design. The experiment contained four crossed within-subject factors. The first factor was base (three levels). This factor corresponds to the three initial syllables that were used as a base for constructing the sets. Each base was used to create a variable and a constant condition. This is the factor word type (two levels). All words within the variable or the constant condition were tested in a homogeneous and a heterogeneous set. This factor will be called context (two levels). In the test phase, subjects responded to each prompt five times, this is the factor repetition (five levels).

The order in which the sets were presented to the subjects was fully counterbalanced. Half of the subjects (Groups A and B) got the variable condition first, the other half (Groups C and D) got the constant condition first. Within the variable or constant condition, subjects could see either the homogeneous sets first (Groups A and C), or the heterogeneous sets first (Groups B and D). As a last step, the order of presenting the three bases was varied across the three subjects within a group, such that a set made from a base morpheme occurred once as the first, second, or third set. The prompts of a set were repeated five times each in a block of trials. The order of presentation was fully randomised per subject and per block, with the constraint that there was no repetition of prompts on adjacent trials.

Procedure. The participants were tested individually in a quiet room. They were given written instructions, which stressed the fact that they should respond as fast and accurate as possible. The experiment consisted of alternating learning and test phases. In the learning phase, subjects were shown the three prompt-response pairs of a set on the computer screen. When they indicated that they had sufficiently studied the pairs (after about half a minute on average), the experimenter started the test phase. Each trial started with an attention sign (asterisk) marking the position of the prompt. The asterisk was displayed for 500 ms, then there was a 500 ms pause, and finally the prompt was displayed. At the same time, the voice key was opened for 1500 ms. Then the prompt disappeared, and after a 2500 ms pause, the cycle started again.

The experimenter sat in the same room and took note of hesitations, voice key errors, wrong responses and time outs. These trials were removed from the analyses. After each of the practice sets, subjects received feedback when necessary (i.e., when they made too many clicking noises that disturb the voice key, or when they had not memorised the pairs well enough).

The total time required for the experiment varied slightly as a function of the subject's learning time. An experimental session lasted 20 minutes on average, in which the subject's response to 60 practice and 180 experimental trials was measured. Subjects that were extremely slow learning the associations (more than a minute per set of three pairs), were removed from the analyses and replaced by others. This happened five times all together in the experiments reported here.

Apparatus. The experiment was controlled by a PC running the NESU program for controlling experiments, which was locally developed at the Max Planck Institute. Stimuli were presented to the subjects on a NEC Multisync30 15 inch monitor, positioned about 50 cm away from them. Their reactions were registered by a Sennheiser ME40 microphone, which fed to a NESUbox voice-key device and a DAT recorder. In the learning phase, the three pairs were displayed in a 20 point typewriter font. In the test phase, prompts were displayed in a 36 point sans serif font.

Analysis. From the output of the voice key device, we removed all responses resulting from voice key errors, all wrong responses, hesitations, and time outs. When in doubt, the experimenter consulted the recordings that were made of the sessions. In all experiments reported here, less than 5% errors were made. No more than half of these errors are undoubtedly due to subject error (these are wrong responses or hesitations, a time out can be caused by the voice key not picking up the speech signal). Error analysis were therefore carried out on all errors, although the number of data points is still quite small and the variance is low because we are dealing with 0–1 data. For these reasons, no interaction terms were computed for the error data.

Difference scores were computed for each response word by subtracting its mean RT in the homogeneous condition from the mean RT in the heterogeneous condition. The mean RTs were computed over the five repetitions of each prompt-response pair. When there were missing observations, the mean RT was computed on the remaining observations.

The statistical analyses included the remaining two factors, base and word type, and their interaction. Because both subjects and items are random variables, F' (quasi F) ratios were computed on the data. Significance of F' means that a replication of the experiment, with different words and different subjects, is expected to yield the same results (Clark, 1973; Forster & Dickinson, 1976; Raaijmakers, Schrijnemakers, & Gremmen, 1999). In general, only significant F ratios are reported. The reported MS_e is the interaction of subjects with the factor at hand and not the pooled term actually used in computing the F'.

To test whether there was an overall preparation effect, we report the test on the so-called ‘‘constant’’ term. This is a contrast comparing the mean of all cells to zero and again a F' test is used to compare this mean to both the within-items and within-subjects variance. Next, simple main effects are computed for the significance of the preparation effect for both levels of the word type factor, again using F'. In other papers using the implicit priming task, F_1 and F_2 ratios are often reported. The F' is slightly more conservative than the combination of these two tests, but its underlying assumptions are more satisfying.

Two further tests have been routinely done to ensure that the assumptions underlying the statistics were met: First, Tukey’s test for non-additivity was run. Second, we did Levene’s test on the homogeneity of variance. The results of these tests are only reported when the assumptions were not met (Maxwell & Bray, 1986; Santa, Miller, & Shaw, 1979).

Results and discussion

In this experiment, 2.2% of the observations were removed because of errors (no response, voice key error, response time out, wrong response). The results are summarised in Table 3. In the variable condition, where a verb was combined with two nouns, a much smaller preparation effect is obtained than in the constant condition with only nouns (30 vs. 67 ms). This is supported by the statistical tests: there was a significant overall preparation effect, $F'(1,20) = 26.30$, $MS_e = 13546$, $p = .000$. The preparation effect was larger for constant than for variable word type, $F'(1, 22) = 5.39$, $MS_e = 8518$, $p = .025$ and significantly larger than zero for each of the two levels of word type: for variable sets $F'(1,21) = 6.19$, $MS_e = 11032$, $p = .018$ and for constant sets $F'(1,21) = 28.53$, $MS_e = 11032$, $p = .000$. Inspection of the mean production latencies shows that subjects were overall slower in the constant conditions. Different materials were used in the two conditions and this prevents us from drawing any conclusions about this difference. The preparation effect in the variable condition was reduced for both verbal and nominal items.

Essentially, what is obtained is an effect of the inflectional frame on the preparation process. In the constant condition, the initial syllable could be prepared and so could the inflectional frame. Together, this led to a sizeable preparation effect of 67 ms. In the variable condition, the initial syllable could also be prepared, but not the inflectional frame. Because the choice between a 3-slot verbal frame and a 2-slot nominal frame could not be made until the prompt appeared and the target word was known, the inflectional frame had to be made available on-line. This reduced the preparation effect to 30 ms.

EXPERIMENT 2

The previous experiment indicated that the preparation effect in an implicit priming experiment is influenced by the predictability of the inflectional frame. Because this experiment is fundamental to the rest of this paper, we thought it necessary to replicate the experiment with partly different materials.

TABLE 3
Mean production latencies (ms), error percentages, and preparation effects (Δ) for Experiment 1

	Context				
Word type	*homogeneous*		*heterogeneous*		Δ
variable	655	2.0%	686	1.7%	30
constant	678	3.3%	746	1.7%	67

Method

The methods for this experiment were similar to those used for Experiment 1. See Experiment 1 for a detailed description of the procedure used for the implicit priming paradigm, the apparatus, and the analyses.

Participants. Twelve subjects from the Max Planck subject pool participated in the experiment. None of them had participated in Experiment 1 or in any other implicit priming experiment in the last three months before the experiment. They were paid for their efforts.

Materials. As in the previous experiment, the materials contained verbs and nouns that shared the initial syllable. The past tense form of the verbs was used, the nouns were all singular. Because few nouns exist that overlap in this way with past tense forms, only two of the three sets from Experiment 1 could be replaced. The three past tense forms tested in this experiment were *doodde (killed)*, *raadde (guessed)* and *heette (be called,* kept from Experiment 1). For each verb, five nouns were selected that overlapped with the first full syllable of the verb. Three of these nouns were assigned to the constant condition, the verb and the remaining two made up the variable condition. See the Appendix for the full set of materials. The nouns on /he/ were the same as those in the previous experiment, with one exception: The noun *Hema* was replaced by an adjective, *hevig*. There are not enough simple nouns on /he/ to make up a full set of nominal controls.

Results and discussion

Overall, 3.8% errors were made. The results are summarised in Table 4. There was a much larger preparation effect in the constant condition than in the variable condition (91 vs. 27 ms), the difference is even larger than in the original Experiment 1. The larger effect leads to low probability values: the overall preparation effect was significant at $F'(1,18) = 11.68$,

TABLE 4
Mean production latencies (ms), error percentages, and preparation effects (Δ) for Experiment 2: replication

	Context				
Word type	*homogeneous*		*heterogeneous*		Δ
variable	659	4.1%	686	2.2%	27
constant	648	3.5%	739	5.6%	91

$MS_e = 48855$, $p = .003$ and the difference between the variable and constant preparation effects reached $F'(1,23) = 7.91$, $MS_e = 13451$, $p = .008$. In an analysis of simple main effects, the preparation effect in the variable condition was not significant $F'(1,14) < 1$, but the preparation effect in the constant condition was $F'(1,14) = 7.86$, $MS_e = 98407$, $p = .013$.

The results of Experiment 1 were replicated with new subjects and mostly new items. The preparation effect in the constant condition is slightly larger than before. The effect in the variable condition is smaller and no longer significant in the simple main effects analysis. The pattern of results is quite the same and we take this to confirm that the inflectional frame plays a role in the production of nouns and verbs: Subjects' reaction times varied depending on whether words with the same or words with different inflectional frames were mixed.

One should in general be careful with comparing effect sizes between experiments, because the actual size of the preparation effect depends on many factors, and not all are systematically controlled for (e.g., subjects). However, because the constraints on the items in this experiment and in Experiment 1 were exactly the same, a statistical analysis of the differences between the two experiments can be done. First, the preparation effects for the constant sets in the two experiments (67 and 91 ms) were compared by submitting the data to an ANOVA with factors experiment (1 or 2), and base within experiment (the three possible word onsets in each experiment). The result shows that the difference between the two constant sets (the factor experiment) did not reach significance: $F'(2,21) = 1.70$, $MS_e = 7525$, $p = .208$. The only significant effect in the analysis was the overall preparation effect: $F'(1,21) = 32.28$, $MS_e = 26706$, $p = .000$. The variable sets from the two experiments were submitted to the same analysis and the factor experiment was again non-significant at $F'(66,17) = 0.25$, $MS_e = 25175$ while there was an overall preparation effect, $F'(1,17) = 5.81$, $MS_e = 24963$, $p = .026$.

Because the item *heette* and accompanying nouns were used in both experiments, we compared the outcomes for the /he/ items between the two experiments. This was done separately for the variable and constant conditions, and again no differences were found in a one-factor ANOVA comparing the levels of experiment: both for variable and constant sets, the factor did not reach significance.

EXPERIMENT 3

Some of the past tense forms that were used in Experiments 1 and 2 are homophonous to nouns. The form *waadde (waded)*, for example, has the same pronunciation as the word *wade (shroud)*. This can be used to our advantage by using these items again, but this time in the nominal sense. If our hypothesis is right, the inflectional frame is tied to the meaning of the word and not to its acoustic form. Hence, replacing *waadde* (the verb) by *wade* (the noun) should remove the difference between the constant and variable conditions that we found before.

In our previous experiments, the homophonous forms were disambiguated by the orthography of the form as presented in learning phase and by the choice of a prompt word that was strongly associated with the verbal sense. The same means were used in this experiment to insure that subjects produced these homophones in the nominal sense. Three verbs from the previous experiments were taken and all that was changed was their orthography and the associated prompt word.

The prediction is that a similar preparation effect will be found in both word type conditions of this experiment: The homophonic nouns will not be treated any differently from normal nouns. They need a 2-slot inflectional frame and no frame difference occurs in the variable condition. The same is true for the constant condition, and the preparation effects should be equivalent.

Method

The reader is referred to the materials section of Experiment 1 for a detailed description of the design, procedure, and analysis.

Materials. From Experiment 1 and 2, the sets that were constructed around the verbs *waadde (waded)*, *doodde (killed)*, and *pootte (planted)* were taken. The verbs were changed into their respective nominal homophones: *wade (shroud)*, *dode (corpse)*, and *poten (feet of an animal)*. New prompts were selected for the homophonic nouns that enforce the nominal reading of the word. The complete materials are listed in the Appendix.

Results and discussion

The results of the experiment are shown in Table 5. In total, 2.7% of the responses were counted as errors. The preparation effect in the variable and constant sets were indeed highly similar (25 vs. 30 ms). The statistical

TABLE 5
Mean production latencies (ms), error percentages, and preparation effects (Δ) for Experiment 3: homophones

	Context				
Word type	*homogeneous*		*heterogeneous*		Δ
variable	631	3.3%	656	3.0%	25
constant	637	2.0%	667	2.6%	30

analysis confirmed this interpretation: The factor word type (variable vs. constant) was not significant at $F'(37, 21) = 0.49$, $MS_e = 10194$, $p = .9728$. The effect for base (type of initial overlap: /*wa*/, /*do*/, or /*po*/) was nearly significant at $F'(3, 34) = 2.70$, $MS_e = 9432$, $p = .0640$. The overall preparation effect also approached significance: $F'(1, 14) = 3.33$, $MS_e = 44247$, $p = .087$. The non-significance of the latter test is not too alarming: further investigation of the data showed that this is due to variance between subjects and the fact that the items starting with /*po*/ did show a less substantial preparation effect across the board, see Figure 1 (the constant /*po*/ items produced sizeable preparation effects of 62, 43, and 58 ms in Experiment 1). The F' test is known to be a trifle conservative by nature and facing two sources of extraneous variance it will easily fail to reach significance even if the separate sources of variance are relatively small.

The nominal homophones did not reduce the preparation effect like their past tense counterparts did in previous experiments. This is what was expected because in the nominal sense, these words carry a 2-slot inflectional frame which is no different from that of ordinary nouns. Only in their verbal sense, a 3-slot frame is associated with the form and this deviating frame prevented the subjects from preparing the inflectional frame in Experiment 1 and 2.

Of course, one could argue that word class itself influences the preparation effects. Even though it is not obvious how and why word class can influence preparation, this alternative has to be taken quite seriously. The next experiment will, among other things, exclude the influence of word class *per se*.

EXPERIMENT 4

We have so far assumed that Dutch verbs are associated with inflectional frames with three slots: one for the stem, one for the tense marker, and one for the person/number marker. We have also claimed that frames are assigned to forms on the basis of syntactic class information. The 3-slot frame should therefore be assigned to all verbal forms. The exception to this rule are irregular forms. Similar to English, Dutch has a small class of verbs that do not receive the regular inflectional past tense marker, but change the vowel of the stem in the past tense in a process that linguists refer to as *ablaut*. Since a stem allomorph expresses the past tense diacritic, an overt regular tense suffix never occurs, and irregular verbs take a 2-slot frame. This implies that there should be no frame difference between irregular verbs and nouns. That is, combining irregular verbs and nouns in a set should not lead to a preparation reduction. Experiment 4 tests this prediction for the past tense and Experiment 5 tests the same prediction

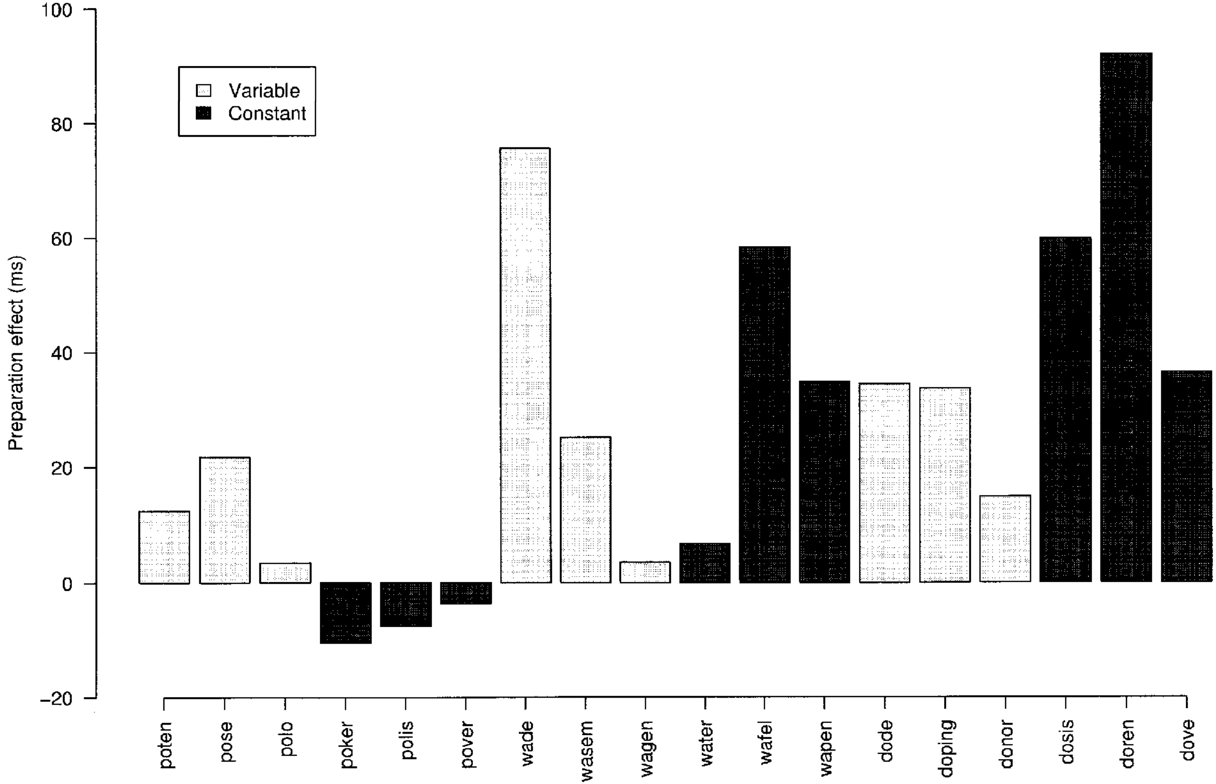

Figure 1. Mean preparation effects for the items in Experiment 3.

for the present tense. Note that if no preparation reduction is obtained in this experiment, this excludes the possibility that the preparation reduction in Experiments 1 and 2 was due to combining nouns and verbs in a set.

Method

Participants. There were 16 participants in this experiment, to enable full counterbalancing of the order of presentation of the four levels of the factor base used in this experiment.

Materials. Sets of four words were used in this specific experiment (all other experiments contained three-word sets). This was done because monosyllabic nouns were used, sharing only the onset rather than a full syllable (as in all other experiments), and smaller preparation effects are to be expected because of this. To keep the power of the experiments comparable, we increased the number of items and subjects in this experiment.

There were 4 practice sets and 16 experimental sets. Four irregular past tense forms were selected to form the basis for the experimental sets: *kreeg (got)*, *spoot (squirted)*, *droeg (carried)*, and *sliep (slept)*. The verbs were not completely irregular but were so-called strong verbs: their past tense is formed by vowel alternation. In all other respects, the verbs selected followed the regular Dutch paradigm. For the response words in the variable condition, we used the past tense and three additional monomorphemic, monosyllabic words that shared the onset cluster with the past tense. For the constant condition, four additional words were selected according to the same criteria.

Some Dutch nouns are homographic to first person, present tense verbal forms. We tried to avoid using these nouns, but seven cases had to be included to complete the materials. We ascertained that they were paired with prompts that activate the nominal reading. For six of them, the nominal frequency was also much higher than the verbal frequency (average of six was 1536 for N and 529 for V, per 42 million).

In the testing phase, each of the four prompts was presented four times to keep the length of the test phase comparable to the other experiments, in which three prompts were presented five times each. In all other respects, the construction of the materials was the same as in Experiment 1.

Results and discussion

Table 6 gives a summary of the reaction time latencies and preparation effects. The preparation effects for variable and constant word type are almost identical (44 and 43 ms). The overall preparation effect was larger than zero, $F'(1,21) = 10.15$, $MS_e = 79387$, $p = .004$. Both simple main

TABLE 6
Mean production latencies (ms), error percentages, and preparation effects (Δ) for Experiment 4: irregular past tense

	Context				
Word type	*homogeneous*		*heterogeneous*		Δ
variable	703	5.1%	746	3.4%	44
constant	715	5.1%	759	3.6%	43

effects were of course significant, but no other effects reached significance. Over the whole experiment, 4.3% errors were made. There were no significant effects in the errors. Levene's test was significant, signalling that the standard deviation varied substantially between conditions. According to Santa et al. (1979), the actual α in our test was slightly too liberal, which need not worry us.

The results are congruent with our assumptions. The equivalence of the effects for both conditions shows that, when the inflectional frame is predictable, no difference between the constant and variable conditions is obtained. Our analysis of irregular past tense forms was confirmed: These forms lack a tense suffix slot because they are inherently past tense.

Additionally, this again shows that word class has no effect on phonological preparation. More specifically, in this experiment mixing verbs and nouns leads to similar preparation effects in both conditions, and this contradicts an explanation of Experiment 1 and 2 in terms of verb–noun differences.

The next experiment tests the same predictions as the present experiment except that it requires the production of the present tense of strong verbs.

EXPERIMENT 5

Method

The experiment closely resembles the previous one: sets of four words were used to compensate for the lesser segmental overlap. For all other details, the reader is referred to Experiment 1.

Participants. Sixteen subjects participated in this experiment.

Materials. The materials for this experiment were similar to those used in Experiment 4 on irregular past tense forms. Four verbs were selected that were monomorphemic, monosyllabic, and did not have a competing nominal reading (this latter constraint made it impossible to reuse the verbs of the previous experiment). Monosyllabic nouns were chosen that had a CCVC or CCVCC sound pattern, and overlapped in initial consonants and vowel with the verbs. Many nouns from Experiment 4 could be reused. Matching prompt words were found that were semantically related to the response words. The prompts were chosen such that they formed strong and unambiguous retrieval cues for the corresponding response word and were not phonologically or semantically unrelated to any other response or prompt.

Four types of sets were made, based on the verbs *kruip, schiet, draag, sluit (crawl, shoot, wear, close)*. The constant sets contained four of the monosyllabic nouns, the variable sets contained three such nouns and a matching verb. An example variable set is *kruip, krent, kraal, kreeft (crawl, currant, bead, lobster)*, see the Appendix for the full list of materials. For heterogeneous sets, prompt-response pairs from four different sets were combined to create a set without overlap. Care was taken not to introduce unwanted semantic or phonological overlap between prompts and response words.

Results and discussion

The error rate in this experiment was even smaller than in other experiments: overall 2.0% of the observations had to be removed. A substantial priming effect was found in both the variable and constant sets: 47 and 29 ms (see Table 7). There was a significant overall preparation effect, $F'(1,17) = 5.53$, $MS_e = 123266$, $p = .0301$. No other factors in the reaction time or the error analysis were significant. Although the difference between the variable and constant sets was 18 ms, this was not significant. The test for the factor word type (variable vs. constant sets) yielded $F'(2,26) = 1.37$, $MS_e = 27291$, $p = .266$ and it should be noted that the direction of the difference is opposite to what would be expected under the assumption that the verbal forms carried regular 3-slot frames.

TABLE 7
Mean production latencies (ms), error percentages, and preparation effects (Δ) for Experiment 5: present tenses of irregular verbs

	Context				
Word type	*homogeneous*		*heterogeneous*		Δ
variable	677	1.8%	724	2.2%	47
constant	663	2.7%	692	1.3%	29

The variable condition (containing the present tense forms of irregular verbs) showed a preparation effect that was slightly, but insignificantly, larger than that in the constant condition. There is no trace of a reduction of the preparation effect in this condition, whereas a reduction was to be expected under the assumption of a regular 3-slot inflectional frame for the present tense of irregular verbs. This supports the idea that the inflectional frame for the present tense forms contains two slots. The frames for irregular verbs (regardless of tense) are irregular, 2-slot frames and in the combination with nouns, this leads to fully predictable inflectional frames and, hence, full-fledged preparation effects.

EXPERIMENT 6

A distinction made in many theories of the comprehension of morphologically complex words is that between storage and computation. We will follow the proposal of Schreuder and Baayen (1995) here. Some, but not all, complex forms are assumed to be stored in the lexicon and their meaning can be retrieved by look-up. Words that are used often and words with a non-decompositional meaning are obvious candidates for storage. The majority of inflectional forms are likely to be generated on the fly because inflectional forms always have a transparent meaning.

The difference between forms that can be retrieved from memory and forms that have to be generated can be used to construe an alternative account of the data: In our first two experiments, we mixed past tense forms with simple nouns. The past tense forms were built up from a stem and a suffix, and can be taken to be the result of computation. The simple nouns we used were bare lexical entries that can be produced on the basis of stored lexical information (note that this account has to deny the obligatory inflectional processing of all lexical form, or the storage-computation distinction is effectively nullified). An account of the equality of conditions in Experiment 3, 4, and 5 can also be given: in these experiments, only stored forms were used and no storage-computation distinction came into play.

In the language production model we advocate, the production of simple nouns involves more than lexical look-up. All words are submitted to an inflectional affixation process, even if there is no discernible result of singular marking on Dutch or English nouns. The two hypotheses can thus be contrasted by looking at plural forms. According to a storage-computation account, plural nouns are constructed on-line from their stem and the affix, whereas singular forms are looked up. According to our inflectional frame hypothesis, all words always undergo affixation and there should be no difference between singular and plural forms in our experiments.

This experiment will mimic Experiment 1 to a large extent. Instead of singular nouns and verbs, plural forms will be used throughout. Under both hypotheses of word production, it is predicted that all words undergo on-line affixation (computation): If the storage-computation hypothesis is correct, all sets in this experiment should behave the same. In contrast, under our hypothesis a smaller preparation effect should be obtained in the sets with a verb, because verbs require a different inflectional frame.

Method

Materials. For this experiment, bisyllabic plural nouns and bisyllabic plural past tenses were used, like *hanen* (*roosters*) and *haatten* (*hate* PAST, PLURAL). These words overlap in the first syllable. We expect to find a preparation effect, which will be reduced in the variable condition, where there is a frame difference between the nouns and the verb.

Three bisyllabic plural past tenses were selected: *haatten*, *raadden*, *doodden (hated*, *guessed*, *killed)*. Each verb was combined with five bisyllabic plural nouns that shared the first full syllable with the verb. Two of these nouns were combined with the verb to form the variable condition, the other three nouns formed the constant condition. In all other respects, this experiment was carried out exactly as described for Experiment 1. See the Appendix for the full set of materials.

Results and discussion

Overall, 3.8% of the items had to be removed because of errors. Table 8 shows the results of the experiment. The preparation effect in the constant condition is large, and the effect in the variable condition is particularly small (70 vs. 22 ms). This difference is significant, $F'(1, 23) = 4.69$, $MS_e = 15642$, $p = .0372$. Overall, subjects responded faster in the homogeneous conditions than in the heterogeneous conditions: $F'(1, 19) = 10.45$, $MS_e = 31612$, $p = .004$, but the preparation effect in the variable condition was too small to reach significance in a test of simple main effects. For variable sets, the simple main effect was $F'(1, 22) = 1.36$, $MS_e = 23155$, $p = .267$ and

TABLE 8
Mean production latencies (ms), error percentages, and preparation effects (Δ) for Experiment 6: plural forms

	Context				
Word type	*homogeneous*		*heterogeneous*		Δ
variable	688	4.3%	710	2.2%	22
constant	675	4.3%	745	4.3%	70

for constant sets it was $F'(1,22) = 12.40$, $MS_e = 23155$, $p = .002$. There were no other significant differences, nor any significant effects in the errors.

This outcome shows that the storage-computation explanation does not hold for the results we have obtained. A difference between variable and constant sets was found in this experiment, although no words could directly be retrieved from the lexicon. The explanation given for Experiment 1 still holds. The presence of a number suffix on all forms does not influence the pattern of results, because inflectional frames contain slots for all *possible* types of inflectional affixes.

Compared to the previous experiments, the preparation effect in the variable condition (20 ms) is rather small. Before, values of 30 ms and larger were found, which were significant in analyses of simple main effects for the variable conditions. In the present experiment, this simple main effect was not significant and this might indicate that the larger variety of affixes used in this experiment imposes an additional workload on the affix encoding stage.

GENERAL DISCUSSION

The goal of the reported research was to obtain evidence about inflectional processes in speech production in general, and to test a specific theoretical proposal, Levelt's (1989) slot-and-filler model, in particular. According to this model, the production of inflected forms involves generating inflectional frames, on the one hand, and stems and affixes, on the other hand, on the basis of an abstract morpho-syntactic specification of the word, followed by morpheme-to-frame association. Generating an inflectional frame costs time, and therefore it should be possible to measure it in appropriately designed chronometric experiments. We tested for effects of inflectional frame generation using the speech-preparation paradigm developed by Meyer (1990, 1991).

Experiments 1 and 2 tested the prediction that the preparation effect is reduced in inflectionally variable sets compared to inflectionally constant sets. In the constant sets, all words were singular nouns whereas the variable sets included singular nouns and inflected verbs. In both experiments a preparation effect was obtained in all sets, which was larger in the constant than in the variable sets. Experiment 3 excluded that this difference in effect was due to a special property of the phonological form of inflected verbs by replicating Experiment 1 with the inflected verb replaced by a homophonous singular noun. Since the critical sets were now inflectionally constant, no reduction should be obtained, which was indeed observed. Experiment 4 tested whether the reduction was indeed due to an inflectional frame difference as compared to different diacritical slots at the lemma level. Strong verbs have the same diacritical slots as weak verbs but differ in inflectional frame. Again, no difference in preparation effect was obtained, which suggests that the reduction is due to differing morphological frames rather than lemma diacritics or to a difference between nouns and verbs per se. Experiment 5 empirically confirmed the same prediction using the present tense of strong verbs. Experiment 6, finally, tested for an effect of a frame difference using words with one or two overt suffixes. Again, the frame difference reduced the preparation effect.

In all experiments a preparation reduction was observed when the overt number of suffixes on the *odd* item in a set was different from the overt number of suffixes of the other items. In Experiment 1, for example, the *odd* item was the inflected verb (with -*de*). This form was compared to bare nouns, and a preparation reduction was obtained. Thus, it would seem that the reduction might be due to a confusion of the participants about when to realise an overt affix. Or alternatively, the reduction might be due to competition between actual suffixes. Competition or confusion between suffixes cannot explain, however, why the results for the constant sets in Experiment 6 came about. Nouns with either of the two plural suffixes in Dutch (-*s* and -*en*) were produced equally often, but no preparation reduction was obtained for these items.

Still, the fact that the reduction is observed when the number of overt suffixes differ raises the more fundamental question of the necessity of null morphemes. Is it possible to define inflectional frames in terms of the number of *actual* suffixes rather than the number of *possible* affixes, as is done now? Inflectional frames that specify the number of overt suffixes (“tailor-made frames”) eliminate the need for null morphemes. Dell (1986) introduced null elements as slot fillers in order to keep the generation of frames simple and general (“one size fits all”). In particular, the generation of inflectional frames only has to take the syntactic class and, in some cases, irregularity, into account. If instead, frames are generated to exactly accommodate the overt suffixes of the inflected form, then the frame generation process also has to consult the values of the lemma diacritics (specifying person, number, and tense). Frame generation and affix spell out become highly redundant in this approach: For generating frames, we require knowledge of which (combinations of) diacritics are not overtly expressed, and for affix spell out we need to know how the remaining diacritics are overtly expressed.

The results of Experiments 1 to 6 do not distinguish between the two theoretical possibilities for frame generation and there are very few possibilities to put this to test in a language like Dutch. As said, in all our experiments a difference in “one size fits all” inflectional frames went hand-in-hand with a difference in the number of overt morphemes, and hence a difference in "tailor-made" frames. Cast in terms of "tailor-made" frames, Experiments 1 and 2 contrasted 1-slot frames (nouns) with 2-slot frames (verbs+tense), which explains the observed preparation reduction. Experiments 3 to 5 compared 1-slot frames (nouns) with other 1-slot frames (verbs), which explains why no preparation reduction was obtained. Finally, Experiment 6 contrasted 2-slot frames (noun+number) and 3-slot frames (verb+tense+number), which explains the preparation reduction that was again obtained. An equally powerful explanation can thus be obtained, which does not have to assume null morphemes but is more complicated in its frame generation part. Further experiments are needed to decide between the two possibilities.

The outcomes of Experiment 5 are of special interest for the on-going discussion of the past tense generation. The present tense of irregular verbs was shown to bear an irregular inflectional frame in this experiment. This finding is at odds with the assumptions of connectionist models, which claim that there is no irregularity in the system. The structured tree representation proposed by Clahsen (1999) is not compatible with our findings for the same reason, because the root of the structured tree (the present tense form) cannot currently bear any traits (Janssen, 1999).

To conclude, together the results from Experiments 1 to 6 support a slot-and-filler account of the generation of inflected forms in speech production. According to this account, inflectional frames are independently spelled out from stems and affixes. Generation of frames and stems is based on an abstract morpho-syntactic specification of the word and involves lexical lookup of the stem followed by morpheme-to-frame association. Generation of regular suffixes is done by morpheme-to-frame association of the suffixes for the remaining abstract morpho-syntactic specifications. When an irregular form is produced, this is reflected in the spell out of both the stem and the frame, but morpheme-to-frame association stays the same.

Manuscript received January 1999
Revised manuscript received March 2001

REFERENCES

Baddeley, A.D. (1997). *Human memory: Theory and practice* (Revised ed.). Hove, UK: Psychology Press.

Bloom, P. (1994). Controversies in language acquisition. In M.A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 741–779). San Diego, CA: Academic Press.

Clahsen, H. (1999). Lexical entries and rules of language: a multidisciplinary study of German inflection. *Behavioral and Brain Sciences*, *22*, 991–1060.

Clark, H.H. (1973). The language-as-a-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, *12*, 335–359.

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

Forster, K.I., & Dickinson, R.G. (1976). More on the language-as-a-fixed-effect fallacy: Monte Carlo estimates of error rates for F_1, F_2, F', and *min* F'. *Journal of Verbal Learning and Verbal Behavior*, *15*, 135–142.

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

Garrett, M.F. (1980). Levels of processing in sentence production. In B. Butterworth (Ed.), *Language production. Vol. 1.* London: Academic Press.

Garrett, M.F. (1982). Production of speech: Observations from normal and pathological language use. In A.W. Ellis (Ed.), *Normality and pathology in cognitive functions* (pp. 19–76). London: Academic Press.

Janssen, D.P. (1999). The place of analogy in Minimalist Morphology and the irregularity of regular forms. *Behavioral and Brain Sciences*, *22*, 1025–1026.

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

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

Marslen-Wilson, W., Tyler, L.K., Waksler, R., & Older, L. (1994). Morphology and meaning in the English mental lexicon. *Psychological Review*, *101*, 3–33.

Maxwell, S.E., & Bray, J.H. (1986). Robustness of the quasi *F* statistic to violations of sphericity. *Psychological Bulletin*, *99*, 416–421.

McQueen, J.M., & Cutler, A. (1998). Morphology in word recognition. In A. Spencer & A.M. Zwicky (Eds.), *The handbook of morphology* (pp. 406–427). Oxford: Blackwell.

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: Phonological encoding inside a syllable. *Journal of Memory and Language*, *30*, 69–89.

Plunkett, K. (Ed.) (1998). *Language acquisition and connectionism.* Hove, UK: Psychology Press. (Reprint of *Language and Cognitive Processes* special issue 13:2/3.)

Raaijmakers, J.G.W., Schrijnemakers, J.M.C., & Gremmen, F. (1999). How to deal with "the language-as-a-fixed-effect fallacy": Common misconceptions and alternative solutions. *Journal of Memory and Language*, *41*, 416–426.

Roelofs, A. (1996a). Morpheme frequency in speech production: Testing WEAVER. In G.E. Booij & J. van Marle (Eds.), *Yearbook of morphology*. Dordrecht: Kluwer Academic Press.

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

Roelofs, A. (1997a). Syllabification in speech production: Evaluation of WEAVER. *Language and Cognitive Processes*, *12*, 657–693.

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

Roelofs, A. (1998). Rightward incrementality in encoding simple phrasal forms in speech production: Verb-particle combinations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 904–921.

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

Roelofs, A., & Meyer, A.S. (1998). Metrical structure in planning the production of spoken words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 922–939.

Sandra, D.M.J. (1990). *Processing and representational aspects of compound words in visual word recognition: An experimental approach and a methodological appraisal*. Unpublished doctoral dissertation, Universiteit van Amsterdam.

Santa, J.L., Miller, J.J., & Shaw, M.L. (1979). Using quasi *F* to prevent alpha inflation due to stimulus variation. *Psychological Bulletin*, *86*, 37–46.

Schreuder, R., & Baayen, R.H. (1995). Modeling morphological processing. In L.B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 131–154). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Shattuck-Hufnagel, S. (1979). Speech errors as evidence for a serial order mechanism in sentence production. In W.E. Cooper & E.C.T. Walker (Eds.), *Sentence processing: Psycholinguistic studies presented to Merrill Garrett* (pp. 295–342). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Tager-Flusberg, H. (1997). Putting words together: morphology and syntax in preschool years. In J. Berko Gleason (Ed.), *The development of language* (4th ed., pp. 159–209). Boston, MA: Allyn and Bacon.

Zwitserlood, P. (1994). The role of semantic transparency in the processing and representation of Dutch compounds. *Language and Cognitive Processes*, *9*, 341–368.

APPENDIX

Experiment 1, variable sets

aardappels–pootte (potatoes–plant PAST) *model–pose* (model–pose) *sport–polo* (sport–polo); *rivier–waadde* (river–wade PAST) *damp–wasem* (steam–mist) *kar–wagen* (cart–wagon); *naam–heette* (name–be called PAST) *paradijs–hemel* (paradise–heaven) *winkel–hema* (shop–Hema)

Experiment 1, constant sets

gokspel–poker (gambling game–poker) *contract–polis* (contract–insurance) *zuinig–pover* (stingy–meagre, poor); *oase–water* (oasis–water) *gebak–wafel* (cake–waffle) *pistool–wapen* (pistol–weapon); *toekomst–heden* (future–present) *afkeer–hekel* (dislike–aversion) *slang–hevel* (tube–siphon)

Experiment 2, variable sets

moord–doodde (murder–kill PAST) *sporter–doping* (sportsman–doping) *nier–donor* (kidney–donor); *gok–raadde* (gamble–bet PAST) *vliegtuig–radar* (airplane–radar) *stof–rafel* (fabric–loose ends); *naam–heette* (name–be called PAST) *paradijs–hemel* (paradise–heaven) *erg–hevig* (very–intense)

Experiment 2, constant sets

medicijn–dosis (drug–dose) *roos–doren* (rose–thorn) *blinde–dove* (blind person–deaf person); *trend–rage* (trend–hype) *onkosten–raming* (expenses–estimate) *toeter–ratel* (horn–rattle); *toekomst–heden* (future–present) *afkeer–hekel* (dislike–aversion) *slang–hevel* (tube–siphon)

Experiment 3, variable sets

voeten–poten (feet–feet (of an animal)) *model–pose* (model–pose) *sport–polo* (sport–polo); *priester–wade* (priest–shroud) *damp–wasem* (steam–mist) *kar–wagen* (cart–wagon); *moord–dode* (murder–corpse) *wedstrijd–doping* (match–doping) *nier–donor* (kidney–donor)

Experiment 3, constant sets

gokspel–poker (gambling game–poker) *contract–polis* (contract–insurance) *zuinig–pover* (stingy–meagre); *oase–water* (oasis–water) *gebak–wafel* (cakes–waffle) *pistool–wapen* (pistol–weapon); *medicijn–dosis* (drug–dose) *roos–doren* (rose–thorn) *blinde–dove* (blind person–deaf person)

Experiment 4, variable sets

kado–kreeg (present–receive PAST) *pater–kruin (father–tonsure) rozijn–krent* (raisin–currant) *schram–kras* (scrape–scratch); *water–spoot* (water–squirt PAST) *zenuw–spier* (nerve–muscle) *baby–speen* (baby–dummy) *wiel–spaak* (wheel–spoke); *kleding–droeg* (clothing–wear past) *snoep–drop* (candy–licorice) *bier–drank* (beer–drink) *galop–draf* (gallop–trot); *nacht–sliep* (night–sleep PAST) *teug–slok* (drink–sip) *klap–slag* (punch–blow) *reptiel–slang* (reptile–snake)

Experiment 4, constant sets

pijn–kramp (pain–cramp) *nieuws–krant* (news–newspaper) *prins–kroon* (prince–crown) *stoel–kruk* (chair–stool); *vogel–specht* (bird–woodpecker) *haast–spoed* (hurry–speed) *trein–spoor* (train–railroad track) *hobby—port* (hobby–sport); *sprookje–draak* (fairy tale–dragon) *wijn–druif* (wine–grape) *slaap–droom* (sleep–dream) *trommel–drum* (drum–drum); *brug–sluis* (bridge–lock) *modder–slib* (mud–silt) *poort–slot* (gate–lock) *knecht–slaaf* (servant–slave)

Experiment 5, variable sets

tunnel–kruip (tunnel–crawl) *rozijn–krent* (raisin–currant) *ketting–kraal* (chain–bead) *zeester–kreeft* (starfish–lobster); *geweer–schiet* (rifle–shoot) *haring–schol* (herring–plaice) *boerderij–schuur* (farm–barn) *laars–schoen* (boot–shoe); *kleding–draag* (clothing–wear) *galop–draf* (gallop–trot) *wijn–druif* (wine–grape) *snoep–drop* (candy–licorice); *poort–sluit* (gate–close) *winter–slee* (winter–sledge) *reptiel–slang* (reptile–snake) *teug–slok* (drink–sip)

Experiment 5, constant sets

stoel–kruk (chair–stool) *peuter–creche* (toddler–kindergarten) *nieuws–krant* (news–newspaper) *vaas–kruik* (vase–hot-water bottle); *geit–schaap* (goat–sheep) *strand–schelp* (beach–shell) *diskette–schijf* (diskette–floppy) *zwaard–schild* (sword–shield); *naald–draad* (neadle–thread) *knal–dreun* (bang–rumble) *poep–drol* (shit–turd) *bier–drank* (beer–drink); *modder–slib* (mud–silt) *olifant–slurf* (elephant–trunk) *boot–sloep* (boat–barge) *knecht–slaaf* (servant–slave)

Experiment 6, variable sets

vijand–haatten (enemy–hate PAST, PL) *kip–hanen* (chicken–roosters) *spijker–hamers* (nail–hammers); *gok–raadden* (bet–guess PAST, PL) *trend–rages* (trend–hypes) *scheur–rafels* (tear–loose ends); *moord–doodden* (murder–kill PAST, PL) *nier–donors* (kidney–donors) *ei–dooiers* (egg–yolks)

Experiment 6, constant sets

konijn–hazen (rabbit–hares) *boot–havens* (ship–harbours) *sultan–harems* (sultan–harems); *venster–ramen* (window–windows) *toeter–ratels* (horn–rattles) *kraai–raven* (crow–ravens); *roos–dorens* (rose–thorns) *kist–dozen* (chest–boxes) *blinde–doven* (blind person–deaf persons)