Producing past and plural inflections

MPI Series in Psycholinguistics

1. The electrophysiology of speaking: Investigations on the time course of semantic, syntactic, and phonological processing
   Miranda van Turennout
2. The role of the syllable in speech production: Evidence from lexical statistics, metalinguistics, masked priming, and electromagnetic midsagittal articulography
   Niels Schiller
3. Lexical access in the production of ellipsis and pronouns
   Bernadette Schmitt
4. The open-/closed-class distinction in spoken-word recognition
   Alette Haveman
5. The acquisition of phonetic categories in young infants: A self-organising artificial neural network approach
   Kay Behnke
6. Gesture and speech production
   Jan-Peter de Ruiter
7. Comparative intonational phonology: English and German
   Esther Grabe
8. Finiteness in adult and child German
   Ingeborg Lasser
9. Language input for word discovery
   Joost van de Weijer
10. Inherent complement verbs revisited: Towards an understanding of argument structure in Ewe
    James Essegbey
11. Producing past and plural inflections
    Dirk Janssen
12. Valence and transitivity in Saliba, an Austronesian language of Papua New Guinea
    Anna Margetts
13. From speech to words
    Arie van der Lugt

Cover illustration by the author, with experimental stimuli drawn by Esther Grabe
ISBN 90-64648-17-4

Producing past and plural inflections

een wetenschappelijke proeve
op het gebied van de
Sociale Wetenschappen

Proefschrift

ter verkrijging van de graad van doctor
aan de Katholieke Universiteit Nijmegen,
volgens besluit van het College van Decanen
in het openbaar te verdedigen op
dinsdag 23 november 1999,
des namiddags om
3.30 uur precies

door Dirk Paul Janssen
geboren op 19 oktober 1969 te Tilburg.

Promotor: prof. W. J. M. Levelt
Co-promotor: dr. A. Roelofs (University of Exeter)
Manuscriptcommissie: Prof. H. Schriefers
dr. J. McQueen
dr. R. H. Baayen

The research reported in this thesis was supported by a grant from the Max-Planck-Gesellschaft zur Förderung der Wissenschaften, München, Germany.

Voorwoord

Terugkijkend op de laatste drie en half jaar is dit een periode geweest waarin ik vooral heel veel plezier heb gehad: Ik werkte aan een project dat me echt interesseerde en waar vernieuwende resultaten uitkwamen; ik had aardige collega's, goede vrienden en leuke huisgenoten; en er was tijd genoeg om me in 1000 andere dingen te storten. Die gelukkige tijd was een plezierige samenloop van omstandigheden, maar is ook te danken aan de inzet en energie van heel veel andere mensen. Jullie wil ik hier bedanken.

Allereerst zijn daar Ardi Roelofs en Pim Levelt, die me voortreffelijk begeleid hebben door me enthousiast te maken, terug te fluiten en goede raad te geven. De talloze discussies met hen hebben me veel geleerd en hebben de inhoud van dit proefschrift voor een groot deel bepaald. Meer informeel heb ik ook hulp gehad van James McQueen, Antje Meyer, en Laura Walsh Dickey. De wetenschap drijft behalve op ratio ook op ervaring en inschatting en een paar extra bronnen van dat soort informatie kwamen me altijd goed van pas, vooral toen Ardi naar Exeter verhuisd was.

Naast al die mensen die ouder en wijzer zijn, zijn er een heleboel mensen op mijn werk waarmee ik even kon ontspannen, bijkletsen of naar de film gaan. Deze groep bevat eigenlijk iedereen die *PhD student* (of iets soortgelijks) aan het Max Planck was: Samen zijn ze medeverantwoordelijk voor de geweldige sfeer op het Instituut. Speciale aandacht voor (in min of meer chronologische volgorde) maestro Arie van der Lugt, mede-nerd Hedderik van Rijn, goedlachse kamergenoot Delphine Dahan, Esther *kruidenthee* Grabe, goedlachse buurkamergenoot Alette Haveman, de enige echte Maarten van Casteren, collega-psychomorfoloog Kerstin Mauth, Fryske-limbo Femke van der Meulen, en bijzonder ontspannen kamergenoot Ger Desserjer.

Mijn leven buiten het werk werd opgevrolijkt door *de vriendenklup:* de meest fascinerende verzameling tegenstrijdige karakters die ik ooit gezien heb en, 11 jaar na dato, het levend bewijs dat het heel veel uit kan maken met wie je je introductie hebt. Speciale aandacht voor 'La' Katja Duurland, waarmee het goed wijn drinken is. Dat geldt trouwens ook voor Michael Cysouw en Nens van Alfen, en voor veel meer andere mensen die ik niet allemaal bij naam ga noemen omdat het toch wel duidelijk is dat ik ze graag mag.

Het laatste half jaar van het project was nogal anders dan de rest: Terwijl mijn contract afliep, ging ik harder werken om de periode van het eindeloze schrijven en herschrijven te bekorten. De werkdruk steeg enorm en de regelmatige avondjes of dagjes uit werden echt een noodzaak. In die hectische tijd besloot ik ook nog mijn relatie met Rianne te verbreken. Geen bijzonder geschikt moment voor het afsluiten van zes jaar liefde en geluk: Toch wil ik je heel erg bedanken voor alle warmte, plezier, inspiratie en wijsheid die je me in de jaren daarvoor gegeven hebt. Merci.

In dat veelbewogen voorjaar had ik (naast Femke en Kerstin) heel veel aan mijn geweldige huisgenoten Esther, Francisco en Sietske. Bedankt voor alles, jullie zijn (en hebben) het leukste huis van Nijmegen! Het leukste huis van Riel is ook niet moeilijk te vinden en Carel en Loes, mijn ouders, kunnen persoonlijk verantwoordelijk gesteld worden voor het feit dat ik nu in de wetenschap zit, bijna altijd goede zin heb, en naar een ver land ga verhuizen. Met vijf reizen per jaar geven jullie gelukkig nog steeds het goede voorbeeld. Mijn 'kleine zusje' Marion doet nog veel spannendere (en nuttigere) dingen dan ik en ik vind het iedere keer weer leuk om te zien hoe we naar elkaar toe zijn gegroeid.

Er valt, blijkt, haast niet te ontkomen aan de vastgeroeste opsomming van mensen die belangrijk waren of zijn, van zakelijk tot privé. Terwijl ik eigenlijk net zo goed in één keer had kunnen zeggen dat ik jullie allemaal zal missen, komend jaar in Arizona en Illinois. Ik hoop dat ik daar weer interessante experimenten kan doen en net zulke leuke mensen zal ontmoeten.

Nijmegen, augustus 1999,

Dirk

I'll put my suitcase here for now
I'll turn the TV to the bed
but if no-one calls and I don't speak all day
do I disappear?
Everything but the Girl, Walking wounded

Contents

Reading guide

Apart from members of the committee, who are obliged to read the whole book, and fellow morphologists, who already know half of what will be said here, few people will read this book from cover to cover. For those, it is probably easiest to first read the first two sections of the first chapter that deal with the definition of inflection, and where an overview of Dutch morphology is given. The rest of Chapter 1 and Chapter 2 provides the background for the study and can be skipped. Next, I advise you to read Chapter 3, which contains questions and predictions that will be tested in the following experimental chapters. Your previous experience with psycholinguistic experiments determines at what speed you may want to browse through these chapters. The penultimate Chapter 6 is in a sense the most interesting, but probably also the hardest to understand. If you are interested in theoretical issues concerning morphology and the conceptual system, this chapter is definitely worth reading. Otherwise, spend your precious time reading Chapter 7, which summarises the results.

The split morphology hypothesis

CHAPTER 1

1.1 Inflection and derivation

In a dissertation devoted to the production of inflectional forms, it is of course of great importance to delineate which forms are inflectional, and which forms are not. Special attention should be paid to the distinction between inflection and its closest neighbour, derivation. However, no good definition of the two can be given. Although *prototypical* inflection and *prototypical* derivation are concepts that are intuitively clear and that are agreed upon by the majority of morphologists, and although for most concrete examples the field has by and large agreed on their status, the two phenomena seem to resist formalisation (Matthews, 1974). Informally, one can say that for every imaginable theory of inflection, language seems to have a counterexample up its sleeve. No attempt at solving this problem will be made in this thesis, although the experimental results of Chapter 4 might help finding a solution one day.

So what *are* inflection and derivation? The prototypical uses of the terms can best be demonstrated by examples from English: *Stormy* in *a stormy night* is a prototypical derivation in that the *-y* suffix creates an adjective from a noun (*storm*). The *-y* suffix plays the same trick with many other nouns, and the result is, usually, an adjective with a meaning that can quite successfully be derived from the meaning of the noun. The hedge is necessary: Notice how *a stormy night* is not at all a night that is like a storm but a night full of storm(s), and how this is fundamentally different from *a creamy taste* or *a flashy outfit*. Derivations can also move away from their base words, up to the point where there hardly is any relation in meaning between the two. Examples are *catty* in *a catty remark*, and *department* in any context.

Compare this to the prototypical inflectional form *sailed:* This is also formed from a base word, but the results are of the same word class. The past tense suffix can be applied to almost all verbs, and the meaning of the inflection is fully predictable. Another prominent case of inflection is number marking, as in *The J-22s approached the buoy on starboard tack.* This example was chosen to demonstrate that newly formed words (like the ship type *J-22*) receive a plural form in a strictly predictable fashion, and no difficulties with the interpretation ever occur. Inflection is not exempt from irregularity, however. Irregular plural formation does exist for a limited number of forms (*thesis, theses*), and marginally there is suppletion: A form standing in for an inflectional variant without there being an obvious relation between the two, like in *go–went*.

It was mentioned above that the distinction between inflection and derivation seems hard to make. This claim needs some support and (while intending the opposite) this was provided by Scalise (1988), who devoted a paper to the issue. Scalise mentions no less than 15 grounds for a distinction, and discusses their merits. Some of the criteria mentioned are specific to the theoretical framework of the Government-and-Binding theory that was popular then, and most of these are already outdated. Other, more traditional, criteria are either hard to use predictively or do not speak on all forms. An example of the latter is the criterion that only derivation may change word class: *Greenish* is of the same word class as its base, but it is nevertheless considered derivational.

The upshot of Scalise's review is that the inflection–derivation distinction cannot be precisely defined by a formal criterion, although the distinction is of undeniable importance to morphology. Scalise does mention the definition given by Anderson (1982), but it is regrettably excluded from his discussion. In this thesis, I will adopt Anderson's definition as a working hypothesis. It will mainly figure in this chapter, where Anderson's work will be surveyed. In the later chapters, prototypical inflection and derivation will be used, and the precise definition is of less importance.

1.2 Inflection and derivation in Dutch

This thesis will largely be about the verbal inflections of Dutch. Dutch has two inflectionally expressed tenses, present and past. The past tense is marked by either *-de* or *-te*, depending on the underlying voicing of the stemfinal consonant, compare column 2 and 3 of Table 1.1. There are 4 markers that express the combinations of number and tense, a fifth marker is reserved to distinguish between first person and second or third person in singular present tense. In the table, Dutch orthography is shown which reflects pronunciation to a large extent. However, the wordfinal /n/ of the inflectional suffix *-en* is no longer pronounced in many contexts.

There are two further regular inflectional forms: The present participle, which is used rather infrequently in constructions like (1.1a), and the past participle, which is used with auxiliary verbs to express the present perfect, as in (1.1b).

(1.1) a. *Hij zat balend achter zijn computer.*
He sat fretting behind his computer
b. *Hij heeft gebaald.*
He has fretted

Table 1.1: Dutch finite tense inflections for *gijpen (to jibe)* and *sturen (to steer)*

	Stemfinal consonant		Meaning of
Pronoun	voiceless	voiced	the inflection[a]
ik	gijp	stuur	1SG present
jij/hij/zij	gijpt	stuurt	2SG, 3SG present
wij/jullie/zij	gijpen	sturen	PL present
ik/jij/hij/zij	gijpte	stuurde	SG past
wij/jullie/zij	gijpten	stuurden	PL past

[a] SG=singular, PL=plural, and 2SG is shorthand for 2nd person singular.

Notice that almost all inflections are suffixes, the past participle *ge-* is the only exception.

Dutch nouns are inflected only for number. Two different affixes exist, *-en* and *-s*, see (1.2a-b). The choice between the two is mainy made on the basis of metrical properties of the word. Note that the *-en* suffix marks plural for both verbs and a large portion of the nouns. Nouns come in two genders, neutral and non-neutral.

(1.2)	a. *peer*	*peren*	pear/pears
	b. *appel*	*appels*	apple/apples
	c. *een groene boot*	*de groene boot*	a/the green boat
	d. *groene boten*	*de groene boten*	ø/the green boats
	e. *een groen schip*	*het groene schip*	a/the green ship
	f. *groene schepen*	*de groene schepen*	ø/the green ships

Dutch adjectives are inflected for a combination of gender, definiteness and number. Adjectives always bear an inflectional *-e* suffix, unless they agree with a singular, indefinite, and neutral noun (cases (c-f), no affix in (e)).

Derivation in Dutch is, by nature, not easily captured in diagrams like those given above for inflection. Similar to English, not all predictable combinations of stems and affixes exist, and neither do they always carry the predictable meaning. Booij and van Santen (1998) give an overview of the derivational suffixes of Dutch and present the perplexing example of the *–ist* suffix, which can combine with a noun to describe a person but the meaning of the combination seems completely free. Compare *violist (someone who plays the violin), behaviorist (someone who belongs to behaviorism),* and *arabist (someone who studies Arabic).* Similar to the inflections, there is an overwhelming preference for suffixing morphology (see also Cutler, Hawkins, & Gilligan, 1985), with some prefixing derivations for verbs.

One derivational suffix is of special importance because it will be used in an experiment later on: The diminutive affix *–(t)je*. This affix can productively combine with every noun of Dutch, and applies as easily with newly borrowed words of foreign origin:

(1.3) *zeiltje* (little sail), *slootje* (little brook), *stormpje* (little storm), *vipje* (little VIP), *computertje* (little computer), *Stalintje* (little Stalin)

There has been some debate over the status of Dutch diminutives, because they resemble inflections to some extent. Most morphologists nowadays agree that they are derivational (Booij & van Santen, 1998), for various sound reasons. First, there is a possibility for meaning specialisation, which is not possible for inflectional forms: Take for example *een knietje geven (lit. to give a knee* DIM, *the act of giving someone a knee in the groin)*, and *briefje (note, lit. letter* DIM*)*. Next, diminutives can combine with adjectives and are class-changing in that case, like *groentje (greenhorn, fresher)* from *groen (green)* and *blondje (a blonde girl)* from *blond*. Diminutives can also incidentally combine with words of other classes, as in *praatje (talk or lecture)*, from the verb *praten (to talk)*, and *uitje (trip)*, from the preposition *uit (out)*. The presence of fully predictable diminutives notwithstanding, such cases indicate that the diminutive formation process of Dutch is derivational.

1.3 The split morphology hypothesis

I have so far discussed inflection and derivation in theory (Scalise, 1988) and in practice (the Dutch examples) and I have shown that the two have differing properties. The split morphology hypothesis, put forward by Perlmutter (1988), takes the distinction between inflection and derivation one step further and claims that the difference between these processes results from the place they occupy in the language system:

(1.4) The split morphology hypothesis (Perlmutter, 1988, p. 95):

a. Derivational morphology is in the lexicon.
b. Stems are listed in the lexicon. Consequently, suppletive stems are listed in the lexicon.
c. Irregular and closed-class inflected forms are listed in the lexicon. Consequently, suppletive inflected forms are listed in the lexicon.
d. Regular, productive inflection is extralexical.[1]

The hypothesis rests on two assumptions: First, regular inflectional processes are treated differently from irregular processes; only the latter are allowed inside the lexicon. Irregular forms and suppletive stems have to be

[1]The term 'extralexical' is introduced to refer to processes that occur after lexical look-up, because the more obvious 'post-lexical' is already in use in a different meaning (Kiparsky, 1982).

stored and this separates them from the otherwise highly predictable regular inflectional forms. Clause (1b) and (1c) are included to handle the commonplace observation that inflection exhibits both extreme regularity (past tense expression by suffixation with *-ed*) and, on the other extreme, suppletion (as in *go–went*).

The second assumption regards a definition of inflection and derivation. Perlmutter adopts Anderson's definition that "inflection is what is relevant to syntax" (cf. Anderson, 1982, p. 587). The phrasing of this has lead to confusion in the field, however, because several things are relevant to syntax that are not part of morphology proper, and are thus not intended to be covered by Anderson's 1982 definition. In his 1992 book, Anderson has consequently rephrased the definition of inflection (in a regrettably less pungent way) to be that part of morphology that forms an exception to the Lexicalist Hypothesis. This is formulated more positively in (1.5):

(1.5) Inflection is that part of morphology that is accessible to and/or manipulated by the rules of syntax (after Anderson, 1992, p. 83).

This definition does not agree with all previous uses of the inflection–derivation distinction. Many characteristics of the two processes have been proposed (see Scalise, 1988, for an overview) and they do not at all converge. The major alternative criterion is that of *generality*, which is a measure of how generally applicable the affix is within the set of stems it can formally apply to. Related measures are productivity and regularity. All inflectional forms are highly general, and generality varies for derivations: The suffix in *Roman+esque*, for example, is less general than that in *vague+ly* because very few nouns combine with *-esque* while many adjectives combine with *-ly*.

Many of the problems with the split morphology hypothesis that were raised in the literature are based on conflicts between the generality criterion and (1.5). Anderson has denounced high generality as a touchstone for inflection: "A high degree of productivity thus does not seem to be either a necessary or a sufficient criterion for calling a morphological category 'inflectional' " (Anderson, 1992, p. 78). I will return to the merits of the generality criterion below.

1.4 Theoretical problems

Gradual criteria

The split morphology hypothesis requires a contrastive definition of inflection versus derivation (as was given above in (1.5)), and this is a point of discussion with the defendants of a more gradual distinction between inflection and derivation, like Bybee (1985, 1995), Kuryłowicz (1964), and Plank

(1994). These authors have argued that there is only *prototypical* inflection and derivation, and the two cannot be separated.

Bybee (1985) has proposed that whether a certain property of the language (eg. number, aspect) is expressed by an inflectional or a derivational affix, is determined by two gradual aspects: Its *generality* and its *relevance*. The latter is a measure of the meaning impact that the affix has on its base. These two aspects are working against each other, high generality correlates with low relevance, and vice versa. Bybee gives several theoretical reasons for this antagonism and a highly negative correlation between the two phenomena was found in a sample of 50 languages she examined. This survey largely confirmed her thesis that inflectional processes are those that combine moderate generality with moderate relevance. More relevant properties become derivational, while more general ones are expressed by clitics or other 'free grammatical processes'.

Bybee further widened the choice of expression by means of inflection and derivation by introducing a more far-reaching scale of *modes of expression*, which also contains lexical forms, 'free grammatical items' (clitics, amongst other things), and syntactic expression. This is a fruitful approach when considering how linguistic categories as various as number, thematic roles, and social status are expressed across languages, and how languages change their preferred mode of expression over time. From Bybee's perspective, a grammaticalisation process changing a word into an affix no longer results in a sudden breach of domain, but instead slowly pushes a linguistic category to one side of the dimension (see Bybee, 1995, for considerations of this kind).

An interesting example that shows the strengths and the weaknesses of Bybee's proposal can be found in Dutch: There is a productive process of making noun–verb compounds that function as verbs in the language. Below, three examples of these compounds are shown that have differing characteristics: The past tense can be formed by inflecting the compound as a whole (1.6a), by separating the noun and inflecting the remaining verb (1.7b), or by a periphrastic syntactic expression using *was aan het* (literally, *was busy doing*).

(1.6) *tafeltennissen*, 'to play table tennis'
- a. *Trinko tafeltenniste de sterren van de hemel.*
- b. **Trinko tenniste tafel (de sterren van de hemel).*
- c. *Trinko was de sterren van de hemel aan het tafeltennissen.*
  Trinko played table tennis like an angle.

(1.7) *brandstichten*, 'to set something on fire'
- a. **Rinus brandstichtte (in) een beroemd gebouw.*
- b. *Rinus stichtte brand in een beroemd gebouw.*
- c. *Rinus was die avond aan het brandstichten in een beroemd gebouw.*
  Rinus set fire to the famous building.

(1.8) *ballonvaren*, 'travel by balloon'

a. **Het mannenkoor ballonvaarde die dag.*
b. **Het mannenkoor voer die dag ballon.*
c. *Het mannenkoor was die dag aan het ballonvaren.*
The choir travelled by balloon that day.

The examples nicely demonstrate three distinct positions on the scale of modes of expression: Past tense is expressed either by inflection, by inflection and a 'free grammatical item', or by syntactic expression. Properties of the individual words may be taken as predictors of which mode of expression is chosen for the expression of tense. In a more traditional analysis, only differences in the derivational history of these examples can cause different ways of past tense formation. This approach is taken by Booij and van Santen (1998), who discuss similar examples and propose that *ballonvaren* is verbal reinterpretation of a noun–noun compound with a nominalised verb as the right hand constituent:

(1.9) $[[\text{ballon}_N+[\text{varen}_V]_N]_N]_V$

Booij and van Santen (1998) give some reasons why this verbal reinterpretation is restricted to the infinitival form and the present participle *ballonvarende*. Whether or not one buys this reasoning, it remains unexplained why *ballonvaren* must be analysed this way, and *tafeltennissen* differently.[2] An explanation that uses word specific properties, like frequency or *lexical strength* (Bybee, 1995) to assign a certain mode of expression to a word seems more satisfying in this respect, but no proposal in this direction has been made so far. It is also unclear why language chooses between either inflectional expresssion (1.6a) or inflectional plus 'free grammatical' expression (1.7b), while the *aan het* construct is always possible.

However useful the modes of expression are in analysing the example given above, the claim that there is a gradual scale of modes of expression can be contested. The fact that variable, non-dichotomous forces influence the choice of expression does not imply that the modes of expressions themselves form a continuum. On the contrary, discrete modes of expression might explain the slow rate at which language phenomena are changing. An example Bybee (1995) mentions, is the diachronic move that the Old English noun *lice (body)* made from a noun which frequently appeared as a part in compounds to a generally applicable derivational affix *-ly*. Changing mode of expression does not happen on a large scale and this particular move took hundreds of years. This reluctance to change might be an effect of the effort and perseverance needed to drag a whole language community over the discrete hurdle of changing a mode of expression.

[2]Even more so since the authors suggest that forms may 'promote' over time and acquire a full inflectional paradigm. Their example is *paardrijden (horse riding)*, which nowadays allows for *paardreed*.

Generality of inflectional processes

A problem with the split morphology hypothesis that has been noted by Booij (1993, 1996) concerns the fact that some inflectional processes are not without exception and irregularity. Some alleged inflectional forms have to be stored in the lexicon because they are not predictably related to their stems, like *clothes*. Also, inflectional paradigms are incomplete for nouns that occur only in singular or plural form *(scissors)*, and these exceptions to the inflectional rules have to be stored somewhere (as *scissor* [–lexical insertion] in the terminology of Halle, 1973).

There is, however, no problem with this for an interpretation of split morphology that allows for suppletion and irregularity in the lexicon, as in (1.4). The fact that regular inflection is outside of the lexicon does not preclude forms that resemble inflections to appear in the lexicon. Under this approach, the lexicon can contain *clothes* as the form associated to the meaning 'garments'. This form *clothes* has to be a monolithic form that differs from the plural inflection of *cloth* in the way it is generated. This point is illustrated by Dutch *letteren (arts)*, which has kept the old plural affix, while the form that once was its singular, *letter (letter, character)*, has acquired the currently regular pluralisation suffix: *letters*.

Further examples given by Booij (1993) involve participles that are not related to any synchronically occurring verb, as shown in (1.10). Booij considers these to be examples of inflectional processes that have lost their generality, but an analysis along the lines proposed for *clothes* above is also possible. This approach removes these problematic forms from the inflectional paradigm.

(1.10) *bezeten* mad
gesmeerd smoothy
ontzettend very
woedend angry

Of special interest in this respect are the countless participles that have taken on an irregular and lexicalised meaning, while still having a homonymic sister with regular meaning:

(1.11) *gezond*, lit. sunbathed, non-literal meaning healthy
Ik heb de hele dag gezond.
I have the whole day sunbathed

gehecht, lit. stitched (together), non-literal meaning attached
De wond moest gehecht worden.
the wound had.to stitched be

laaiend, lit. burning, non-literal meaning angry[3]
Hij is omgekomen in de laaiende vlammen.
He has passed.away in the burning flames.

The non-literal meaning is the current and most frequent one for all these participles. Nonetheless, the example sentences are perfectly normal Dutch and the appropriate reading of the word will be understood without problem. A rather ad-hoc way of explaining this is to say that these inflectional forms have acquired an irregular meaning, which can be cancelled by the context. Alternatively, this can be considered a case of homonymy between an inflectional form with regular meaning, $gezond_a$ *(sun bathed)*, and a lexicalised form with irregular meaning, $gezond_b$ *(healthy)*. This analysis predicts that the lexicalised form is free to move away from its inflectional origin, and has the possibility to become an isolated form like the ones in (1.10).

Class changing inflection

A traditional observation is that only derivation may change word class. Henceforth, I will call this the class change criterion for distinguishing inflection and derivation. The class change criterion was disputed by Haspelmath (1996), who analysed participle formation in German (and seven other phenomena) as examples of class changing inflection. Consider the principle example:

(1.12) *der im Wald laut* $sing_V$*-*$ende_{Adj}$ *Wanderer*
the in.the forest loud singing.PART hiker
the hiker (who is) singing loudly in the forest
(adapted from Haspelmath, 1996, p. 44)

The participle bears an ending that agrees with the gender, number and case of the noun *Wanderer*, like all other German adjectives, and there is little doubt that a class change has indeed occurred. The question remains whether participles are formed by a inflectional process. Haspelmath argues that this is the case by virtue of the generality criterion: Participles can be formed of all verbs by a regular, productive process and are 'thus' inflectional.

According to Anderson's criterion, this use of the present participle is surely derivational because the presence of the participle form is not required by syntax, nor has it repercussions for any syntactic process or agreement relation. Even though participles like *singend* exhibit a high degree of generality, they are not inflectional because generality is neither a necessary nor a sufficient condition for calling a process inflectional (Anderson, 1992).[4]

[3]It is remarkable how many present participles of Dutch now have a meaning similar to *angry* (Booij, 1993, gives another two). Many of these have been created from verbs that are no longer in current use, which is why this example has a different structure from the previous ones.

[4]Haspelmath is of the opinion that participles are inflectional according to Anderson's definition, because they change word class. Word class, it is reasoned, is surely relevant to syntax.

Example (1.13) below contains Dutch fragments that clearly demonstrate that the *-end* suffix is lexically determined when it is used adjectivally. In this position, the suffix is not governed by or relevant to syntax. In (1.13a-b), the participle *zingende* can be replaced by any adjective that bears the correct noun agreement ending *-e*. Compare this to the verbal use of a participle in (1.13c), where it can only be replaced by other past participles because it is part of a construction which expresses tense. Substitution by the adjectival form is not possible, as shown in (1.13d):

(1.13) a. De zingende wandelaar
'the singing hiker'
b. De vrolijke wandelaar
'the happy hiker'
c. Al zingend struikelde de wandelaar.
'While singing, the hiker stumbled.'
d. *Al vrolijk struikelde de wandelaar.
'While happy, the hiker stumbled.'

The upshot of this discussion is that the generality criterion is sometimes contradictory to other criteria for the inflection–derivation distinction. Conflicts with Anderson's definition in (1.5) and with the class change criterion occur in the case of participles described above and also for many highly frequent affixes like English *-ly*. The point of discussion then is whether generality or a more syntactic criterion like (1.5) must be used to define inflection and derivation. There are many advantages to adopting (1.5): First, there is no need to create a new category of class-changing inflection. Second, Anderson's criterion is easier to evaluate than generality, because it gives either a positive or a negative outcome, and not a scalar measure that has not been quantified yet. Third, high generality can be considered a consequence of the fact that regular inflection is driven by syntax, and thus required. No lexical gaps are allowed here, which is very different from the situation for derivation.

A remaining problem is how the lexicon can contain a form like *singende* that resembles an inflectional form in appearance. One way is to propose storage of all participle forms in the lexicon, in a fashion similar to *clothes*. The drawback of this account is that massive duplication of forms occurs, whereby all participles are stored in full in the lexicon, and can be derived by inflectional rule. This situation appears certainly unaesthetic to the linguistically trained mind, which always seeks for economy of storage. As Henderson (1989) has argued and as has become apparent from the experiments by Baayen, Dijkstra, and Schreuder (1997), the language system does not favour economy of storage above all other things. Massive duplication of

---

This confusion rests on a misinterpretation of Anderson's (1982) definition of inflection, that was alluded to above. Haspelmath goes even further and claims that a word class change can be a sufficient condition for inflection. This would reassign many cases that have always been considered derivational to inflection, and it is hard to see the benefit of this move.

forms is one possibility, an alternative account will be considered in Chapter 6.

Inflection inside derivation and compounding

It has been noted before that inflected forms do sometimes occur in compounds. Kiparsky (1982) has proposed that this process is limited to irregularly inflected words, but this claim does not hold. Booij (1996) lists many examples from Dutch that are all similar in structure to the following:

(1.14) *student- en- huis*
student PL house
student house

Although this type of formation is definitely widespread and can be extended at will (*metselaarsverbond*), it is restricted to a certain class of heads that seem to promote a plural interpretation of the left-hand constituent. Even within this type of head, a plural nonhead is not applied consistently. In a survey of about 250 compounds in the CELEX lexical database of Dutch (Baayen, Piepenbrock, & van Rijn, 1993) that consisted of two parts of which the second was *huis*, I counted 36 cases in which a plural form was used (*botenhuis*, 'shiphouse'), 76 cases in which a singular form was used (*jachthuis*, 'hunting lodge', and 11 cases in which a singularis tantum was used (*melkhuis*, 'milk house'). The remaining cases were compounds formed from non-nominal first parts, like adjectives or verbs. I have been very liberal with assigning words to this category, and it contains a number of cases that others might want to classify as containing a singular noun (eg. *danshuis*, 'dance house', which can be either formed of the nominal or verbal meaning of *dance*). Under this conservative approach, the majority of N–N compounds was formed with a singular nonhead, but the class of plural nonheads is substantial. I have not been able to extract any regularity in the assignment of singular or plural nonheads, and there is some process at work here that is not fully understood yet.

One attempt at explaining these forms is to say that compounds are lexicalised phrases. There are many compounds for which this is clearly the case, consider *hit-and-run*, *the language-as-a-fixed-effect fallacy*, and Dutch compounds with possibility for in-built agreement:

(1.15) a. *blije mensen feest*
happy.INFL humans party
b. *holle bomen bos*
hollow.INFL trees forest

The majority of Dutch compounds is formed from two nouns, however, and this leaves open the question why N–N juxtaposition is allowed in the phrases that form the basis for these compounds, whereas N–N juxtaposition is normally not found in Dutch.[5]

The conclusion of this is that inflections do appear in compounds, but we are in uncharted waters where it comes to the place of compounding relative to derivation and to the processes that are used for compounding. The *hit-and-run* type compounds shows that far more is allowed here than elsewhere in morphology. This undermines any implications of the compounding examples for the inflection–derivation distinction.

More problematic for the split morphology hypothesis are those cases in which low-frequent regular inflections, which have no independent reason for being stored in the lexicon, appear as the basis for derivation. Booij (1996) gives three derivational suffixes that may take inflections as their base:

(1.16)	a.	*held-*	*en-*	*dom*
		hero	PL	hood
		'heroism'		
	b.	*boek-*	*en-*	*achtig*
		book	PL	like
		'like books'		
	c.	*weten-*	*schap*	
		to.know	NOMINALISER	
		'science'		

Booij notes that all three derivational suffixes developed from lexemes. The question is whether these derivational forms have roots back to compound formation which involved a second part that was still lexical, that is, whether *heldendom* is the modern-day form of a word that entered the language as a compound created from the plural *helden* and the noun *dom*. This would explain these derivational cases to the extent that they have entered the language under the guise of compounds, and have been reanalysed as derivations since. But this is certainly not tenable for all the examples Booij gives, as some must be relatively recent (*leerlingdom*).

The pattern of pluralisation of the nonhead is not consistent, even with these three suffixes combinations exist that take a singular first part. Pairs that are almost minimal are:

(1.17)	a.	*grootvorst-*	*en-*	*dom*	vs.	*groothertog-*	*dom*
		great.king	PL	hood		great.duke	hood
	b.	*pap-*	*en-*	*dom*	vs.	*priester-*	*dom*
		papist	PL	hood		priest	hood
	c.	*engel-*	*en-*	*dom*	vs.	*mens-*	*dom*
		angel	PL	hood		man	hood

[5]On the other hand, one might want to claim that all N–N juxtaposition has previously been considered compound formation.

I can provide no good explanation for the words formed from these three suffixes under the split morphology hypothesis. The examples can not be explained by proposing a 'leak' from the inflectional component to the lexicon. The fact that only plural forms occur as non-heads in derivations, and only before specific suffixes, suggest that there is some regularity here that cannot be captured by allowing any inflectional form to be stored in the lexicon.

Further data are certainly necessary to solve this issue. Of extreme interest are productivity measures of these derivations (the existence of *draagmoederschap (surrogate motherhood)* suggests that some are productive), and especially how productive they are with plural first constituents. If productivity is attested, the issue of why combinations with and without plural first constituents are found can be approached experimentally. If not, linguistic scrutiny over the existing cases has to provide us with an answer.

1.5 Psycholinguistic perspectives

The Levelt/Roelofs-model

In this section, I will discuss the merits of the split morphology hypothesis from a cognitive viewpoint. The Levelt/Roelofs-model of language production will be used as a framework. The details of this model are discussed in Chapter 2, here we are only concerned with some underlying principles to the model, specifically the assumption of a *lemma level.* A sketch of the model is given in Figure 1.1.

Most of the components of the Levelt/Roelofs-model are the same as or similar to what is assumed in other psycholinguistic models and in linguistic thinking about language production. An exception to this is the presence of representations called *lemmas* between grammatical encoding and morphological encoding. Lemmas are abstract lexical entries that form the intermediate level in the connection from a conceptual specification to a word form. The lemma itself contains all the syntactically relevant properties of the word, but it does not contain its form specification. Instead, it points to the morphemes that express this lemma. In an early version of the theory (Levelt, 1989), the lemma also contained the semantics of the word. In more recent proposals (eg. Roelofs, 1992a; Levelt, Roelofs, & Meyer, 1999), the meaning of a lemma is represented by its connection to a lexical concept.

The reasons for assuming an intermediate level between conceptual and form entries are manyfold. Most compelling are the empirical results, ranging from speech error data to carefully designed experiments with evoked brain potentials (Kempen & Huijbers, 1983; Jescheniak & Levelt, 1994; Garrett, 1988; Dell, 1986; Van Turennout, Hagoort, & Brown, 1997). Lexical concepts (semantics) are associated with lemmas and not with morphemes

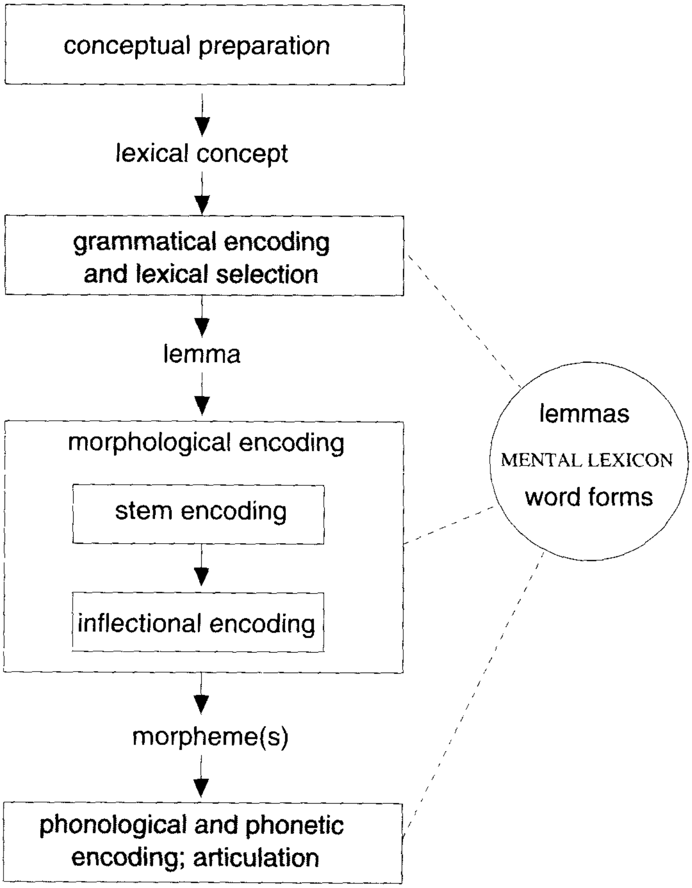


Figure 1.1: The Levelt/Roelofs-model with the morphological extensions proposed here. Shaded boxes summarise multiple levels in the model.

in the model, and in effect the Levelt/Roelofs-model implements a word-based morphology (Aronoff, 1976).

Modularity

What are the psycholinguistic advantages of accepting the split morphology hypothesis? Most psychological theories, like linguistic theories, distinguish between various components in the language production process. Components are thought to be modular; they should interact only through interface representations. Separating inflection from derivation brings us closer to this goal: If inflection is separated from derivation, if syntax can interact with inflection only, and if the creation of new words is specific to derivational morphology, the syntactic component is effectively isolated from the word formation component.

The influence of syntax on inflection is also tightly constrained in the Levelt/Roelofs-model: For one, syntax has access only to the lemmas, and no properties of the word form (phonology) can play a role at the syntactic level. The lemma contains all formal properties, and these can be freely accessed by the syntax. Gender, for example, is read off the lemma by the agreement rules. To store accidental properties of the lemma, which are valid only in the current context of uttering, the lemma contains *diacritical parameters* or *diacritics.* To continue the example, the gender specification of the noun is copied to the diacritic for *gender* of the agreeing adjective in a Dutch NP. Diacritics also store specifications that originate at the conceptual level, like the specification *singular* or *plural* for nouns. In sum, the interface between syntax and inflection comprises two parts: The formal properties of the lemma, which can be read only, and the diacritics, which are given a value by syntax or semantics.

Diacritics are in some respects comparable to the morphosyntactic representations (MSRs), proposed by Anderson (Anderson, 1988, 1992). The MSR functions as the terminal node of the grammar, and it contains the morphological specifications that are operated on by syntax. No formal distinction is made within the MSR between the fixed properties of the word and those that can be changed by syntax. The lexicon contains a list of pairwise combinations of an MSR and a stem. After retrieval of the stem, the features represented at the MSR trigger the application of extralexical, inflectional rules.

In the Levelt/Roelofs-model, diacritics reside on a lemma. The lemma points to one or more morphemes and the stem is created by concatenating these morphemes. After stem generation, inflectional processes are invoked by the content of the diacritics. Lemmas have a meaning by virtue of the fact that every lexical concept points to a lemma that expresses this concept.[6] In Anderson's generative approach the semantic specification of the

[6]Some concepts point to more than one lemma, examples are the English prefix verbs like

lexical entries is left to some unspecified interpretation component: A state of affairs that is not fully satisfying from a psycholinguistic perspective.

In contrast to the MSR, the set of diacritical parameters form an unstructured list of values. Anderson has argued that this cannot be right for languages where verbs agree with both subject and object: Two specifications for number, gender, etc. are necessary and he proposes a system to keep the two apart. Because the Levelt/Roelofs-model was developed for Dutch and English, structured diacritics have not been included in the Levelt/Roelofs-model yet, but it is clear that some modifications in this direction will have to be made to account for other languages than English and Dutch.

Weakening the split morphology hypothesis

There are some reasons to weaken the interpretation of the split morphology hypothesis, which follow from work by Schreuder, Baayen, and colleagues (Baayen, Burani, & Schreuder, 1997; Baayen et al., 1997; Schreuder & Baayen, 1997). Working on language comprehension, these authors have attested that frequently encountered words may become stored as a whole. In their model (Schreuder & Baayen, 1995), the language input is simultaneously parsed by two separate routes: A lexical route, which consults the lexicon for a stored representation, and a parsing route, which tries to disassemble the word into its parts. The last route is always successful for morphologically complex words, but takes a considerable amount of time and effort. The lexical route is faster, but cannot be used for every single word of the language or we would face massive storage of perfectly predictable forms (the problems connected with massive storage are further discussed in the context of Bybee's model, on page 37).

The model of language production can be equipped with a similar device: Words that are used often enough may be stored in the lexicon as a whole. Because they appear unanalysed, they are in all respect similar to suppletive and irregular forms. Utterance of a pre-stored past tense form (say, *sailed*) bypasses regular tense inflection because the tense diacritic was already cleared when accessing the stored past tense form (cf. the treatment of irregular forms in Anderson, 1992, and in Chapter 4).

Data on the inflection–derivation distinction

Evidence for the psychological reality of the inflection–derivation distinction has been found by many researchers. A good overview of the experimental literature is given by McQueen and Cutler (1998). To demonstrate how wide a range of paradigms, experiments, and observations converge on the fact that inflection is treated differently from derivation, I will mention some of these studies here.

---

to speak up, and the French negation construct *ne... pas*.

First of all, there is evidence from speech error corpora. It has been observed by Garrett (1980) that so-called stranding errors mostly involve stems and inflectional morphemes, while stranding in derivations occurred rarely. Next, a number of aphasic patients have been reported, for whom derivational and inflectional forms were differentially affected (Badecker & Caramazza, 1989; Laine, Niemi, Koivuselkä-Sallinen, & Hyönä, 1995; Niemi, Laine, & Tuominen, 1994). For one patient, Miceli and Caramazza (1988) found that no less than 97% of the morphological errors he made in a word repetition task were of inflectional nature.

Further evidence has been obtained from lexical decision with normal subjects. In a lexical decision task, subjects are asked to give a word/non-word decision to a stimulus. In the relevant experiments, the stimulus is preceded by a morphologically related word and the influence of this *prime* on the decision time for the stimulus is measured. Schriefers, Friederici, and Graetz (1992) found that inflectional primes made subjects consistently faster to react to other inflectional stimuli. They used German adjective forms with either *-es* inflection (for gender) or null inflection for this part of the experiment: *kleines (small)* as a prime facilitated reaction to *klein (small)*, and vice versa. No reliable effects were obtained for derivational words priming each other, as in *Wahrheit–wahrlich* (*truth, truthful*). Similarly, there was no effect of a derivation priming an inflection (*wahres–wahrlich*, 'true–truthful') or of priming the other way around (*wahrlich–wahres*). This argues for a special status of the inflectional form.

A similar task was employed by Laudanna, Badecker, and Caramazza (1992), with the twist that homographic Italian stems were used to construct the primes: For example, the effect of an inflectional prime created from *mutare (to change)* was compared to the effect of a derivational prime created from the same verb. In both cases, the stimulus was the semantically and morphologically unrelated, but overlapping word *mute (mute)*. It turned out that the effects of the two primes were radically different: only inflectionally related words affected the reaction times to *mute*.

Data on the regular–irregular distinction

Central to the split morphology hypothesis is the distinction between regular and irregular inflectional morphology: Many cases of suppletion exist that can clearly not be solved in an extralexical component (cf. the relevant clauses of Perlmutter's definition, (1.4b) and (1.4c)). Experimental data showing that this distinction is psychologically valid has been collected in the cause of the psycholinguistic debate on connectionism versus symbolism. The issue of this debate is whether past tense formation is based on symbolic processing of rules and exceptions, or can be captured by a connectionist model that produces both the regular and the irregular forms on the basis of analogy. The debate was initiated by the appearance of the connectionist model of Rumelhart and McClelland (1986), which has been subsequently improved and extended upon by Daugherty and Seidenberg (1994), Elman et al. (1996), MacWhinney and Leinbach (1991) and others. Both the workings of these connectionist models and the past tense acquisition data that they were based on have been challenged by defendants of the symbolist view, for instance Pinker and Prince (1988) and Marcus et al. (1992).

The empirical evidence against one device that produces both regular and irregular inflectional forms comes from various sources. In a brain imaging study on normal subjects, Jaeger et al. (1996) found different regions activated when regular and irregular verbs were produced. Similar results were obtained in an ERP study on correctly and incorrectly inflected regular and irregular German verbs by Penke et al. (1997): A specific frontotemporal negativity was found for regular verbs that were incorrectly affixed with the irregular *-en* affix. Overgeneralisation (applying the regular *-t* to irregular stems) did not give rise to this negativity. Specific impairment of regular or irregular forms was found in aphasic patients (Marslen-Wilson & Tyler, 1997) and for children with William's syndrome (Clahsen & Almazán, 1998). Data from language development were obtained by Bybee and Slobin (1982) indicating that irregular forms were indeed stored in the lexicon, although some generalisation of these forms still occurred in the form of *schemas*. Schemas, however, are very different in nature and application from regular rules.

Experiments reported by Kim, Pinker, Prince, and Prasada (1991) showed that subjects are sensitive to the regularities of the language when they had to judge novel verbs for acceptability. The regularity tested was that denominal verbs never have an irregular past tense. New complex verbs were formed from irregular verbs and prepositions. When the new verb was introduced in a nominal context, subjects preferred a regularised past tense, and when it was introduced in a verbal context, they preferred the irregular past tense of the underlying verb (double question marks precede the form that was disfavoured by the subjects):

(1.18) a. I have had so many light beers.
I'm totally *lighted-out* / ??*lit-out*.
b. She had been trying to light up her face.
She was totally *lit-out* / ??*lighted-out*.

This indicates that subjects are indeed sensitive to the rule that denominal verbs are always regular, while the same is not true for new verbs created from a verb and a preposition. The question now is what the underlying reason for this rule is. Under split morphology, the answer is straightforward: The irregular verb is represented by a lemma that points to the allomorphs of the stem (*light* and *lit*). When the verb is nominalised, a new lemma is created because the nominal form has other inherent properties (its word class is different, for one) and other possible diacritical parameters. This new lemma points at only one morpheme, because allomorphy does not occur for the noun. When this noun is subsequently turned into a verb again, the allomorphy and hence the irregularity cannot be retrieved from the verb it originated from.

Data on the split morphology hypothesis

In Chapters 4 and 5 of this thesis, experiments will be reported that directly bear on the issue of split morphology. The two underlying assumptions were confirmed: First, evidence was found for a lexical treatment of irregular forms in terms of allomorphs, as was claimed by (1.4b-c). Second, inflection and derivation behaved differently in two experiments that were otherwise very similar. The split morphology hypothesis itself was confirmed by data indicating that regular inflectional processing does indeed occur after derivational morphology. Specifically, the derivational complexity of forms did not influence the representation I coined the *inflectional frame*. At this level of representation, only inflectional suffixes matter. Such a representation will be of prime use to an extralexical process of regular inflection, which is separated from the derivational rules in the lexicon.

Word-based morphology

The theory developed by Anderson and the Levelt/Roelofs-model are both word-based, which means that the morpheme is not considered the smallest meaningful element (Aronoff, 1976). The term 'word-based' is an historical accident and not a proper characterisation of these theories: It is the stem (the word minus inflection) that is considered the smallest element that carries meaning. No attempt is made to systematically derive the meaning of complex stems from their parts, and this solves some problems that have haunted morphological theories for a long time. Aronoff's examples against semantic decomposition up to the morpheme involves Latinate prefixes like *re-*, that are notoriously hard to assign meaning to when considering their appearance in *remit, resume, receive, reduce, etc.* Further problems for traditional theories were the bound morphemes that occur in 'cranberry' compounds, meaningless morphs (eg. the theme vowels of Greek), and the general ill-treatment of morphemes in fusional languages.

Word-based morphology is incorporated in the Levelt/Roelofs-model by the intermediate position of lemmas between the conceptual system (semantics) and the morphological encoding component (which accesses the form lexicon). A compound is produced by a lemma pointing to more than one morpheme, as in the STARBOARD pointing to the morphemes *star* and *board*. All pointers are numbered to create the necessary sequential ordering of morphemes and to prevent **boardstar*. A similar reasoning holds for derivational morphology: *readable* is produced by activation of the morphemes *read* and *able* from the appropriate lemma. The meaning of the combination is stored by the link between the conceptual level and the lemma READABLE. The compound or derivation does not have to be transparent in meaning: The same system is applied to forms that are semantically opaque combinations of two morphemes. Experimental evidence for the fact that morphemes play the same role in opaque compounds as in transparent compounds was reported by Monsell (1985), Roelofs (1997a), Roelofs, Baayen, and van den Brink (submitted). The models of Dell (1986) and Chialant and Caramazza (1995) also contain a word-based approach to morphology.

The notion of a *morpheme*, as used in the context of the Levelt/Roelofs-model, is a slightly extended one: Morphemes are the building blocks of word form, and there is no set minimum on the number of word types the morpheme must occur in. This allows representation of the word *capacity* as consisting of the (meaningless) morphemes *capac* /kəpɑs/ and *-ity*. No harm is done by the fact that *capas* is confined to this particular form and its close relative *capacious*. There need also not be a synchronic morphological overlap with obviously related words such as *capable* and *capability*, which would involving the morpheme *cape*. Of course, an analysis that can capture all four forms as instances of one morpheme plus the application of some general phonological rule is in principle preferable, but this seems no feasible alternative in this case and many others.

In his discussion of compounds, Anderson (1992, Chapter 11) is much more reluctant to allow morphs with a very limited distribution the status of a separate entry in the form lexicon. He argues that the semi-productive prefix *Sino-* in *Sino-Japanese trade treaty*, *Sino-Tibetan war*, and *Sinology* has no independent status and new forms with *Sino-* are created by analogy to the old ones. However, this move requires the creation of a new, analogical word formation device, which has no application beyond this process, because all other words are stored in their maximally decomposed form. Second, the form lexicon has to be burdened with many unanalysed forms (*Sino-Tibetan* will be one entry), which are given half-baked morphological complexity by the fact that the analogical device can isolate parts from them. There are many more examples of this type than the rather exotic examples that Anderson gives: The readers of this thesis all have the bound morpheme *morpho* in their lexicon, which appears in *morphosyntactic*, *morphophonological* and, to speculate a little, in *polymorphous* and *a-morphous*.

1.6 Conclusion

The theoretical problems with the split morphology hypothesis that have been raised in the literature seem to boil down to conflicting opinions on what the criteria for inflection and derivation are. By adopting Anderson's definition and rejecting the generality criterion, many of the problems disappear. Another group of problematic cases involves compounds, and it was argued that these cases do not have direct consequences for inflection and derivation. What remains are two 'level ordering' problems, where inflection seems to feed derivation. One case is formed by the three Dutch affixes *-schap*, *-dom* and *-achtig*, that can all attach to plural nouns. The other case concerns participles (and similar phenomena), that are undoubtedly inflectional in one context, but undoubtedly derivational in another. The form identity of these two appearances cannot straightforwardly be explained under the split morphology hypothesis.

Next, the split morphology hypothesis was approached from a psycholinguistic perspective, and it was shown that the hypothesis helps to maintain modularity in the language production system. The notion of a *lemma* was discussed, and the way it is used to invoke the psycholinguistic version of word-based morphology. Empirical data from a number of sources were reviewed, and these data speak for a strict distinction between inflection and derivation, for a difference in the production of regular and irregular inflectional forms, and for the separation of inflection and derivation in language production.

The split morphology hypothesis deserves more positive attention than it has been given in linguistics so far. The hypothesis agrees with the results of many experiments on morphology, and it is favoured by a major processing model of language production. The linguistic problems that have been raised are not strong enough to reject the thesis, and with some modifications to the theory many can be accounted for.

Psycholinguistics: data and models

CHAPTER 2

In this chapter, several psycholinguistic aspects of the production of morphology will be dealt with. The first issue taken up concerns the speech error data and the evidence that they supply for the existence of certain morphological units in language. Together with these data, proposals on how speech errors should be accounted for will be reviewed. The following three sections are devoted to models of language production and comprehension. First, the analogical models are discussed. These are based on the premise that there are no rules in the language, just analogies of differing strength. Next, dual-route models are treated that contain a built-in distinction between rules and exceptions. The last model explained is the Levelt/Roelofs-model, which will function as the theoretical framework for the following chapters. One final aspect of psycholinguistics that is dealt with are the experimental findings on the role of morphology in language production.

2.1 Speech errors

Much of the early research on language production was done by the collection and careful description of spontaneous errors in speaking.[1] In this section, the data from naturally occurring errors involving morphemes or morpheme-like elements will be reviewed. To explain the occurrence of some errors and the absence of others, some processing models of speech production were put forward. After discussing these, I will briefly review some more data on (mostly elicited) speech errors that involve malfunctioning of agreement processes.

Speech errors and morphology

From a mere glance at a speech error corpus, it is immediately obvious that most errors involve phonological units. There are, however, at least two

[1] There are some fundamental problems connected with this methodology, see Cutler (1982) and Meyer (1992).

types of errors that involve morphemes. Here are some examples of the first type:

(2.1) a. I thought the truck was parked → I thought the park was trucked (Garrett, 1980)
b. I cooked a roast → I roasted a cook (Fromkin, 1973)
c. Fancy getting your nose remodeled → Fancy getting your model renosed (Garrett, 1980)

These errors are called *stranding exchanges*, because some portions of the words exchange but the rest does not take part in the exchange and is stranded. In (a), the two stems *truck* and *park* have changed position, but the past tense inflection stayed behind. Example (b) provides further insight into what the exchanging representations look like: Again, the past tense morpheme *-ed* stayed in place while the stem morphemes *cook* and *roast* were exchanged. The reported pronunciation of *roasted* as /rəʊstɪd/ confirms that the correct level of analysis is morphological. If not morphemes but segments had been exchanged, there is no reason to expect that the *-ed* suffix, which is realized as /t/ at the end of *cooked*, would accommodate to its new verbal context and surfaces as /ɪd/. Next to this, the stems *cook* and *roast* must have been phonologically encoded before the suffix *-ed*, for the closing consonant of the stem is needed for the expression of the past tense suffix as /t/ or /ɪd/. Example (c) shows that these stranding errors may create new words like *renosed*. Although most stranding errors involve only inflectional affixes (Garrett, 1980; Stemberger, 1985),[2] this example contains the derivational prefix *re-* that is also stranded. The involvement of derivational morphemes and other morpheme-like elements *(semi-morphs)* in speech errors is problematic for most theories of morphology in language production. This issue will be skirted in this chapter, and treated in more depth in Chapter 6.

In the second type of morphological error, affixes themselves are moved. This is shown in the following examples:

(2.2) a. Even the best teams lost → Even the best team losts (Garrett, 1988)
b. What that adds up to → What that add ups to (Garrett, 1988)
c. Lots of moisture and protection → Lots of moisture and protecture (Shattuck-Hufnagel, 1979)

In (a), the inflectional ending of /ti:mz/ shows up at the end of the verb, which is realised as /lɔsts/. Notice how the *-s* was voiced in its original context, but not so after it was erroneously combined with *lost*. These error have been called *shifts*, and according to Garrett (1988), they adhere to

[2] Stemberger reports that inflectional affixes were stranded in 89% of the errors "where it [stranding] would have been possible" (Stemberger, 1985, p. 162). According to Garrett, 64% of the stranding errors in the MIT corpus involve only inflectional morphemes, and another 23% involve a derivational and an inflectional morpheme (Garrett, 1980, p. 198). It should be noted that Garrett's definition of morphemes is rather loose from a linguistic viewpoint: It includes semi-morphs like in Example (2.6b).

different constraints than the stranding exchanges in (2.1). One such constraint is that only shifts displace inflectional endings and "minor category words". This can be seen in (b), where the inflectional ending is attached to a preposition. Prepositions rarely occur as the landing site in exchange errors. Other differences between the two types of errors include the fact that stranding exchanges preserve stress (or, in other words, stress is stranded too) while shifts typically distort stress (Cutler, 1980). Next, exchanges can occur between phrases, and preserve word category if they do. Shifts show no tendency for preserving word category in any context. Example (c) shows that affixes do not always exchange place in shift errors, but can also be persevered. Here, the *-ure* suffix of *moisture* replaced the suffix of *protection.*

For both stranding exchanges and shifts, the unit that is displaced is often syllable sized. This raises the question whether we are really dealing with morphological exchanges, or whether these errors can also be syllable exchanges. Most of the previous examples contained an extra clue towards the morphological status of the units by token of the sound alternation, most clearly shown in /t/ changing to /ɪd/ in (2.1b). Generally, however, speech error theories are reluctant to allow syllables the status of an exchangable unit. "With respect to speech error patterns, syllables appear to constrain error rather than indulge in it [...] Movement errors do not generally involve syllable-sized chunks, except where the latter are ambiguous as to their classification (ie. they coincide with morphemes, or the segmental make-up of the error unit is ambiguous)." (from Garrett, 1988, p. 82)

Models for speech errors

From various cases of this sort, it can be inferred that stems and inflectional suffixes constitute separate units of representation at some level of the language production process (Garrett, 1988; Levelt, 1989; MacKay, 1979). Shattuck-Hufnagel (1979) has proposed a mechanism for language production that allows for a distinction between types of units at various levels. For the morphological level, this mechanism would state that there is a set of target units, which contains both stems and inflectional affixes. Next, there is a frame that contains slots labelled for a particular type of unit. A *scan-copier* works on these representations: For each slot in the frame it scans the set of units for one that is appropriate for the slot. The scan-copier always matches the type of the unit with that of the slot (this is called *the unit similarity constraint*), and this prevents stems to appear in inflectional slots and vice versa. When the intended unit is found, it is copied to the slot and a separate *checkoff monitor* marks the unit as being used in this frame. Speech errors are, in Shattuck-Hufnagel's view, the result of the rare failure of the scan-copier or the checkoff monitor. Specifically, the scan-copier can select the wrong unit from the set (resulting in an exchange as in (2.1a)), and the checkoff monitor can mark a unit prematurely or too late (resulting in repeated use of the same unit (2.2c)).

Garrett (1980, 1982, 1988) has proposed a planning frame in which the open class words (nouns and verbs) are inserted, while "all non-lexical formatives [this includes inflectional affixes, *dpj*] are features of planning frames, and hence not subject to exchange processes" (Garrett, 1988, p. 76). Shattuck-Hufnagel's two types of slots for stems and inflections have effectively been replaced by a difference in the way stems and inflections are generated. When the features of the planning frames are converted to segments, it is possible (though rare) that the segments end up in the wrong place, leading to the shift errors reported in (2.2). These errors arise at the latest form encoding stage in Garrett's view, and cannot interfere with the encoding of open class items that occurs much earlier.

A problem with this latter proposal is that there is a linguistically necessary dependency between the expression of the open class words and that of the inflectional suffixes for irregular forms. The encoding of a strong (irregular) past tense, eg. *swim*, cannot be done by first expressing the verb, then expressing the affix *-ed* and finally combining the two to form *swam*. Only by accessing the lexicon with a combination of a specification of the verb and an abstract specification of time, one can arrive at the correct allomorphic stem (cf. Levelt, 1989 and the treatment of irregulars in this thesis). In Garrett's (1988) model, an abstract (functional) specification of the utterance is made. But the retrieval of the planning frame with built-in inflectional endings seems independent of the retrieval of stems from the mental lexicon, and this leaves open the question how *swam* is generated.

In this thesis, I will propose a frame representation that has something of both Shattuck-Hufnagel's and Garrett's hypotheses. Stems and inflectional affixes constitute two different classes to account for the fact that they hardly ever exchange. But the morphological encoding process for stems takes place at an earlier stage than that for inflectional affixes, and there is thus also a processing difference between the two. The specification of the inflectional affix will not be done by features of the frame, but by features of the abstract word (open class item), and this allows for a straightforward treatment of irregular forms. Details of this proposal will be presented in Chapter 3 and Chapter 4.

In both Shattuck-Hufnagel's and Garrett's proposals, the elements that can exchange in an error must be, at some level of language production, competing for slots of the same frame. For "I thought the park was trucked" (2.1a) to appear, there must be a sentential frame with a slot for the object NP and one for the verb, such that the lexical content for these slots can be exchanged. Erroneous insertion of *park* in the NP slot leads to the checking off of this stem, and the subsequent insertion of the only remaining unit *truck* in the verbal slot.

Without invalidating this possibility, Levelt (1989) claims that another cause for errors is the parallel construction of similar parts of speech. The following example shows how this can be effected:

(2.3) The child gave the mother the **cat** → The child gave the cat the **mo**ther
(Levelt, 1989, p. 246, bold face marks the primary sentence accent)

According to Levelt, the NPs *the mother* and *the cat* are simultaneously being constructed at a certain moment during production of this sentence. When the lemmas are generated, they can get assigned to the wrong NP. There is some error checking done by the NP, but this only regards the word class of the inserted lemma, which should be nominal. *Cat* is a noun, and there are no formal objections to its insertion in the indirect object NP. Sentence accent is a property of the direct object NP and the subsequently incorrectly inserted lemma *mother* accommodates this property without further ado.

Allocation to the wrong sentence frame slot can also be the fate of more abstract sentence properties. Consider the following German example, in which the diacritic that triggers the subjunctive inflection of *könnte* is anticipated and erroneously applied to *haben*:

(2.4) *Es ist so schade, daß wir kein Bett dort haben, daß ich könnte ...* →
Es ist so schade, daß wir kein Bett dort hätten, daß ich könnte ...
It's such a shame that we don't have a bed over there, that I can ...
(From Meringer's corpus, cited in MacKay, 1979)

An explanation of this error in terms of exchanging material of one single frame is rather contrived because it forces us to postulate a frame that contains a main clause and two subordinate clauses. Given the freedom that language users have to add subordinate clauses at will, an analysis in terms of three separate but ordered sentential frames (one for the main clause, two for the subordinate clauses) is more convincing. The error can then be explained by a spurious assignment of the subjunctive diacritic to the subordinate verb *haben*.

The frames that will be proposed in this thesis are relatively small: They encode only for one stem and its affixes. Because the frame distinguishes stems from affixes by having typed slots, no exchange between these two classes can occur. When multiple frames are constructed in parallel, the inflectional frame can still play a role in the explanation of speech errors. Consider (2.5):

(2.5) a. people read the backs of boxes → people read the backses of boxes (Shattuck-Hufnagel, 1979)
b. if that was done to me → if I was done to that (Fay, 1980)

The abstract specification for plural that goes with the word *box* is, in keeping with the stem it was to attach to, expressed as /ɪz/ in (a). This form (allomorph) of the suffix was retrieved from the lexicon and inserted into the affix slot of the *box* frame. The form *backs* that was constructed at the same time, was also soliciting for a plural suffix, and accidentally received /ɪz/ too. Other errors involve mixing up intermediate representations, like the error in (b). To encode the stem, a pointer is followed into the form lexicon.

To explain (b), I assume that the frames and the case specifications stayed in place, while the pointers were mistaken. This will lead to the observed insertion of *that* in object position and to a correctly case-marked form for *I* in subject position. This explanation is tentative, the error can also be accounted for by a within-frame exchange in a sentence frame.

Levelt's proposal of *crosstalk* between frames constructed in parallel allows us to make use of smaller frames, while maintaining the possibility for errors that span several words. Levelt's approach has certain other benefits. Garrett (1988) remarked that the level of syntactic processing that is needed in his theory must be multiphrasal, because exchanges can cross phrase boundaries. A two-clause limit is suggested to account for the fact that exchanges usually involve adjacent phrases. This constraint would be rather *ad hoc* in Garrett's theory, but can fall out naturally from a parallel processing units account: It would be a rare thing for a speaker to already plan a clause far ahead from the one currently under construction.

Problematic cases still remain, whatever framework is chosen. These involve the exchange of part of a derivational form with an inflected stem, as in (2.6a), or, worse, the exchange of an inflected stem with something that has no morphological status in this word at all (2.6b). A similar problem occurs in (c), where an affix is attached to a preposition, a word class that normally never bears affixes:

(2.6) a. they were talking Turkish → they were turking talkish (Garrett, 1988)
b. I have got a load of chicken cooked → ... cooken chicked (Garrett, 1980)
c. I'd forgotten about that → I'd forgot abouten that (Garrett, 1980)

The occurrence of these types of errors suggests that at some level of language production, arbitrary morphemes and morpheme-like entities can be exchanged. The first proposal that will be made in this thesis holds that inflectional and derivational morphemes are treated separately, and there is no role for non-morphemic entities like *-en* in (b). This makes the errors in (2.6) hard to explain. In Chapter 6, I will try to bring together word-based morphology, which allows us to introduce 'meaningless' morphemes, and the speech error evidence for *semi-morphs* like the abovementioned *-en* in *chicken*.

In summary, the natural speech error evidence indicates that morphemes must constitute a unit at some level in the language production process. Inflectional endings appear to be treated separately, and encoded later, than stems. The models developed by Shattuck-Hufnagel (1979) and Garrett (1980, 1988) contain a frame that holds both types of units and that can explain why only certain types of errors are made. Levelt (1989) has proposed that, apart from possible errors in the assignment of fillers (lexical material) to the proper slot of a frame, fillers can also end up in another frame, if frames are constructed in parallel.

Agreement errors

A whole body of experimental work has been devoted to cases in which agreement is failing. An example of such an agreement error is:

(2.7) Grammatical findings indicate that the grammatical *number* of English verbs *are* determined by agreement with the noun. (from Allerton, 1992, cited in Eberhard, 1997; italics added)

In this error, the plural number of the adposition *English verbs* interfered with the number of the subject NP *number*, and the correct agreement of the verb *to be* with the subject NP was tampered with. Studies by Bock and Miller (1991), Eberhard (1997), Vigliocco and Nicol (1998), and others have revealed under which circumstances this error predominantly occurs. Bock and Miller (1991) found that the errors are not influenced by the animacy of the local noun (*English verbs* in the above example), or by the distance between this noun and the verb. Eberhard (1997) found that explicitly marking the number of the subject by using *one* instead of *the*, decreased the error proportion, while marking the local noun with *one* increased the likelihood of errors. In both papers, it was found that plural local nouns were far more likely to induce errors than singular local nouns.

According to Vigliocco, Hartsuiker, Jarema, and Kolk (1996), number marking on the verb in Dutch, French, and Italian is not only determined by the grammatical number of the subject, the conceptual number of the subject concept can also play a role: Fewer erroneous plural verbs were found for (2.8a) than (2.8b).

(2.8) a. The owner of the suitcases.
b. The collar of the coats.
(Translated materials from Vigliocco, Hartsuiker, Jarema, & Kolk, 1996)

The explanation for this effect is that the people interpret *the owner* in (a) as a singular person owning more than one object, while in (b) they presume multiple copies of *the collar*, each on every coat. When making an agreement error, the speaker seems to be in two minds about whether to have the verbal number diacritic agree with the number specification of the subject NP or with the number specification at the concept classification node. This presupposes the availability of the subject (agent) concept when encoding the verb, and timing factors that influence the likelihood of these errors have been proposed.

Bock and Miller (1991) and Vigliocco, Butterworth, and Garrett (1996) have been unable to find an effect of conceptual number on English agreement errors, where these errors have been found for French, Spanish, and Dutch. Vigliocco et al. (1996) have argued that this difference across languages is due to the fact that conceptual number is hardly ever expressed on English verbs: In the present tense, the *-s* encodes singular third person, all other forms are unmarked for number. The authors argue for the fact that subject–verb agreement in English has no *number meaning* and number marking on verbs therefore relies solely on the number specification on the syntactic subject and never on the number specification at the conceptual level. British English, however, does show an influence of the conceptual number on verbal inflection in the case of names for larger groups or companies, even if the name itself is a singular form, as in *Shell have withdrawn their plans to...*

Agreement also features in natural speech errors. It has been observed that the inflectional marking on words surrounding the error is sometimes adapted to the error. This is called *contextual accommodation* by Berg (1987), and is commonly found for inflections in English, but not for the choice of article in German:

(2.9) a. your tongue is all red → your teeth are all red (Stemberger, 1985)
b. *wie der Ort heißt* → *wie der Wort heißt*
how the.M village.M is.called → how the.M word.N is.called
(Berg, 1987)

In (a), *are* matches the plural number of the mistakenly selected concept TEETH and *to be* is thus said to be accommodated to the error. If the claim by Vigliocco et al. (1996) is correct, the incidence of this type of error for which the verb is accommodated should be smaller in Dutch and French, because for these languages there is a counteracting influence of the conceptual number of the intended concept TONGUE. For English, no conceptual influence on verbal inflection is assumed and accommodation can occur more easily.

In (b), a German noun of another gender than the original nouns was inserted. No accommodation of the article took place, for the masculine article *der* remained while *Wort* requires the neuter article *das.* Berg (1987) points out that gender is generally unpredictable from conceptual factors because most nouns do not have a natural gender. By implication, the form of the German article cannot be predicted on conceptual grounds. Number marking for English verbs, on the other hand, is predictable from the grammatical specification of number as well as from its conceptual specification. According to Berg, only a predictable process can accommodate itself to the insertion of an erroneous element.

Notice that the explanation of accommodations as given by Berg runs against the explanation for cross-linguistic differences in agreement errors reviewed above: Berg claims that number marking of English verbs is predictable from conceptual specification and that this facilitates the chance of accommodation taking place. Vigliocco et al. claim that the conceptual specification of subject number does not matter for English verbal inflection, because number is only marginally expressed on the verb.

With this unsolved problem, we leave the area of speech errors and move on to an overview of the morphological aspects of various models of language production and perception that have been proposed.

2.2 Analogical models of language production

Apart from some early generative approaches that tried to explain away irregularity by posing very specific phonological rules (Chomsky & Halle, 1968; Jackendoff, 1975), linguists have traditionally recognised the importance of both rules and exceptions in describing language phenomena. When applying this insight to psycholinguistic on-line models of language production, one straightforwardly arrives at positing a rule based device for regulars and a lookup procedure for irregulars. In several theoretical and empirical works, Clahsen and Pinker have argued at length for such a direct application of linguistic theories to psycholinguistics (see Clahsen, in press; Pinker, 1991). Below, the parsing model of Frauenfelder and Schreuder (1992) will be reviewed, which rests on similar assumptions.

There is, however, also a strong linguistic tradition of trying to find patterns in the irregular forms of a language. Many irregular forms derive from regular patterns, and obey a rule that has become extinct. Some have been influenced by a variety of historical processes that in effect obscure the relationship of the form to its historical rule (eg. *kabeljauw (codfish)* which is, according to De Vries, 1971, related to modern Spanish *bacalao*). Lacking a systematic relationship, analogy and similarity can be used to describe the relations between the irregular forms. This analogy has marked effects, however, that imply that there is more than descriptive force to it. First, newly created strong verbs are learned best when they are similar to existing strong verbs (Clahsen, in press). Second, some regular forms have become irregular over time but this happened only for verbs that shared consonants with a subregular pattern (Bybee & Slobin, 1982). Next, strong verbs that do not belong to any of the subregular patterns must be extremely frequent to survive the pressure to (sub)regularise (Hooper, 1976; Bybee & Slobin, 1982; Hare & Elman, 1995). Lastly, in a production experiment by Stemberger and MacWhinney (1986), regular verbs that resemble strong ones were attracted to the strong pattern and were produced in error more often than their controls.

Analogy also plays an independent role in word formation: Every now and then, new words like *workaholic* are created that, under a rule-like approach, seem to involve a reanalysis of an existing word into previously unattested parts, *alco+holic* or *alc+oholic*, and the application of one of the new parts to another stem. The process is usually a one-off, compare **computeroholic*, *?sailaholic*, but *chocoholic*.[3] Clearly, the explanation in terms of ad-hoc rules is indeed rather ad-hoc. Instead of this cumbersome and rather contrived analysis, a more natural approach seems to be that hearers are able to guess at the meaning of a new word by making analogies to existing words (Anshen & Aronoff, 1988; Bybee & Slobin, 1982). If this

[3] This last example is due to Anshen and Aronoff (1988).

analogy is strong and unambiguous, the resulting word may have a similar chance of being accepted into the language as a new word derived by rule. When an analogy is made often enough (both in types and tokens), a process with rule-like properties may emerge. Novelty foods are a common source for this and, with the support of expensive marketing campaigns, Dutch as acquired *-naise* for *type of mayonnaise,* has in *halvanaise, yoghonaise, olijvonaise.*

Past tense debate

Much of the discussion in the field of analogical modelling has focused on the problem of generating the regular and irregular past tense of English. With its mixture of rule-governed regular cases and a diverse set of irregular cases that can be grouped into subregular patterns, the English past tense system is an ideal testbed for any analogical model. This was realised by Rumelhart and McClelland (1986), who were among the first to propose an analogical solution to the problem (for which see below).

Their model learns the past tense by repeated exposure to it. Rumelhart and McClelland compared the learning curves of their model to those obtained in language acquisition research. Children, they claimed, show a typical *U-shaped learning curve* in the use of irregular forms. During the first stages of learning, all words are learned by rote and because many high frequency verbs are irregular, they appear in a substantial proportion of the child's utterances. In the second stage, the child acquires the regular *+ed* rule and overapplies this rule to previously learned irregulars. The result is that the proportion of correctly inflected irregular forms drops dramatically. In the third stage of learning, it is realised that there are exceptions to the rule and the irregulars regain their relatively frequent occurrence in the utterances (see Kuczaj, 1977; Bybee & Slobin, 1982, for more observations on the U-shaped curve and Marcus et al., 1992, for a critique on its interpretation by connectionists).

A third observation that has been frequently called upon is the generalisation of irregular patterns to new verbs that resemble irregulars ones. Connectionist models perform such generalisations readily. When subjects were asked to produce the past tense of nonce words or to judge the welformedness of nonce strong past tenses (MacKay, 1976; Kuczaj, 1978), generalisation of irregular patterns to new forms appeared quite acceptable. These experiments are related to the classic study by Berko (1958), who found that children prefer *glinged* as the past tense to *gling*, whereas adults are torn between *glinged* and *glang*. The adult data should be treated with care, however, because new irregular verbs are hardly ever introduced into the language. The willingness of subjects to produce or accept new strong past tense forms could well have been induced by the structure of the materials or the nature of the task.

The schema approach of Bybee

The resemblance between analogies and rules can be extended to the point where all rule-governed behaviour is a case of strong analogy. This approach is taken by Bybee (Bybee, 1985, 1988, 1995) and Skousen (Skousen, 1989; Derwing & Skousen, 1994). I will start with reviewing the first approach. In Bybee's model, all words and inflectional forms are listed in full. Semantic, morphological, and grammatical similarities are expressed by connections between individual segments of the word forms. These connections will relate the suffix *-ing* in *flying* to all other occurrences of *-ing* in participle position (*steering, rounding, etc*), but not to *ing* in *swing*. Partial and inconsistent relations can also be expressed. A hackneyed example is the putative sound symbolism in *slide, slip, slither*. When a new form is encountered, its relation to all existing forms is evaluated and a generalisation is made from the closest existing form.

From a psycholinguistic perspective, Bybee's model suffers from the problem that all phenomena in her model are gradual, and it is hard to make exact predictions from inspecting the model. As Schreuder and Baayen (1995) have noted, the model is also extremely hard to implement, which prevents us from using computational means to draw predictions from the model. The model freely mixes semantic, grammatical, and morphological information. This is considered a virtue by some, but psycholinguists tend to prefer a separation of these effects.

Lastly, all possible word forms are stored in full in the lexicon. The massive redundancy that results from this can be remedied by assuming that the redundancy that is captured by connections does not add to the size of the lexicon (similar to Jackendoff, 1975), but how this must be implemented, either by a model or by the human brain, is not obvious. Alternatively, one may want to argue that the capacity of the brain is large enough to simply list all forms without removing any redundancy. This requires, however, that the connections do not only indicate which forms resemble each other, but also that the connections reliably link to all occurrences of a morpheme in the word inventory. Only in this way, it can be assured that changes in, for example, the phonological form of a morpheme are carried over to all uses of that morpheme in the language. Whether a full-listing lexicon with exhaustive connections is preferable over a smaller lexicon with procedures (rules) to generate the remaining forms is mainly a question of aesthetics, the descriptive power of both alternatives is the same.

Skousen's analogical modelling approach

Skousen has laid out a general method to model analogy in his 1989 book. In Derwing and Skousen (1994), the model is applied to the English past tense to compare it with the connectionist approaches. The method used by Skousen's model entails that whenever the past tense form of particular verb is needed, this verb is looked up in a database of words (the lexicon). When a match is found, a present tense–past tense pair is retrieved from memory and the stored past tense is produced. When no match occurs, the *vicinity* of the verb is examined, looking for nearby examples. The nearest example is most likely to be taken as a guide, with the proviso that examples can strengthen each other (gang effect) and can be cancelled by intervening counterexamples.

There are at least two aspects in which Skousen's model differs from many others: It deals with analogy in a computational manner and, as a result, it renders a set of alternative outcomes of the analogy process with associated probabilities. Whereas connectionist models must revert to running each simulation many times to create a Monte-Carlo distribution of possible outcomes, Skousen's model directly reports the alternatives with their probabilities. Put more informally, when the connectionist result is "in 80% of the cases, the regular outcome was obtained," Skousen's model claims "this is regular, with 80% certainty." The second difference is that Skousen's model contains a lexicon that is in every sense the traditional store of words, meanings, and allomorphs and is thus more compatible with traditional linguistic approaches.

What is needed on top of a lexicon is a similarity mapping that gauges the proximity of items. The similarity mapping is the meat of the analogy model, and Derwing and Skousen (1994) discuss several possible similarity measures for their simulation of the past tense problem. The similarity mapping must be based on a multi-dimensional measure, because the traverse of the similarity space is done by relaxing one of the dimensions at a time. This is the Achilles' heel of the model: Given a psychologically valid similarity measure, the model can be used to predict behaviour in a particularly elegant manner, but finding a more or less correct measure, which captures the language data and accords with the formal constraint of multi-dimensionality, is not at all trivial and might turn out to be the principal problem in analogical modelling. As said, this problem is recognised by Derwing and Skousen (1994). Compared to other modelling approaches, Skousen's can be credited with a fair sense for its more arbitrary or problematic parts, but this does not seem to have enhanced the use of the model.

Connectionist models

A similar focus on analogy is entertained within the computational framework of connectionist models. The influential work of Rumelhart and McClelland (1986) has caused an upheaval in connectionist approaches to language production in general and strong past tense formation in particular (MacWhinney & Leinbach, 1991; Plunkett & Marchman, 1991; Daugherty & Seidenberg, 1994). All these models take a phonological representation of the present tense as input and convert it to a regular or strong past form. The strong points of connectionist models are their sensitivity to subregularities in the past tense system, and the fact that they acquire all their knowledge through learning. There are no rules in the model; the regular application of *-ed* is no more than the most frequent of the many mappings between present and past.

As with any new or re-emerging approach, the results of the connectionist models were at first taken very literally and were overenthousiastically applied to new domains. The past tense model tried to simulate the typical U-shaped pattern in the acquisition of strong past forms, but the way in which this was done has been severely criticised (Pinker & Prince, 1988; Ling & Marinov, 1993) and so have the data that the modelling was based on (Marcus et al., 1992). A fundamental problem concerns the similarity measure, which was also featured in the discussion of Skousen's model. The connectionist models build their own measure of what is related and what is not from patterns in the data, but the representation in which the input (present tense) and output (past tense) are presented to the model can highlight or hide these patterns. In the original Rumelhart and McClelland model, a complicated representation was used that consisted of sound-triplets and which probably made the model more sensitive to the ablaut patterns.

As was referred to above, several improved models have seen the light of day, which addressed the criticisms that have been levelled against connectionism in general and Rumelhart and McClelland (1986)'s model in particular. A persistent problem is that over a hundred years of linguistic thinking, experimental data (see Clahsen, in press, for an overview), speech error evidence (eg. Stemberger, 1985), and neuro-imaging techniques (eg. Jaeger et al., 1996; Penke et al., 1997) all point to a fundamental difference between regular and irregular patterns and the rule-like properties of the former. Although alternative networks have been devised that can handle *default* cases (Hare, Elman, & Daugherty, 1995), this discrepancy is still unsolved. It is also hard to envision how the connectionist simulations on past tense formation carry over to the domain of morphology at large and what the implications of this work for the encoding of inflection are. Connectionist models of word recognition often deny that morphology plays in independent role (Seidenberg, 1995), but this can clearly not be true for language production.

What remains is a surprisingly simple mechanism that can deal with analogy, make predictions, and store a great many associations in a nonsymbolic way. As was mentioned above, this sensitivity to analogy is not present in many psycholinguistic models because they are focused on deterministic behaviour (but see Skousen's model above and Ling & Marinov, 1993, for an alternative technique). The Levelt/Roelofs-model currently contains a spreading-activation network for modelling the conceptual, lemma, and word form system. Although interactive activation networks are not identical to connectionist models (they usually do not learn, and they use symbolic rules to augment the network's working), there is no principled reason why a form lexicon implemented as an interactive activation network could not exhibit associative properties similar to the connectionist nets. This way, the strengths of more traditional models could be combined with the natural ease with which connectionist models can explain the subregularities in the strong past tense system.

Stemberger's model

In his 1994 paper, Stemberger has proposed an unimplemented connectionist model that can deal with many morphological findings (see also Stemberger, 1985). Most interestingly, the model denies that there is anything morphological in the speaker's mind: All effects of morphology are explained as phonological effects and gang effects (see also Seidenberg, 1995). The term 'gang effect' describes a situation in which a group of weakly activated nodes is able to determine the end state of the model because they all emit similar patterns of activation and inhibition. Stemberger's model uses two types of gang effects: First, top-down gangs are caused by words that have similar meanings. Together they excite a pressure for certain sounds to be encoded. The regular past tense is formed this way: Three huge gangs compete for suffixing each word that has a past tense meaning with /ɪd/, /d/, or /t/. Phonological rules can cancel the /d/ and /t/ gangs in certain contexts.[4] Depending on cancellation and the respective strengths, one gang will win and affix the word accordingly.

The second type of gang is the bottom-up gang and this type is called in to help strong verbs to survive competition. The first pass through the network will render strong verbs correctly, but the regular gangs triggered by the past tense meaning component will kick in strongly enough to regularise the past tense form before the network has settled down. To counteract the regular top-down gangs, there are bottom-up gangs that group around similar sounds in the output. Irregular *sang* and *rang* will support the irregular pluralisation *drank* to resist regularisation to *drinked*. Bottom-up gangs thus encode the subregularities in the strong past tense system.

The strength of a bottom-up gang principally depends on the number of shared segments. This must be so to disarm competing bottom-up gangs centred around the regular endings /t/, /d/, and /ɪd/. Fortunately, the first two groups, which are the largest in number, also have the smallest segmental basis. But the importance of shared segments for the strength of the bottom-up gang is hard to bring to terms with findings by Stemberger and MacWhinney (1986): In a production experiment, subjects were asked to form the past tense of the verbs they were presented with. The materials included many weak verbs that resemble strong verbs and the result of the

[4]Stemberger's model is hybrid in that it contains nodes and rules operating on them. Because a gang is an emergent property of a group of nodes, it is not at all obvious how a rule that cancels a gang could be implemented.

experiment was that the verbs *flunk*, *snore*, and *screw* appeared relatively often without the required *-ed*. In these so-called non-marking errors, a past tense form is produced which is identical to the present tense (cf. *hit–hit*). The authors argue that the error rate in their experiment was a function of the resemblance of the present tense forms *flunk*, *snore*, and *screw* to past forms like *shrunk*, *wore*, and *blew*. In terms of the Stemberger model, this means that *blew*, *knew*, *grew* (and similar forms) exerted a substantial influence towards irregularity on *screw*. This occurs even though the segmental overlap with *screw* is as low as one or two segments and is thus comparable in size to the overlap of the regularising bottom-up gangs on /t/, /d/, and /ɪd/ with the regular past *screwed*. The balance between the effects of gang size and segmental overlap must be exactly right to ensure that the bottom-up gang around *blew* will sometimes outweigh the combined forces of the top-down gang pushing for regular suffixation with /d/ *and* the bottom-up gang of (regular) words ending in /d/.

A distinct problem with Stemberger's model is that it is not implemented. Although gang effects frequently occur in connectionist networks, their behaviour is hard to predict in practice. Also, it is stipulated that one gang should be stronger than another (cf. the discussion of *screw*), but whether the behaviour of a real network will obey this specification depends on factors like the size of the gang, the frequency of its members, the resemblance between the members, etc. Whether the language data agree with the constraints set forth by Stemberger's model has yet to be shown.

The interactive-activation model by Dell

Another, more influential model was proposed by Dell (1986, 1988). Dell's model is an spreading activation model in which nodes for syntax, morphology, syllables, phonemes, and features exchange activation. All levels of the model work together in producing a word, although the uppermost levels (concepts, syntax) are likely to be one word ahead when the final decisions on the phonetic level are made. There is full interactivity between adjacent levels, with activation flowing both upwards and downwards. Within each level, the most active node is selected and inserted into a frame. The frames are symbolic representations, similar to linguistic trees. Contrary to Stemberger's proposals, Dell's model is implemented and the model's predictions have been successfully compared to the results of an experimental study of sound slips (Dell, 1986).

The focus of the model is on phonology, not on morphology. Still, some properties are relevant for the discussion here and I will review them one by one. Dell discusses the in-between status of inflectional morphology, which is affected by syntax but operates at the morphological level. His conclusion is that (specifications for) inflectional suffixes are generated by the syntax, not by the conceptual system or the morphological system. This means that the representation for *some swimmers* comprises three syntactic nodes:

Some, Swimmer, and Plural. The syntactic node for plural is linked to the plural *-s* at the morphological level. Derivational morphemes come into play no earlier than at the morphological level. At the syntactic level, *swimmer* is an atom, while at the morphological level, it is represented by a stem (*swim*) and a derivational suffix (*-er*). Although Dell does not discuss this issue, his proposed representation allows semantically opaque derivations like *casualty* to be represented in the same way as transparent ones, such as *casualness*. This is compatible with the word-based approach to morphology that was argued for in the previous chapter.

All nodes at the morphological level are typed to ensure the correct operation of morphological rules. Derivational and inflectional affixes are of different types, which makes speech errors that confuse inflectional and derivational affixes very unlikely. According to Dell, this is borne out by the data, but note that Garrett (1980) has claimed that these affix types do mix in speech errors, with the constraint that two derivational affixes never seem to exchange (see (2.1c) and footnote 2). Prefixes, suffixes, and stems are also distinct categories in Dell's model and the proposal made in this thesis will feature a similar three-way distinction.

Of all models discussed in this section, Dell's is certainly the most interesting one. It is a general model of language production and has the distinctive advantage that it makes falsifiable predictions that can, and have been, tested. Regrettably, the model does not specialise in morphological issues. The combination between a spreading-activation network and rules that operate on this network to select and process the most active node has proven to be fruitful, and the same combination of mechanisms is encountered in the formalisation of the Levelt/Roelofs-model.

2.3 Dual-route models of language comprehension

In this section, some of the more influential models of language comprehension will be discussed. Although there are fundamental differences between the role of morphology in parsing and in producing language, some of the issues that are addressed by models of comprehension are relevant for our purposes. One issue concerns the way in which morphological complexity is reflected in the mental lexicon: Although the comprehension process may use different or additional form representations (or *access representations*), it is held by many that at least lexical concepts and lemmas are shared between the production and comprehension system (eg. Kempen, 1999; Levelt et al., 1999).

Another interesting property of dual-route models is that the existence of two independent routes for reaching the outcome (ie. the meaning of the word) leads to *statistical facilitation*: The average processing time for two routes together will be shorter than the average processing time for each of the individual routes.

The morphological race model (MRM)

Frauenfelder and Schreuder (1992) proposed a psycholinguistic model of language comprehension that contains a lexical lookup route and a concurrent parsing route. The two routes are called the direct route and the morphological parsing route. The direct route tries to look up the complete form in the lexicon. The lexicon is organised by token frequency, so frequent words will be easily looked up and assigned a meaning by the direct route. Less common words can also be assigned a meaning by the direct route (all existing word forms are listed in the lexicon), but this will take time.

It is therefore more likely that the competing morphological parsing route produces a meaning based on analysis of the word. The parsing route analyses the word into morphemes, and retrieves the meaning of these morphemes from a separate lexicon. The parsing route can only be successful if the word can indeed be analysed into morphemes and if the meaning of the word can be computed from combining the meaning of the parts. If the parse is successful, the *resting activation* of (the access representation for) the morphemes will be increased. This will facilitate further attempts to parse words that contain this morpheme, because the time it takes to parse a word depends on its morphological and semantic complexity, and on the activation levels of its parts.

The Morphological Race Model has interesting dynamic properties that ensure that high frequent words will end up being always recognised by the direct route, and that productive affixes will be easier to parse than unproductive ones. Frauenfelder and Schreuder devote ample discussion to the synchronic and diachronic changes that may occur in the lexicon. To name just two of these: Words may acquire a non-transparent meaning only if they are always recognised via the direct route, and morphemes may disappear from the language if they are seldom used, forcing the direct route recognition of all remaining words that contain that morpheme. The dynamic nature of the model allows for a larger degree of linguistic sophistication than was hitherto attained.

The meta model

A model with this slightly confusing name was proposed by Schreuder and Baayen (1995). The model builds on the MRM by Frauenfelder and Schreuder (1992), and contains a similar division of labour between a direct route and a parsing route. The major difference with the MRM is that both routes are now cooperating instead of working independently. To this extent, both routes are implemented in an interactive activation framework and work on the same network of representations. The meta model distinguishes access representations, intermediate representations (called *concept nodes* by Schreuder and Baayen), and syntactic and semantic representations. The end result of a parse is a combination of semantic and syntactic properties. For example, *proudness* is expressed by (amongst others) features <proud>, <abstractness>, <character-trait>, and <noun>. There are also (implicit) rules that ensure that matters are correctly adjusted to the new situation which occurs when two morphemes are combined. The suffix *-ness*, for instance, has a node that triggers the cancellation of the <adjective> feature of *proud*, to ensure that *proudness* will only bear the <noun> feature.

One of the problems with the MRM was the rather forced introduction of a feedback mechanism to dynamically account for frequency effects. The language data indicate that all high-frequent forms are recognised via a full-form representation. To explain this, the feedback mechanism must enhance the whole-word representations more than the morpheme representations. In the meta model, this problem is solved by postulating that feedback is the strongest to those representations that aided most in the recognition of the word. This appears to be a fair principle and it has the desired effects. When a morphologically complex word W is recognised, its whole-word representation is, by nature, connected to all semantic and syntactic nodes that express W. Each of the morphemes that is part of W is connected to only a subset of the syntactic and semantic nodes. Hence, the individual morphemes will receive less feedback than the whole-word representation.

If this mechanism of feedback is freely applied, whole-word representations for all inflectional variants of high frequent words would be constructed. Schreuder and Baayen feel that this would burden the lexicon too heavily, and while they assume separate access representations for all inflectional forms (as is customary for language comprehension models), no intermediate nodes (concepts) are created for those words that can be transparently derived from their constituents. Schreuder and Baayen formulate an explicit criterion on what can be transparently derived.

What is the status of the distinction between inflection and derivation in the models of Frauenfelder and Schreuder (1992) and Schreuder and Baayen (1995)? In essence, there is no such distinction. Whether a form is parsed or looked up depends solely on its frequency of use and its transparency, and not on its derivational or inflectional status. Of course, inflectional forms are transparent and will therefore be parsed if they are not too high frequent in the MRM. In the meta model, roughly the same holds. We have also seen that there is an explicit ban in the model on intermediate representations for transparent combinations. This precludes intermediate representations for inflected forms, because they are all transparent. Precisely the same constraint, however, holds for transparent derivational forms and there is no fundamental difference between the two in language comprehension (see McQueen & Cutler, 1998, for an overview of the experimental literature that confirms this stance).

The augmented addressed morphology model (AAM)

A host of studies on the role of morphology has been produced by Caramazza and co-workers, both on patients and on morphological processes in normal subjects. In two early papers (Caramazza, Laudanna, & Romani, 1988; Laudanna et al., 1992, see also Chialant & Caramazza, 1995), they have outlined a dual-route model of language comprehension that is particularly geared towards Italian. This means that it has built-in capabilities for extensive allomorphy, a rich inflectional morphology, and declensional classes. The model contains a direct route and a decomposed route, much like the dual-route models described earlier. The difference is that the AAM assumes that the direct route is faster than the decomposed route for all known words (in the 1989 formulation, words gained their own representation after a small number of exposures; in the 1995 version, low-frequent words of high-frequent constituents may remain without a whole-word representation). Only new words are recognised via the decomposed route.

A further property of the AAM is that complex relations are assumed between stems, allomorphic stems, and inflectional affixes. This structure is used to account for the data on nonword stimuli of various morphological structure, as examined by Laudanna et al. (1992). The default case is that a stem has facilitatory connections to all its inflectional suffixes. Some Italian verbs have a stem allomorph (minor stem) that is used instead of the regular stem (major stem) for a limited number of forms. An example is *correre (to run)*, which does not take the regular participle *corr+uto* but has a form created from the minor stem and a minor affix *cors+o*.[5] The minor stem *cors* is not connected to all inflectional suffixes, but only to those with which it may co-occur. The major stem *corr* is linked to all inflectional suffixes, but inhibitory links are assumed to those suffixes that are exempted by the minor stem, like *-uto*, to prevent overgeneralisation to occur. This pattern of links precisely describes the data obtained by Laudanna et al. (1992).

With respect to the distinction between inflection and derivation, the paper by Laudanna et al. (1992) contains results that point to a differing role of inflectional and derivational forms in priming targets that were orthographically related but morphologically unrelated to the prime. Hence, the authors proposed that the smallest element in the lexicon is the stem and not the root. This implies that derivational forms are stored in full in the mental lexicon, while inflectional forms are stored in a decomposed manner. Both types are fully listed in the access representations.

[5]Note that the minor stem is not irregular in the same sense as the English strong past tense forms, only 'predictable irregularity' was tested by Laudanna et al. (1992).

2.4 The Levelt/Roelofs-model of language production

We have already encountered the Levelt/Roelofs-model in the previous chapter, where it was introduced to serve as an example of a psycholinguistic model. The underlying principles of the theory are its parsimony, its modularity, and its feedforwardness. This last aspect implies that influence from lower level process on higher level processes is avoided where possible.

The top level in the model (see Figure 2.1) is formed by the conceptual stratum, where the message is created. This message is an abstract, prelinguistic representation of the sentence(s) that will be uttered. According to the model, the message must be specified in terms of *lexical concepts* only. Lexical concepts are those concepts that are expressible by a word in the target language, and this makes the message language-specific (Levelt, 1989). Roelofs (1997d) has argued at length on why, contrary to many other proposals, the conceptual stratum must contain a node for each lexical concept (eg. TABLE), and cannot represent lexical concepts in decomposed or distributed way, eg. as <furniture>, <has four legs>, ...This is relevant for our purposes in that conceptual number cannot be expressed by adding another feature, <±plural>. Instead, the conceptual stratum must contain two types of nodes: First, lexical concept nodes that point to a specific word and refer to an object, action or property in the world (eg. TABLE → *table*). Second, there are concept classification nodes, which only have an interpretation in conjunction with a lexical concept node, and which modify the extension (meaning) of that lexical node (eg. TABLE + PLURAL → *tables*). These nodes typically refer to linguistically salient aspects of objects and actions, and they express properties like number, event time, person, mood, or social distance.

How is an inflected word form generated in this framework? Because verbs hardly ever appear in isolation, I will consider the generation of the following sentence:

(2.10) The yachts rounded the buoy.

First, a preverbal message is constructed along the lines suggested by Levelt (1989). I will not go into the structure of these messages here, as it suffices to state that they contain the content of the sentence(s), and include one or more lexical concepts, here YACHT, TO ROUND, and BUOY. On the basis of a preverbal message, a syntactic structure is created. Generation of syntactic structures is lexically driven (Kempen & Hoenkamp, 1987) and the lexical concepts of the preverbal message will be used to guide production. In the above example, *to round* is the matrix verb. Its formal properties are needed to construct the utterance and these are retrieved by accessing the lemma for *to round* in the lexicon (lemmas contain an abstract word specification,

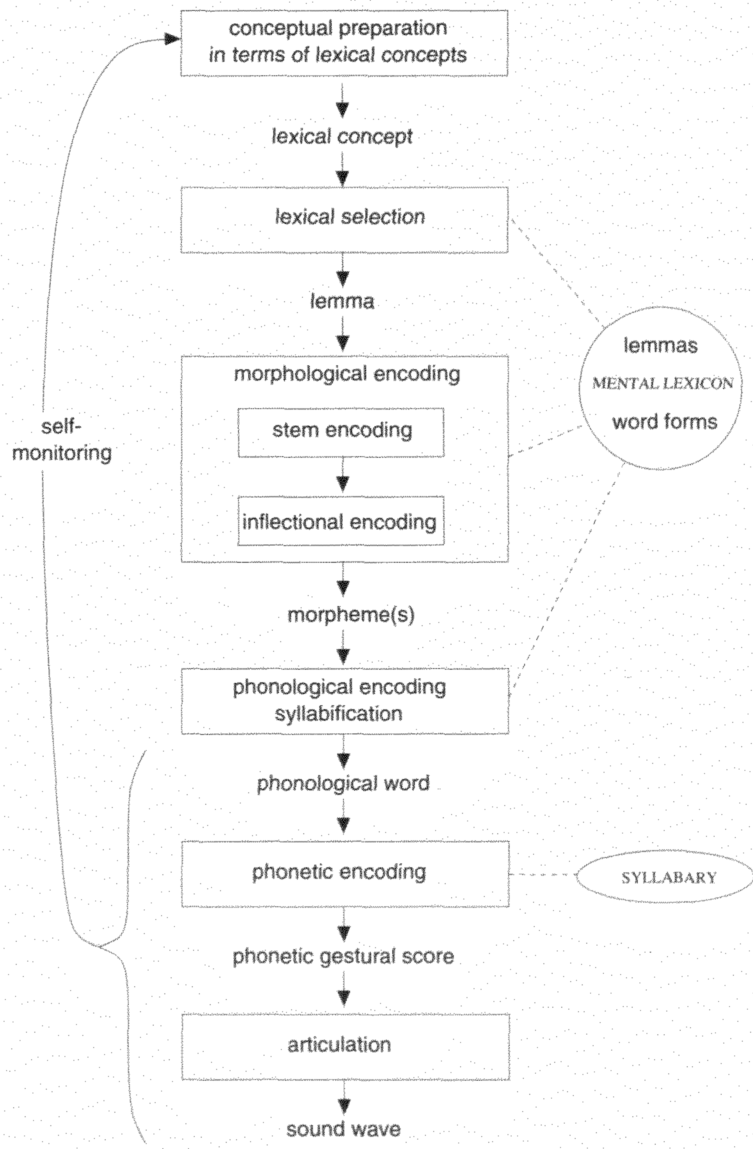


Figure 2.1: The Levelt/Roelofs-theory in outline (after Levelt, Roelofs, & Meyer, 1999)

I will return to them below). At the lemma level, the argument structure of TO ROUND is stored and it turns out to be a transitive verb, which requires a subject NP and an object NP. A syntactic structure will be generated that can host these required arguments, and the lemma for *to round* will be inserted into the V slot. Next, the two empty NPs will be filled by the lemmas that belong to the two other lexical concepts of the message.

Morphologically complex words

What results is a bare and unadorned utterance: The syntactic structure is already present, but it contains lemmas (abstract words), that are not yet inflected. One of the following steps is retrieving the word forms for the lemmas. Each lemma contains one or more pointers to morphemes in the form lexicon. More than one morpheme is needed for words that exhibit morphological complexity.

In the Levelt/Roelofs-model, three types of morphological complex words are distinguished, which differ in the way the complexity of the word comes about: Complexity may be principally due to more than one concept, more than one lemma, or more than one morpheme being activated (Figure 2.2). To start with conceptual complexity, Levelt et al. (1999) claim that whenever the speaker forms a new compound, like *paper pile*, this is done by simultaneously activating two concepts. Each concept will activate its lemma, and this in turn will activate its morphemes (Figure 2.2, Panel A). There is thus also more than one lemma active for this type of morphologically complex word, but this is not the principal source of complexity.

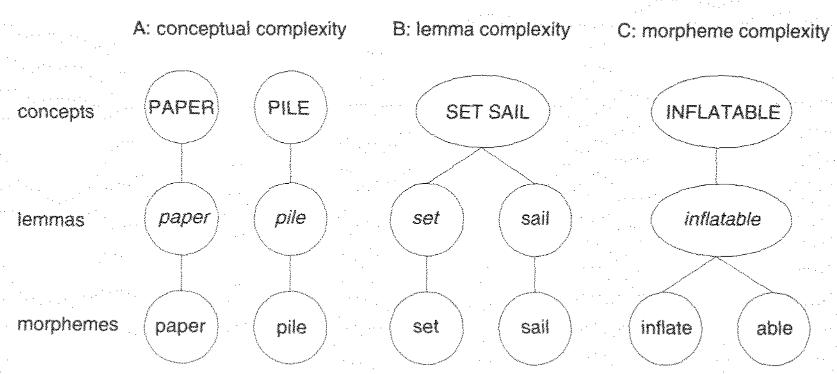


Figure 2.2: Three types of morphological complexity in the Levelt/Roelofs-model of speech production (after Levelt, Roelofs, & Meyer, 1999)

A well known example of the next type of morphological complexity, one lexical concept pointing to two lemmas (Figure 2.2, Panel B), is formed by complex verbs. Complex verbs like *to set sail* form a semantic entity, but are separable forms. This can be seen from the fact that inflectional suffixes appear between the two parts: "And our hero sets sail for another adventure." Other examples include particle verbs (*to run down*) and fixed adjective–noun combinations like Dutch *de zwarte doos (the black box)*.[6] For English, so-called exocentric compounds are good examples of a lexical concept pointing to two lemmas: *a passer-by; two passers-by*. Although marginal in English, this type of compounding is not rare at all. In French, for example, it is probably more frequent than any other: *planche à voile (windsurfing board)*.[7]

The third type of morphologically complex words is of the linguistically run-of-the-mill kind: One concept activates one lemma, which happens to be expressed by more than one morpheme (Panel C). This is the type of morphological complexity that will be dealt with in this thesis. It includes the familiar subtypes inflection, derivation, and compounding. As was indicated in the previous chapter, inflection is the most restricted of the three and occurs, according to the split morphology hypothesis, after lexical access (extralexical). Derivation and compounding, on the other hand, are inside the lexicon. In the Levelt/Roelofs-model, this means that the latter processes can create new lemma entries. Such lemmas will be expressed by two or more morphemes, like *inflatable* and *freeze-dry*.

There are two possible sources of compounds in this classification: multiple lemmas or multiple morphemes. This matches the finding that the process of compounding is more powerful than that of derivation, as was observed in Chapter 1. It is an open question whether the extra possibilities that compounding has above those available to derivation are all explainable in terms of the two possible representations that can underlie a compound.

After this digression on the complexity of words, the retrieval of a word form is straightforward. Following the pointers stored at the lemma level, the system enters the form lexicon and retrieves the morphemes that constitute this word. The Levelt/Roelofs-model uses tagged links in many places, and here the links into the form lexicon must be numbered to account for the serial order of morphemes. Once the morphemes are incrementally retrieved, they are simply concatenated and can serve as the input to further inflectional encoding. Word form retrieval is not always this simple, the mechanism must be slightly extended to deal with irregular forms, see below and Chapter 4. Note also that the specification of the word form is still

---

[6]The latter is a fixed combination and not a compound because the adjective *zwarte* is inflected. Compare *het zwarte boek (the black book, a noun phrase)* with *het zwartboek (the black book, in the non-literal sense)*. Anderson (1992) uses the term *composites* for similar compound-like phenomena.

[7]It should be noted that a new entry, the complex lemma node, has recently been introduced for the description of fixed expressions like *to bend over backwards* (Levelt, personal communication). This would slightly change the representation of *to set sail*, but would otherwise not affect the claims made here.

rather abstract at this level: Although the morphemes are made concrete, no phonological segments are present yet.

Agreement

Before a word form can be looked up, an abstract specification of the lemma's inflectional categories needs to be made. The lemma representation used in the Levelt/Roelofs-model contains *diacritics* to store the current value for each inflectional category, like plural or present. The diacritics store accidental properties of the lemma that depend on the context. The lemma also contains inherent properties that remain over time, like the word's syntactic class, the declension it belongs to, etc. Inflectional affixes are generated on the basis of diacritics and inherent properties.

There are four ways in which a diacritic can be given a value. The first and most common one is by agreement between two words of the utterance. Diacritic features and inherent lemma properties are copied from one word (the agreement source) to the other (the agreement target). The classical example of this is gender agreement for adjectives, which is instantiated by copying the noun's inherent gender property to the gender diacritic of the adjective. French adjectives agree both in gender and number with their nominal heads (*les filles heureuses*, 'the cheerful.FEM.PL girls') and this exemplifies that both inherent properties and diacritics of the agreement source can be copied to the agreement target.

The second type of agreement is similar, but here the agreement target and agreement source are conflated. This means that no copying is necessary: The information that will be used for the selection of the proper affixes is already present at the lemma. Latin noun lemmas, for example, must be inherently marked for declensional class to ensure selection of the proper inflectional affixes. Classifier systems probably also operate via inherent properties: The lemmas for words that are of the same classificational type (eg. that are classified as 'made of wood') bear the same inherent specification for class, which leads expression of that classifier at the morphological encoding stage.[8]

The third type of agreement concerns copying the information of a concept classification node at the conceptual level to the diacritics of the lemma at hand. Next to word choice enforced by the selection of certain lexical concepts, this is the only connection between the conceptual system and the language production system.

The fourth and last type of agreement occurs between the syntactic level and the diacritics of certain head lemmas. Although many seemingly syntactic influences on inflection can be solved lexically, others require a direct

[8]This presumes that classifier systems have no consistent semantic content and are, in fact, as arbitrary as gender systems. The only remaining difference between a gender system and a classifier system would be that gender is expressed on agreement targets, while classifiers are expressed on the word itself.

influence of syntax on the inflectional forms. An example of such an influence is the required omission of the Dutch second person singular affix -*t* if the order of subject and verb is reversed:

(2.11)	a.	*Jij*	*schrijft*	*een*	*boek*	*over*	*zeilen.*
		You	write.2SG	a	book	about	sailing
	b.	*Schrijf*	*jij*	*een*	*boek*	*over*	*zeilen?*
		write.2SG	you	a	book	about	sailing

The resulting verb form *schrijf* is identical to the first person, singular present tense.

The focus of this thesis will be on how the result of successful agreement is encoded by generating inflectional affixes. It was mentioned above that the specification of the diacritics has to be present before stem encoding can commence. The reason for this lies in the irregular words of the language. Sometimes, diacritics are not coded by regular affixation but by choosing a variant of the stem (an allomorph) that inherently expresses this diacritic. Irregular pluralisation of the *ox–oxen* type is an example of this. Allomorphs cannot be derived from the default stem by a simple operation, and must be listed in the form lexicon. This necessitates the presence of all diacritics before the stem is retrieved from the lexicon. A detailed mechanism for the retrieval of irregular forms will be given in Chapter 4.

Summary

The language production process has now gone through the following steps: First, a message was generated. The message was cast in terms of lexical concepts and concept classification nodes. Next, a syntactic structure was built by the grammatical encoder, with help from the argument structure of the matrix verb. The formal properties of the words in the sentence become available by accessing the lemmas that are linked to the selected lexical concepts. The diacritics of each lemma have to be given a value by agreement with other words or with higher levels in the system. The penultimate step is to retrieve one or more stem morphemes that express a lemma. If the lexicon specifies that this word is irregular, it lists the possible alternatives and the correct one is taken.

We are now in the position to generate a fully inflected form by combining the stem morphemes with those inflectional affixes that code for the specified diacritics. It is widely held that the encoding of regular inflectional affixes from diacritics is rule-governed (Clahsen, in press), and the rules for Dutch and other Germanic languages are extremely simple (see Chapter 1). Exactly how the combination of stem morphemes and inflectional affixes is made and how the system deals with irregular expression of inflection is not specified by the Levelt/Roelofs-model yet. In the following chapters, I will extend the theory to accommodate these two processes.

WEAVER++

Roelofs has implemented the Levelt/Roelofs-model (Roelofs, 1992b, 1997c; Levelt et al., 1999) under the name WEAVER, which stands for Word form Encoding by Activation and VERification. The implementation features a spreading-activation network of the type also encountered in the models of Collins and Loftus (1975) and Dell (1986). In addition to the network, there are production rules that verify whether the outcome of the activation spreading process is allowable. In contrast to many other models, WEAVER++ uses *binding-by-checking*: This means that the assignment of a unit to a slot is effectuated by verification, and not by selecting the most active candidate at a certain moment in time, so-called *binding-by-timing* (see Dell, Ferreira, & Bock, 1999, for a critique).

The first part of the simulation captures lemma retrieval. Activation flows between concept nodes and lemma nodes, and the model addresses some problems in the correct mapping from concepts to lemmas. The model is concerned only with lexical access in one-word utterances, and hence no simulation of diacritics and similar inflectional processes is done.

The second part of the implementation focuses on form encoding of words and syllables. In this domain, WEAVER++ does not simulate the workings of the Levelt/Roelofs-model in all its detail, but concentrates on the time it takes to proceed through the various stages of the model. In this respect, WEAVER++ is most akin to the models of response time that have been proposed for various perception tasks. In Roelofs (1997b, 1997a) and various other papers, it is demonstrated that WEAVER++ can model the reaction times that have empirically been observed, and can be used to predict effects that were not conceived of before (Levelt et al., 1999).

WEAVER++ contains a level of morphological representations, but does not address the issue of inflectional encoding yet. Concatenative morphology in the domain of compounding and derivation is handled by the above mentioned mechanism of a lemma pointing to more than one morpheme. These individual morphemes are converted into segments one by one, while the usual phonological processes of cross-morpheme and cross-word syllabification apply.

2.5 Empirical studies on inflectional encoding

A great many experiments pertaining to morphology have been run over recent decades, and—overstating things a bit—evidence has been found in favour of almost every conceivable theoretical stance. As Marslen-Wilson, Tyler, Waksler, and Older (1994) have noted, the contradictory results might be due to the fact that not all linguistic and psychological factors have always been accounted for. They mention the difference between representation and access, the language and modality of the materials, the exact morphological description of the effects at hand, and the semantic and phonological transparency as possibly disturbing factors for comparing the results of different experiments. I will not try to review the vast literature on the comprehension of morphologically complex words here, because this has been excellently done by Sandra (1990), and, more recently, by Marslen-Wilson and Tyler (1998) for the English past tense and by McQueen and Cutler (1998) for the status of morphology in comprehension in general.

With regard to the literature on language production, we have already encountered many of the crucial findings when discussing theoretical issues in this and the previous chapter. Here, I will mention only those findings that relate to the encoding of inflectional and derivational suffixes. The first source of empirical data on language production comes from speech error observations and, as was shown in the beginning of this chapter, this research has shown that stems and affixes are treated differently (MacKay, 1979; Shattuck-Hufnagel, 1979). Also, inflections seem to be bound more loosely to the stems than derivational affixes (Garrett, 1980).

A second source of evidence comes from research with patients. This has confirmed the previous claim that inflections and derivations are treated differently (Niemi et al., 1994; Goldberg & Obler, 1997). Miceli and Caramazza (1988) describe a patient who, next to an overwhelming bias towards inflectional errors over derivational errors, showed greater difficulty with the plural than with the singular of nouns, and similarly for the feminine and masculine gender of adjectives. This may point to a special status for the singular noun and the masculine adjective: In the AAM model, these forms occupy a central position and are therefore easiest to produce. Miceli and Caramazza (1988) also report that there was no preferred form of the inflected verb in their patient's utterances (cf. Haarman & Kolk, 1992). This runs contrary to what is found for language acquisition by children, where the (initial) preference for the present tense form is ubiquitous and suggests a central position for the present tense (Bybee & Slobin, 1982; Kuczaj, 1977). All observations in this vein of research are vulnerable to the critique that language contains marked and unmarked forms. Lacking the resources to find the correct form (patients) or the active knowledge of the correct form (children), subjects revert to the unmarked form because this is, by token of the markedness theory, less contradictory with the required form than any other.

The encoding of inflectional morphology is steered by information from the conceptual, grammatical, and lexical level. The information is assembled from its various sources and copied to the lemma of the to-be-inflected word by agreement processes. Agreement errors were elicited by Bock and Miller (1991), Eberhard (1997), Vigliocco et al. (1996) and others. The possibilities this methodology opens for influencing the inflectional encoding process have not been pursued yet.

An interesting point is that a strict interpretation of the theory of conflicting sources of number agreement would allow that, after having read the preamble *the owner of the suitcases*, subjects might sometimes produce a verb with both a singular affix (triggered by the lexical concept OWNER) and a plural affix (triggered by the word *suitcases*). To my knowledge, doubly inflected forms hardly ever occur in these experiments. Of course, this claim cannot be tested for English because it lacks overt plural number inflection. The experiment by Vigliocco et al. (1996), however, used Dutch and induced the required type of competition between the third person singular ending *-t* and the plural ending *-en*. No double marking errors seem to have occurred (R. Hartsuiker, personal communication)[9] and this argues for a strong constraint on conflicting affixes. The inflectional frame proposed here implements such a strong constraint.

Further experimental evidence bearing on the issue of inflectional and derivational encoding has been obtained by Feldman (1994) and Feldman, Frost, and Pnini (1995). They used the segment shifting task, in which subjects have to produce an utterance by applying the highlighted affix of one stimulus to another stimulus, see (2.12). Note that the examples use English words similar to the Serbian materials used by Feldman (1994).

(2.12)

stimuli		*utterance*
blast**ed** tilt	→	tilted

This task demands from the subjects that they create an inflected form on the spot, and the results are therefore of great interest to us. It was found that shifting a suffix from an inflected word, for example /-ɪd/ from *blasted*, was easier than shifting the same word final string /ɪd/ from a morphologically unanalysable word, *sordid*. For derivational suffixes, there was no difference between the ease of shifting from truly morphologically complex words *(own+er)* and the ease of shifting from morphologically simple words *(over)*.

Feldman claims that this difference is due to the more transparent morphemic structure of the inflectional form. This transparency might have its effects both on the perceptual part of the task (isolating the suffix) as on the production part (joining the stem and the suffix). In a split morphology model, the relative ease with which an inflectional affix can be attached is due to the processing assumption that inflectional affixes are always added on-line to lexical stems. Adding a derivational suffix, however, is not such a common operation because derivational forms are stored in the lexicon and derivational rules may only have descriptive power (see also Chapter 6). Assuming that the results Feldman obtained are not solely due to reading effects, these experiments support a distinction between inflectional and derivational morphology in language production.

[9] Note that the erroneously doubly marked forms *werk+en+t* and *werk+t+en* coincide with existing Dutch verbal forms, making the *a priori* likelihood of these errors even larger.

In summary, the data on language production reviewed here support an inflectional encoding component that is very different from the derivational encoding component: The results of the processes are usually not mixed and the way they work might be fundamentally different. Within the inflectional paradigm, there are data that suggest that singular or masculine inflectional forms have a special status, but this might be due to general markedness constraints. Furthermore, stacking of two conflicting inflectional suffixes (eg. singular and plural) is extremely rare, and a theory of inflectional encoding should explain why this is so.

Implicit priming

Many of the empirical results that relate to the Levelt/Roelofs-model have been obtained with the implicit priming paradigm, as developed by Meyer (1990, 1991). Here, I will overview the major findings within this paradigm. Although originally developed to investigate the time course of phonological encoding in language production, Roelofs and co-workers (Roelofs, 1996, 1998; Roelofs et al., submitted) have shown that the method can also be used to measure effects of the morphological encoding level. This allowed me to use implicit priming for all experiments reported in this thesis. In this short overview, the focus will also be on results pertaining to morphology.

The implicit priming paradigm involves the production of words that are part of a list of previously learned paired associates. The subjects first learn a small set of prompt–response pairs (see the first column of Table 2.1). Then, on each trial one of the prompt words appears on a computer screen and the subject produces the corresponding response word as fast and accurately as possible. Across trials, the prompts are presented in a random order and each prompt appears several times. Finally, a new set of prompt–response pairs is shown and the cycle starts all over again.

Production latency (time between onset of prompt and onset of speech) is the main dependent variable. The experimental manipulation concerns the overlap between the response words that appear together in a set. The example experiment shown in Table 2.1 contains three *homogeneous* sets with initial segment overlap between the response words (read down), and three *heterogeneous* sets with no overlap (read across). Meyer (1990) observed faster responses in the homogeneous sets than in the heterogeneous sets. Henceforth, I will refer to this difference as the *preparation effect*, because it reflects whether subjects can partially prepare the response word. Many experiments have been done with this task and there is ample evidence that the locus of the preparation effect lies in the phonological encoding process: Next to the early finding that only initial segments can be prepared, which invalidates a straightforward memory account, the finding that preparation can be cancelled by the metrical structure of the words (Roelofs & Meyer, 1998) indicates that preparation in the implicit priming task is due to language production processes proper (see Levelt et al., 1999, for an extensive review).

Table 2.1: English examples of prompt–response pairs similar to those used in the Meyer (1990) study: Homogeneous sets in columns, heterogeneous sets in rows.

	Homogeneous		
Hetero-geneous	fruit – melon	comrade – fellow	sponge – bucket
	iron – metal	river – ferry	rabbit – bunny
	grass – meadow	bird – feather	milk – butter

Effects of morphology

The preparation effect due to shared initial segments is called the *phonological preparation effect.* Roelofs (1996) showed that subjects exhibit an additional *morphological* preparation effect for shared initial morphemes. In his experiments, the critical condition contained morphologically complex words (compounds) sharing the first morpheme: *bijrol, bijvak* and *bijnier (supporting role, subsidiary subject, adrenal gland),* where *bij* is a frequent Dutch preposition, comparable to English *by.* A matching heterogeneous condition was used as the baseline; this condition contained non-overlapping compounds. The obtained preparation effect for compounds sharing the initial morpheme was extremely large.

To establish that the preparation effect for shared initial morphemes was larger than for shared segments alone, Roelofs compared the effect to that of a control condition. The items in this condition started with the segments *bij,* but the items did not contain the morpheme *bij* proper, because they were monomorphemic. This set contained words like *bijbel, bijna,* and *bijster (bible, nearly, unduly).* Using an appropriate baseline, a preparation effect was found in this condition, but it was significantly smaller than the preparation effect for the condition with compounds. This is evidence for a morphological preparation effect due to shared morphemes, which occurs on top of the phonological preparation effect due to shared initial segments.

In a follow-up experiment to this, Roelofs tested compounds that shared the second morpheme (*deegrol* 'rolling-pin', *koprol* 'somersault', *bijrol* 'supporting role'). No preparation effect was obtained for these words, neither phonological nor morphological. The consistent finding that non-initial overlap does not help the subjects (see also Meyer, 1990), has been explained by the fact that language production proceeds in a strictly left-to-right fashion. Subjects are able to prepare part of the utterance if the response words overlap initially and this will enhance their response latency, but preparation of non-initial parts is impossible.

The occurrence of a morphological preparation effect seems to hinge on the linear order of the morphemes in the utterance, and is not influenced by morphological headedness or 'underlying word order'. Verbs with separable particles were used by Roelofs (1998) to research this issue. In the infinitive, the particle precedes the stem and it is written like a prefix (*opzoeken* 'look up', infinitive). When inflected for number, person, or tense, the particle is separated and appears after the stem (*zoek op!* 'look up', imperative). Linguistically speaking, one of the two morphemes must be the head of the particle–verb combination. The stem *zoek* is the most likely candidate because it is of the same word class as the resulting combination. In a generative approach, the varying word order of particle and stem is accounted for by assuming one underlying word order, which can be reversed by a movement operation.

In a first experiment, (Roelofs, 1998) compared two conditions with the particle in first position. A significant morphological preparation effect was obtained when the verbs in the set shared the same particle (*opzoeken, opdraaien, opgeven*), but not when they shared the same stem (*opzoeken, afzoeken, uitzoeken*). The reverse was found in a second experiment, in which the infinitival verbs were replaced by imperatives. Now the verbal stem is in first position and a morphological preparation effect showed up for shared stems (*zoek op, zoek af, zoek uit*), and not for shared particles (*zoek op, draai op, geef op*). Clearly, it is the actual and not the underlying word order that determines the effects.

The size of the morphological preparation effect was also examined by Roelofs (1998). Again using particle verbs, he manipulated the frequency of the verbal part while keeping the particle identical: *geef op! (give up, high frequent)* vs. *veeg op! (clean up, low frequent).* The low frequent stems gave rise to the largest preparation effects, similar to what was predicted by the implementation of the Levelt/Roelofs-model. The reason for this reverse correlation between frequency and effect size is that low frequent items take longer to encode, and therefore a larger reduction of the reaction time is obtained when they can be (partly) prepared.

Similar effect for compounds were obtained by Roelofs (1997a). In this study, the frequency of the first part of compounds was varied, while the frequency of the compound itself was kept constant: *schuimbad (bubble bath)* was contrasted with *schoolbel (school bell)*, where *schuim* is low frequent (419 per million), *school* is high frequent, and both compounds are of similar low frequency. Subjects showed a significantly larger preparation effect for response words with a low frequent first part than for words with a high frequent first part, again showing subjects' sensitivity to morpheme frequency. As was mentioned in the previous chapter, this result confirms the hypothesis that the form of a compound is represented in terms of its constituting morphemes. This is true for transparent items, like *bloedproef (blood+test)*, as well as opaque ones, like *klokhuis (clock+house → core of an apple)*.

Metrical frames

In this thesis, evidence is presented for the presence of frames in the production of inflectional forms. The search for such a frame was in part motivated by the finding that *metrical frames* appear to be used in the phonological encoding of a word. Roelofs and Meyer (1998) showed that the phonological preparation effect, which is normally obtained when the response words show initial segment overlap, depends on sharing a *metrical frame*. The metrical frame is used during the phonological encoding stage of speech production. Part of the metrical frame is a specification of the main stressed syllable. In one of their experiments, Roelofs and Meyer used sets of response words that overlapped in their initial segments, but with position of stress varied within a set. Such a set would, for example, contain the words *marine (navy)* and *madelief (daisy)*, where underlining indicates the primary stressed syllable. No phonological preparation effect was found in this condition with variable metrical structure, whereas the same words produced the expected preparation effect for shared segments 'ma' when grouped in sets with constant metrical structure, ie. *marine* in the context of *materie*, 'matter'.

Roelofs and Meyer explained this dependency by proposing that encoding of segments can only be done after the retrieval of the metrical frame, so that segment-to-frame association is possible. I report a similar dependency of morphological encoding on the morphological (inflectional) frame in the following chapters.

The inflectional frame hypothesis

CHAPTER 3

Morphological frames?

In the previous chapter, several studies were mentioned that have successfully applied a slots-and-fillers type of theory to experimental findings on language production. The prime example is the study by Roelofs and Meyer (1998), who found evidence for phonological frames that are used for the construction of phonological words. In the Levelt/Roelofs-model of language production (Levelt, 1989; Roelofs, 1992a, 1997c; Levelt, Roelofs, & Meyer, 1999), there are two separate encoding steps for morphology and phonology. The obtained evidence for frames in the phonological encoding step makes one wonder whether frames are also used for morphological encoding.

What could be the role of such a hypothetical frame? For this, we will take a look at what the duties of morphological encoding are. The ultimate goal of the process is to convert a lemma specification into one or more morphemes. This process is deceivingly simple for everyone's favourite example words, like *ship:* Part of the lemma SHIP is a pointer into the form lexicon, where all morphemes are listed. This lemma points to the morpheme *ship*, and after this morpheme is selected and phonologically encoded, the word can be uttered.

No frame is necessary for a process this simple, but there are two complications to this picture which may change this. One complication is that there are words which are more complex than *ship*, and the other suggests that the process sketched above is not adequate, not even for a word like *ship*.

The first extension that has to be made concerns lemmas that are expressed by more than one morpheme. Take for example the lemma MAINSAIL: Two pointers into the form lexicon are necessary to produce the compound form *main+sail*. A similar construct has to be employed for words like *durable*, *department* and *devilish*. All are formed from a stem[1] and a derivational suffix. Although the amount of transparency varies, the Levelt/Roelofs-model claims that all three of these words are represented by one lemma and two morphemes.

[1] There are several levels of analysis to which 'stem' can apply. Its net meaning seems to be

The second extension that needs to be made concerns words like *ships.* This word is generated from the same lemma SHIP by changing the *diacritic* for number. The diacritics reside at the selected lemma, and are used to convey information from the conceptual level that could not be expressed by the mere selection of the lemma. How does this affect the word *ship* itself? Although it is not visible at the segmental level, the singularity of the concept has to be specified by a diacritic too. I propose that (for regular words) an inflectional affix is inserted for every diacritic, irrespective of its value. For singular nouns in English, this means that a null morpheme is generated (*ship+ø*).

In sum, there are at least two types of affixes that need to be attached to the stem: derivational and inflectional affixes. The obvious function for a frame in this process is to hold the stem and the affixes, and guide the process of combining them. This is indeed what is proposed: First, a frame is created with a slot for every necessary affix, next, those slots are filled by morphemes. The proposed system has some clear advantages for language production, the foremost being that there is a check on whether all affixes are indeed expressed. If the frame specifies two affix slots, the result of lexical access should be precisely two affix morphemes.

Inflectional frames

But a general question about the *why* of frames (that was addressed for phonological frames by Levelt et al., 1999) turns up here: There clearly is redundancy in separately specifying structure (the number of affixes) and content (the affix morphemes). This redundancy might help to prevent errors, but so would generating the word twice. What has to be shown is that specifying the structure (the frame) can be done at relatively low cost, with relatively large benefits for error checking on morpheme generation. In other words, the frame has to be an efficient way of solving the problems that language production faces.

Does a morphological frame meet the efficiency constraint? I will argue that this depends on the type of affixes for which the frame contains slots. For the derivational affixes, the efficiency is dubious at least. Consider a form like *self+ish+ness,* which contains a noun and two derivational affixes. There is no other way to obtain the number of affixes than by counting them. It is not possible to make a generalisation that would predict the number of affixes from word class, meaning, or other sources. This is shown by the word *ego+ism,* which is a derived noun with a similar meaning but only one derivational affix. The word class of the stem does not help either: Consider the related word *fool+hard+i+ness,* which also has a nominal

something to which affixes can be applied. In the following chapters, 'stem' will mostly mean the base for application of *inflectional* affixes, ie. derivational affixes form part of the stem. In this section, the word will be used in its widest meaning.

stem, but takes three derivational suffixes to become a noun with a similar "having the property of X" meaning.

For inflectional affixes (the second extension) there are certainly better prospects. Prediction of the number of inflectional affixes is generally possible from the word class of the stem: In Dutch, nouns are inflected for number only, verbs are inflected for number and tense, and adjectives are inflected for gender. The number of slots is equal to the number of inflectional affixes under the assumption that unmarked cases, like English singular, are expressed by -ø. This assumption is widely held by morphologists (Bybee, 1985; Anderson, 1992) and will be further discussed in Chapter 6. If it is adopted, the number of inflectional affix slots in the frame can be predicted for all regular words.

The bottom line is that a frame is an efficient way to check the presence of necessary inflectional affixes, but not so for derivational affixes. I therefore propose not a general morphological frame, but a specific inflectional frame. This frame should contain an affix slot for every diacritic that is specified for the stem. Our working definition will avoid the theory-internal notion of a diacritic:

(3.1) **The inflectional frame** is used to guide the process of combining the stem with the inflectional affixes. The frame contains one slot for the stem and one slot for every possible type of inflectional suffix that occurs with that stem. The stem may contain derivational affixes and may be a compound, but will always occupy exactly one slot of the inflectional frame.

The distinction that has been made between inflectional and derivational morphology at the level of the inflectional frame will have consequences for the processing of either type of affixes. If the last part of the definition does indeed hold (experimentally), the processing of derivational morphemes must at least be partly separate from that of inflectional morphemes. This makes a subdivision of the morphological encoding component necessary, which is what I will suggest in Chapter 4, after presenting the data that bears on this issue. From the experimental data reviewed in the previous chapter, and the linguistic arguments for a categorical difference between inflection and derivation presented in Chapter 1, the splitting of morphological encoding should not come as a surprise.

The proposal made above is explicitly restricted to the processing of regular inflectional forms. The behaviour of irregular forms is also of great interest, and this issue will be taken up in various places in the next two chapters. In general, the irregularity of a word consists of it not taking certain standard inflectional affixes. If an affix is never present, there should also be no slot reserved for it. Thus, irregular words are accompanied by irregular inflectional frames. A detailed proposal about this will be made in Chapter 4.

Slots

One last question that needs to concern us is which types of slots will be used in the proposed inflectional frames. In Shattuck-Hufnagel's account, slots carry labels and so do all possible fillers. Part of the error checking done by the frame is to match the label on the slot with the label on the filler. If they do not match, insertion of the filler is not allowed. Labels are thus an important error checking device in this approach.

The phonological frames proposed by Roelofs and Meyer (1998), however, do not bear labels at all. The frames found by these authors are sparser and lack error checking through labels. Part of the reason for this is that, apart from their linear order, there are few intrinsic difference between the abstract syllables of a word. The benefit of this approach is that it is no longer necessary to label all possible fillers.

For inflectional frames, an intermediate solution is proposed. Some labelling of slots and fillers will be assumed, but not to the specific detail of, for example, a slot being reserved for tense suffixes. There are few advantages of such a specific encoding, especially when affixes occur always in the same order, like in Dutch.[2] A more generic division of slots in three classes combines essential error checking with sparseness of representation. Three major classes of slots in the inflectional frame will be distinguished: prefix slots, stem slots, and suffix slots.

[2]This claim may not generalise to languages with much more complicated inflectional systems, where more advanced checks on the contents of the frame might be necessary.

Evidence for inflectional frames

CHAPTER 4

In this chapter, 6 experiments will be reported that address the proposed inflectional frame. The first experiment will look for an influence of the inflectional frame on the morphological preparation effect, using inflected verbs. Having found this influence, the next two experiments explore the limits of this effect by looking at derivational forms. These two experiments also serve as control experiments for Experiment 1.

The preparation effect in the first experiment was greatly reduced by an unpredictable inflectional frame. In Experiment 4, the size of this effect will be further examined by comparing it to a phonological preparation effect. The results of that experiment will call for a timing proposal, to explain the outcomes without introducing a violation of the modularity of phonological and morphological encoding.

Experiment 5 will then apply a similar methodology to plural forms. The results confirm the definition of the inflectional frame in terms of *possible* affixes, and refutes an explanation of the results so far by a storage–computation distinction. In the last experiment, the question of how to incorporate irregular forms is raised. The obtained results for strong past tense forms will lay the foundation for further treatment in Chapter 5.

Implicit priming with an odd man out

In all the implicit priming experiments reviewed in Chapter 2, one type of word was used for each of the two homogeneous conditions. Roelofs' (1996) experiments, for example, contained the following two types of homogeneous sets: In one condition, the three words in the set were all morphologically complex, and in the other condition, they were all morphologically simple.

The experiments reported in this chapter will employ a slightly different paradigm: Instead of two conditions containing different types of words, one condition will contain only neutral words and the other will contain

The text of this chapter is an adapted version of a manuscript submitted for publication in Language and Cognitive Processes. Special thanks to Gerard Kempen for bringing the possible influence of the storage–computation distinction to my attention (Experiment 5).

neutral words plus one special word. This last word is called the odd man out. See Table 4.1 for example materials from the first experiment that will be reported on. The condition with the odd man out is called the variable condition, the other is called the constant condition.

Table 4.1: Example materials from Experiment 1. Homogeneous sets only, *werkte* is the odd man out.

prompt – response		prompt – response	
Variable Condition			
prompt	response	prompt	response
arbeid	*werkte*	labour	worked
stress	*werkdruk*	stress	workload
opdracht	*werkboek*	exercise	workbook
Constant Condition			
prompt	response	prompt	response
honing	*werkbij*	honey	worker bee
overleg	*werkgroep*	consultation	study group
garage	*werkplaats*	garage	workshop

The hypothesis tested will always be whether the odd man out spoils the preparation effect that is induced by the neutral words in the variable condition. Running slightly ahead of things, in Experiment 1 the hypothesis is that the odd man out *werkte* will spoil the morphological preparation effect that would normally occur due to the shared initial morpheme *werk* in this set. To assess this, the preparation effect in the variable condition will be compared to the preparation effect in the constant condition, which contains only neutral words and which should give rise to a solid morphological preparation effect. For the case that the odd man out indeed suppresses the preparation effect in its condition, it is expected that the reaction times to the odd man out and the surrounding neutral words will be similar: Preparation effects arise from consistency in the set of possible response words and a reduction or cancellation of the preparation effect should therefore affect all response words equally.

Experiment 1: Compounds and verbs

As explained in the previous chapter, our goal is to find evidence for an inflectional frame in language production. The hypothesis that such a frame exists rests partly on the finding of a phonological frame by Roelofs and Meyer (1998). The predictability of the phonological frame turned out to influence the preparation effect in an implicit priming task: The well-established finding of a preparation effect occurring when the response words share the same initial segments (Meyer, 1990, see also Chapter 2) could not be replicated when some of the response words had different phonological frames than the others. The predictability of the phonological frame appears to be a prerequisite for phonological preparation to appear.

In this experiment, evidence is sought for the proposed inflectional frames by testing whether a similar condition holds for the morphological preparation effect (Roelofs, 1996). In a condition with initial morphological overlap, which would normally induce a morphological preparation effect, the inflectional frame will be made unpredictable. The (possibly too strong) prediction is that no morphological preparation effect will be found when inflectional frames vary, because there is no place to store the preparable morpheme when the inflectional frame itself cannot be prepared.

The materials used in this experiment were similar to the ones used by Roelofs (1996). In one half of the experiment (the constant condition), the response words are compounds that share the initial morpheme. A morphological preparation effect is expected in this condition. In the other half (the variable condition), an inflectional frame difference was induced. To this end, two compounds were combined with a past tense verb that was formed from the initial morpheme of the compounds. Example variable and constant sets were shown in Table 4.1 above.

Nouns (ie. compounds in this experiment) and verbs take different numbers of inflectional suffixes. If our assumptions are right, this means that these word classes have different inflectional frames. As can be seen in (4.1), nouns are only inflected for number in Dutch. Verbal forms are inflected for tense and number.[1]

(4.1)	Nouns:	Stem	+	Number		
	Verbs:	Stem	+	Tense	+	Number

Of course, compounds and verbs differ in another important aspect: Verbs have a simple stem, containing one morpheme, and compounds have a complex stem, containing two morphemes (as in *work+shop*). It is argued in Chapter 3 that the complexity of the stem (in the sense of 'stem before inflection', that will be used throughout this thesis) does not affect the inflectional frame. This question will be experimentally addressed in the next chapter. If the assumption about complex stems does not hold, an additional frame difference between compounds and verbs exists.

For the present experiment this does not matter and it suffices to state that there is at least one inflectional frame difference between compounds and past tense forms: the number of inflectional affixes. The prediction is that this difference makes it impossible to prepare the inflectional frame in

[1] Dutch verbs are also inflected for person. Both in Dutch and many other languages (eg. almost all Indo-European languages), the markings for number and person are mutually dependent if not intertwined. Dutch marks 2nd or 3rd person, but only for singular. I therefore assume one suffix for the combination of number and person specification. The predictions for this experiment stay the same if a separate person slot is assumed.

the variable condition (where nouns and verbs are mixed), and a smaller or zero preparation effect will be obtained here.

The logic behind this experiment (and all other experiments in this thesis) can be summarised by the catch-phrase "Equivalent preparation effects means equivalent frames". If the frames of nouns and verbs are the same, no reduction in the variable condition will occur and the variable and constant condition should show the same preparation effect. Conversely, if the preparation effects are different, this must be due to a reduction in the variable condition which is caused by the inflectional frame difference that was introduced.

Method

Participants Twelve participants took part in this experiment, as in most other experiments reported in this thesis. 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.

Materials All experiments were carried out in Dutch. For this experiment, which will serve as a standard for most others in this thesis, 16 sets (4 practice and 12 experimental) were created of three word pairs each. Each pair consisted of a prompt and a response word. Three base morphemes were selected that formed the basis for constructing the experimental sets. In this experiment, these were *bouw, straf, werk (to build, to punish, to work).* An example of the three pairs constructed from *werk* is shown in the top half of Table 4.1 on page 64.

There were two conditions, variable and constant. For the response words in a constant set, I selected three bisyllabic compounds that all have the base morpheme in first position. No semantically opaque compounds were used. For the response words of a variable set, two more compounds were taken, together with the past tense of the base. Under the hypothesis that we entertain, the inflectional frame will vary between items in this condition. 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 described above 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 derived from the three different base morphemes. Care should be taken not to confuse the constant–variable contrast (whether or not the inflectional frame was the same for all items in the set) with the homogeneous–heterogeneous contrast (which determines whether the response words showed initial segmental overlap).

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. To keep the association strength between prompt and response within reasonable limits, no prompts were allowed that started with the same segment(s) as the response (eg. 'gold–gift' is not allowed), that could form a compound with the response word (eg. 'sun–shine' is out), or were possibly lexicalised combinations. All response words were bisyllabic compounds, except for the three past tense forms that were bisyllabic root–affix combinations. All response words had initial syllable 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. The pairing of prompts with responses was the same for homogeneous and heterogeneous conditions. 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.

Design The experiment contained four crossed within-subject factors. The first factor was Base (three levels). This factor corresponds to the three morphemes 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 base morphemes 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 5 times each in a block of trials. The order of presentation was fully randomised per subject and per block, with the constraint that repetition of the same prompts on adjacent trials was forbidden.

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 as accurately as possible. These instructions were summarised by the experimenter, who stressed the fact that making performance errors was inevitable in this experiment, but that all associations had to be thoroughly studied. 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 (ie. when they made too many clicking noises that may 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 who were extremely slow to learn the associations (more than a minute per set of three pairs), were removed from the analyses and replaced by others. This happened very infrequently; 7 times in all experiments reported in this thesis.

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 Courier (typewriter) font. In the test phase, prompts were displayed in a 36 point Helvetica (sans serif) font.

Analysis From the output of the voice key device, all responses resulting from voice key errors, all wrong responses, hesitations, and time outs were removed. When in doubt, the experimenter consulted the recordings that were made of the sessions. In this experiment, 4.7% (101 cases) of the trials were removed. Of these, less than half (42 cases) were 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 analyses were carried out on all errors, although even then the number of data points is 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; Raaijmakers, Schrijnemakers, & Gremmen, in press). Only significant F ratios are reported, together with those that approach significance in a very liberal sense. If an F ratio is not reported for an experiment, its probability was .10 or higher. Together with the F' an MS_e term is reported to give the reader insight in the amount of variance in the data. Because the real error term used in a F' is a pooled sum, the MS_e reported is the interaction of Subjects with the factor at hand, which is part of the actual MS_e.

Occasionally, F_1 or F_2 values will be reported to help the reader evaluate the outcome of the F'. Because repeated measurements were taken from the same subjects, all F_1 computations underwent the Huyhn-Feldt correction for sphericity (Huyhn & Feldt, 1976; Rietveld & van Hout, 1993). This correction makes the F_1 tests more conservative to a degree that varies with the lack of sphericity in the data. The Huyhn-Feldt correction is implemented by modifying the degrees of freedom of the F-test. The reported degrees of freedom are always the original ones.

To test whether there was an overall preparation effect, the test on the so-called 'constant' term is reported. This is a contrast comparing the mean of all cells to zero, using items for estimating the within cell variance. Significance of this test indicates that the mean difference score is larger than zero, which implies that there was an overall preparation effect (ie. an effect of Context; homogeneous sets are faster than heterogeneous sets). This can be done with a regular F test instead of an F', because the expected value of the constant term does not contain any subject or item terms (see Appendix B on the issue of expected mean squares). The constant effect is tested against the mean square for subjects. For all experiments that showed a significant overall preparation effect and a significant effect of Word Type, the size of the preparation effect within each level of Word Type (ie. for the variable and the constant condition) was computed. An F' is reported which simultaneously tests the preparation effect for a specific level of Word Type against the overall Items error term and against a pooled error term obtained from the factors Subject and Subject by Word Type (Hays, 1994).

Two further tests were done routinely to ensure that the assumptions underlying the statistics were met: First, Tukey's test for non-additivity was run. Second, Levene's test on the homogeneity of variance was done. The results of these test are only reported when the assumptions were not met

Table 4.2: Mean production latencies (ms), error percentages, and preparation effects (Δ) for Experiment 1: Compounds and verbs

	Context				
Word Type	homogeneous		heterogeneous		Δ
variable	672	*6.9%*	708	*4.1%*	36
constant	619	*3.2%*	741	*4.6%*	122

(Maxwell & Bray, 1986; Santa, Miller, & Shaw, 1979; Rietveld & van Hout, 1993).

Results and Discussion

The results of this experiment are summarised in Table 4.2. Subjects were faster in the homogeneous than in the heterogeneous condition, but this difference was far larger for the constant Word Type than for the variable Word Type condition (122 vs. 36 ms difference). The existence of an overall preparation effect is confirmed by the statistics: $F'(1,19) = 31.28, MS_e = 31518, p = .000$. The difference in the preparation effect between the variable and constant Word Types was also significant: $F'(1,23) = 17.54, MS_e = 13842, p = .000$. The preparation effect was significant for both levels of the Word Type factor: In the variable condition, $F'(1,21) = 4.47, MS_e = 21461, p = .043$; in the constant condition $F'(1,21) = 48.48, MS_e = 21461, p = .000$. Overall, 4.7% errors were made. As can be seen in Table 4.2, there were slightly more errors in the variable–homogeneous condition than in all other conditions. This resulted in a significant effect for Word Type in the error data: $F'(2,23) = 4.18, MS_e = 0.29, p = .039$. Because most errors were made in the slower of the two homogeneous conditions, the errors cannot be due to a speed–accuracy trade off.

It is not the case that the reduction of the preparation effect was specific to the *odd*, inflected words. All other words in the variable condition (the compounds) showed similar reduced preparation effects, as can be seen from the following mean preparation times over items: verbs (variable condition) 39 ms; compounds (variable condition) 34 ms; compounds (constant condition) 122 ms (see Figure 4.1). Clearly, the preparation effect for compounds in the variable condition is comparable to the preparation effect of verbs that occur in the same condition, and not to the compounds that occur in the constant condition.

The results of Experiment 1 point to an influence of the inflectional frame

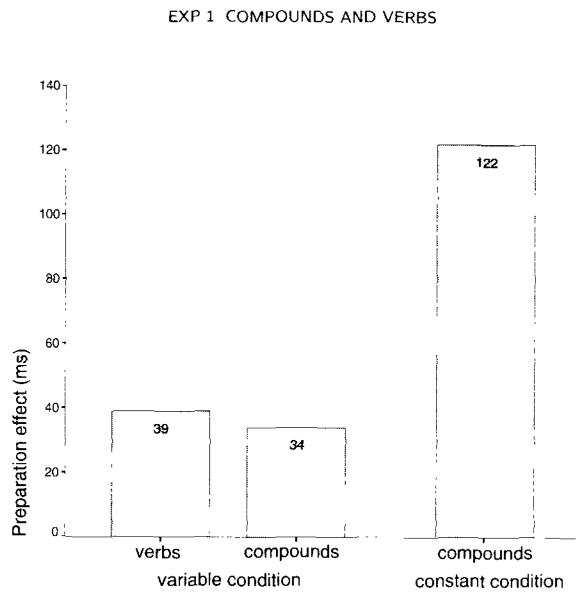


Figure 4.1: Mean preparation effect in Experiment 1 for verbs and compounds in variable and constant conditions.

on the preparation effect. In the constant condition, a stable morphological preparation effect was observed. In the variable condition, which contained varying inflectional frames, the preparation effect was strongly reduced but still significantly larger than zero. Variability in the inflectional frame reduces the morphological preparation effect, similar to what was found for phonological frames. Cast in terms of the slots-and-fillers account, one can say that the benefit from preparable morphemes goes away when there is no frame to store these morphemes in.

No perfect symmetry between the phonological and inflectional frame was found: Where variability of the phonological frame suppressed the preparation effect completely in the experiments of Roelofs and Meyer (1998), variability of the inflectional frame only reduced the preparation effect in this experiment. Note, however, that the morphological preparation effect is always additional to the phonological preparation effect, because segmental overlap is a logical necessity for morphological overlap. I want to argue that the full preparation effect due to morphemes was cancelled out by variability of the inflectional frame, and only the preparation effect due to segmental overlap remained.

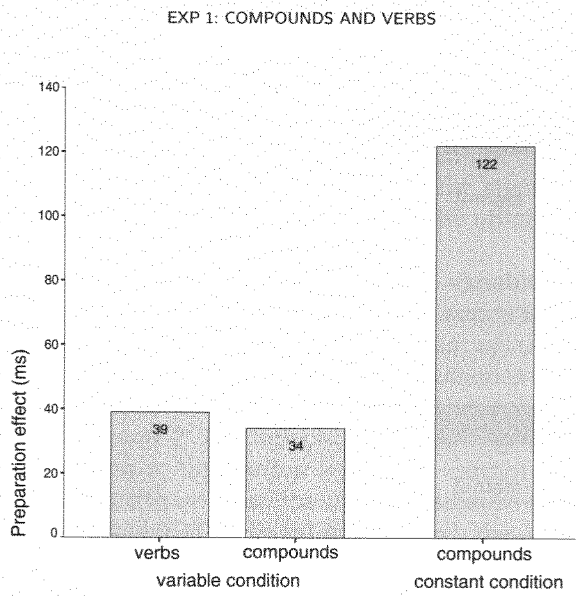


Figure 4.1: Mean preparation effect in Experiment 1 for verbs and compounds in variable and constant conditions.

on the preparation effect. In the constant condition, a stable morphological preparation effect was observed. In the variable condition, which contained varying inflectional frames, the preparation effect was strongly reduced but still significantly larger than zero. Variability in the inflectional frame reduces the morphological preparation effect, similar to what was found for phonological frames. Cast in terms of the slots-and-fillers account, one can say that the benefit from preparable morphemes goes away when there is no frame to store these morphemes in.

No perfect symmetry between the phonological and inflectional frame was found: Where variability of the phonological frame suppressed the preparation effect completely in the experiments of Roelofs and Meyer (1998), variability of the inflectional frame only reduced the preparation effect in this experiment. Note, however, that the morphological preparation effect is always additional to the phonological preparation effect, because segmental overlap is a logical necessity for morphological overlap. I want to argue that the full preparation effect due to morphemes was cancelled out by variability of the inflectional frame, and only the preparation effect due to segmental overlap remained.

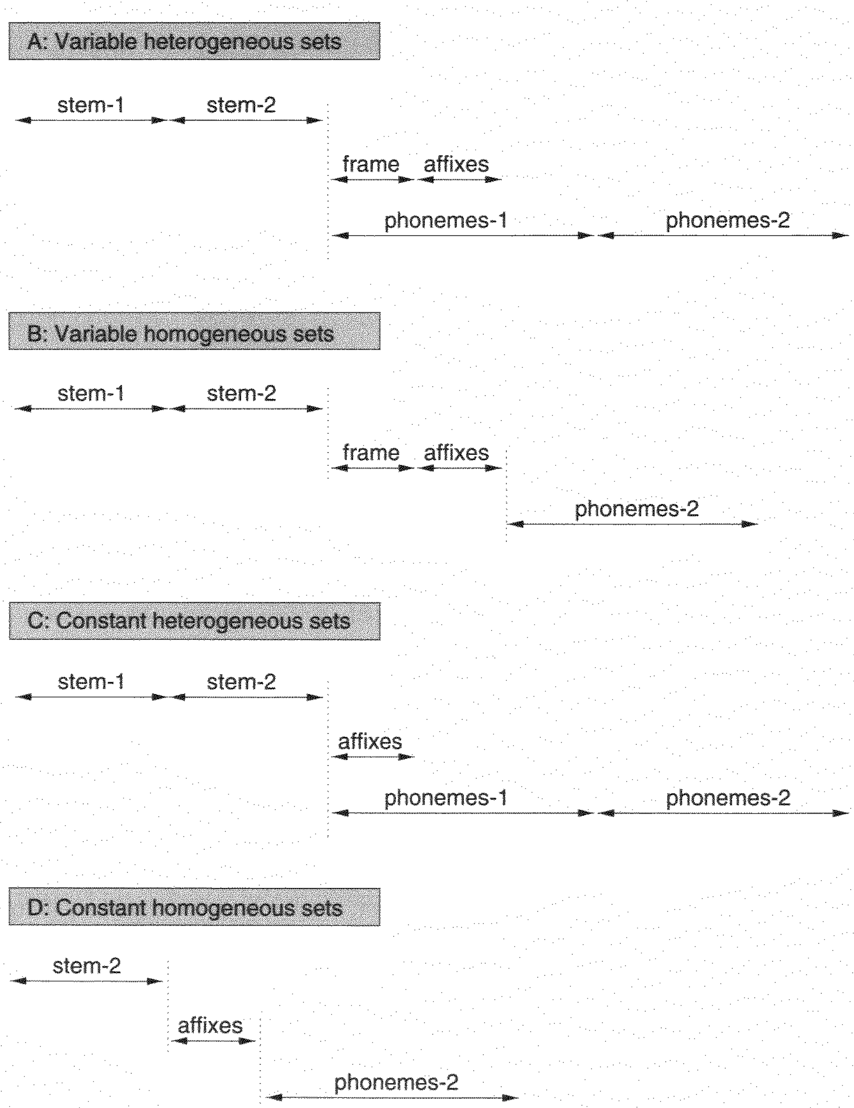


Figure 4.2: Ordering of on-line processes in Experiment 1. Phonological encoding has been split into *Phonemes-1*, the part that can be prepared in the homogeneous sets, and *Phonemes-2*, the part that always has to be done on-line.

A timing proposal

A proposal on the ordering of the various processing steps involved in the production of a word is shown in Figure 4.2. The aim of this proposal is to show how preparing certain steps can influence the steps that remain to be done on-line in a way that corresponds with the obtained preparation effects.

In Panel A, the heterogeneous condition with variable morphological frames is shown. In this condition, the past tense *werkte* is combined with compounds *strafschop (penalty)* and *bouwjaar (year of construction)*. Neither the frame, nor the first morpheme, nor any phonemes can be prepared. The on-line production processes starts off by retrieving the two stem morphemes, *stem-1* and *stem-2*. The inflectional frame is generated from the word class information at the lemma level. This cannot be done before the morphemes are retrieved from the lexicon, because the lexical entry might contain an irregular inflectional frame. Next, the affix morphemes are retrieved in the *affixes* step. The inflectional frame is essential for the storage and checking of affix morphemes and this implies that affix encoding must be done after the inflectional frame has become available. Meanwhile, phonological encoding has started working on the stem morphemes. Phonological encoding has to proceed from left to right (Roelofs & Meyer, 1998; Roelofs, 1998). In the figure, the many segment-sized phonological encoding steps have been grouped in two larger units, *phonemes-1* and *phonemes-2*. The first comprises all segments that can be prepared in the homogeneous condition, the latter the segments that can never be prepared. Because there is no phonological overlap, all phonological encoding is done on-line in Panel A.

In Panel B, the homogeneous variable condition is shown, in which *werkte* is combined with the two compounds *werkdruk (workload)* and *werkboek (workbook)*. Now, the initial phonemes of the morpheme *werk* can be prepared. Notice that the morpheme itself cannot be prepared, because there is no preparable inflectional frame to store the morpheme in. The process labelled *phonemes-1* has been moved to the preparatory stage (not shown in the figure) and this leads to faster reaction times. The encoding of the second part of the word *(phonemes-2)* does not directly start after the retrieval of the stem morphemes because it also depends on the affix morphemes. This dependency is self-evident for the affixed response words like *werk+te*. The same holds for the compound response words which bore no discernable inflectional affixes: Encoding of the second part of the word still has to await the null outcome of the *affixes* step. The difference between the endpoint of the *phonemes-2* processes in Panel A and Panel B is the preparation effect obtained (36 ms).

The situation for the constant conditions is similar to what is described above. Because compounds are used throughout, the inflectional frame can be prepared in the homogeneous and in the heterogeneous condition. The heterogeneous constant condition is shown in Panel C. It features non-overlapping compounds *(werkplaats, bouwkunst, strafrecht)*. The inflectional frame can be prepared even though there is no overlap: All response words are nominal and the inflectional frame is entirely predictable and therefore preparable. In Panel C, both parts of the stem and all phonemes are encoded on-line. The prepared inflectional frame does not affect the net reaction time because phonological encoding proceeds from left to right and the phonological encoding of the affixes *(phonemes-2)* happens too late to benefit from the earlier availability of the affix morphemes.

The matching homogeneous condition (Panel D) contains three compounds that all have *werk* as their first constituent *(werkplaats, werkgroep, werkbij)*. This leads to an entirely different picture: The first morpheme of the stem and the segments that encode this morpheme can be prepared (processes *stem-1* and *phonemes-1*). Again, phonological encoding of the second part of the word has to await the morphological encoding of the affixes, and this can only be done after all stem morphemes are retrieved (this is so because some stems have irregular variants that replace regular affixation). Still, because the inflectional frame can be prepared, morphological encoding of the affixes can start right after the retrieval of second part of the stem and phonological encoding *(phonemes-2)* follows hot on its heels. This leads to a much shorter reaction time in this condition and, consequently, a large preparation effect (122 ms; the figure is not to scale).

The explanation given depends on several assumptions: First, a relevant property of the Levelt/Roelofs-model is its *suspend–resume* mechanism. This mechanism was introduced to provide an explanation for the preparation effect itself: It is assumed that the language production process proceeds as far as it can when it faces partial input. It then halts until the required information becomes available. In the constant homogeneous condition (Panel D), for example, the specification of the initial morpheme forms a partial input that is present before the prompt word is displayed. With this input, the production process can go through part of morpheme retrieval process and part of the phonological encoding process. After this, the production process halts. It is resumed when more information arrives, ie. after the response word is determined from the displayed prompt.

A further assumption that was made is that the actual contents of the inflectional frame do not matter for its preparation. What is crucial is the number of slots because this guides the encoding of inflectional affixes. Next to that, the inflectional frame is needed for storing the stem morphemes. If no frame is available, no shared initial morpheme can be prepared because there is no place to hold it. The relatively small preparation effect in the homogeneous variable condition (Panel B) is explained by the failure to prepare the shared morpheme. Lastly, the duration that was assumed for the various processing steps can be changed but must obey some constraints to allow the timing diagram to be correct: The duration of the *af-*

fixes step, for example, influences the preparation effect in both conditions and if *affixes* plus *frame* were longer than *phonemes-1*, no preparation effect in the variable condition would be predicted.

Some disclaimers have been made about the behaviour of irregular forms, which can be associated with irregular frames. Other forms have irregular variants that replace the regular stem plus inflectional affix combinations. A more detailed analysis of these irregular forms will be given in the General Discussion.

In Experiment 4, this timing proposal will be tested for phonological preparation effects only. But first, two possible alternative explanations of the data have to be tackled. For one, there is a word class difference in the variable sets that is not present in the constant sets. Could it be that mixing words of different classes caused the reduction of the morphological preparation effect? This question will be addressed in the following experiment.

A second issue that needs further attention is syllable strength. The stimuli in this experiment were carefully designed to have matching metrical frames: they were all bisyllabic with stress on the first syllable. Because of the nature of the Dutch past tense affixes however, the past tense forms had a weak second syllable (strong–weak pattern), where the compounds had a strong second syllable (strong–strong pattern). This metrical difference might have an effect on morphological or phonological preparation. Experiment 3 addresses this question.

Experiment 2: Derived adjectives

Experiment 2 was designed to test for a word class effect. I wanted to exclude the possibility that mixing words of two word classes, rather than mixing words with two types of inflectional frame, induced the effects found in Experiment 1. To this end, nouns were combined with adjectives. These two types of words have the same metrical and inflectional frames, as will be explained below. We expect to find morphological preparation under both conditions, irrespective of the word class difference between nouns and adjectives.

All Dutch adjectives are inflected for gender. The two genders (neuter and non-neuter) cause different marking on the adjective in indefinite NPs: *een mooi huis (a beautiful–neuter house)* versus *een mooie kat (a beautiful–nonneuter cat)*. Adjectives in definite NPs always bear the *-e* suffix because the neuter/non-neuter distinction is made by the article: *het mooie huis (the–neuter beautiful house)* versus *de mooie kat (the–nonneuter beautiful cat)*. A slot in the inflectional frame is needed to accommodate this gender suffix.

This leads to a 2-slot frame for Dutch adjectives, with a stem slot and a slot for gender. This frame is fully comparable to a noun frame:

(4.2)	Nouns:	Stem + Number
	Adjectives:	Stem + Gender

The adjectival forms in this experiment were formed from verbal stems by the affix *-baar* (comparable to English *-able*). As in the previous experiment, bisyllabic compounds were used for all the other response words. An example variable homogeneous set is *brandbaar, branddeur, brandpunt (combustible, fire exit, focus)*. The prediction is that subjects will benefit from the morphological overlap between these three forms as much as they will benefit from the overlap between three compounds in the constant homogeneous sets.[2]

Method

Participants, apparatus, procedure, and analysis were the same as in Experiment 1, with the exception of the number of replications: Each prompt was presented eight times in each set in this experiment, as opposed to five presentations in other experiments. The total number of trials went up from 240 to 384.

Materials Three base morphemes were selected to construct the sets from. These were: *lees, brand, draai (to read, to burn, to turn)*. For the variable sets, an adjective was formed from the base with the suffix *-baar*. Two bisyllabic compounds were chosen, which contained the base as their first morpheme. For the constant sets, three additional such compounds were selected. See the Appendix for the full set of materials.

Results and Discussion

After removing the errors from the data (4.8% of all data points), difference scores were computed. The mean reaction times are reported in Table 4.3. There was a preparation effect both for the variable and the constant sets. The preparation effect in the variable condition was slightly larger than in the constant condition (92 vs. 77 ms). The overall preparation effect was significantly different from zero, $F'(1.17) = 29.31. MS_e = 41758. p = .000$, but the size of the preparation effect did not differ significantly between the variable and constant Word Types: $F'(3.21) = 1.13. MS_e = 5469. p = .354$. The interaction between Word Type and Base was insignificant (F' (3,27)=2.50, MS_e =7604, p=.0842) and so were all other factors.

There was, in other words, no effect of mixing word types. We conclude that word class per se has no influence on the preparation effect and was

[2]The linguistically inclined reader might wonder why we do not allot a slot for the comparative and superlative suffixes. It is doubtful whether these processes belong to core inflection, but apart from that, the *-baar* forms frequently occur with the circumlocutory comparative form *beter leesbaar (more readable)* instead of the complex form *leesbaarder*.

Table 4.3: Mean production latencies (ms), error percentages, and preparation effects (Δ) for Experiment 2: Derived adjectives

	Context				
Word Type	homogeneous		heterogeneous		Δ
variable	676	*5.2%*	768	*6.0%*	92
constant	675	*4.5%*	752	*3.7%*	77

not of influence in Experiment 1. As an aside, our analysis of the comparative and superlative adjectival forms was warranted by the data. Given the comparable preparation effects in the two conditions, the inflectional frames for nouns and adjectives appear to be comparable too.

Experiment 3: Diminutives

The second possible counterargument to Experiment 1 involved metrical structure. The past tense verb forms (like *werkte*) have a strong–weak syllable pattern, whereas compounds (like *werkdruk*) are strong–strong. To make sure that it was not this particular metrical difference that caused the effects found in Experiment 1, an experiment was designed that contains *odd* words that have the strong–weak pattern of the verbs used in Experiment 1, but are otherwise comparable to compounds. If it was the metrical structure of the verbs that caused the pattern of results we obtained, the same pattern will be found in this experiment. It is expected that this will not be the case, because under our hypothesis, the results were caused by a difference in inflectional frames. In this experiment, all words will have the same inflectional frame.

It might seem contradictory that no effect of metrical structure is expected here, whereas Roelofs and Meyer (1998) have shown an effect of the metrical frame on phonological preparation. But according to these authors, it is only the number of syllables and the place of stress that gives rise to different metrical frames, and *werkte* and *werkdruk* have the same metrical frame (2 syllables, initial stress). It is the possible effect of syllable strength that we will be testing here.

The *odd* words in this experiment should be bisyllabic, morphologically complex words with strong–weak syllable patterns. Diminutives meet all these criteria: Dutch has a productive derivational affix *-(t)je* (with some phonological variants) to make diminutives of nouns. The affix can even be applied to neologisms and loan words, as in *computertje (little computer).*

When combined with a monosyllabic stem, the diminutive process creates a new bisyllabic stem with a strong–weak syllable pattern. The resulting word is a noun, and can only be inflected for number.

(4.3) Compounds: Stem + Number
Diminutives: Stem + Number

This means that the inflectional frame of a diminutive is equivalent to that of a compound, while the syllable pattern is strong–weak, like that of the past tense form. An example set from the variable homogeneous condition is *vakje, vakbond, vakman (box* DIM, *union, expert).* Our prediction is that the metrical difference between the words has no effect on preparation, and the preparation effect for the variable condition will be similar to the preparation effect in an appropriate constant condition with no metrical difference.

There has been some debate over the status of (Dutch) diminutives. Diminutive formation has indeed some characteristics of an inflectional process, but morphologists nowadays agree that it is derivational (see for example Booij & van Santen, 1998, the latest textbook on Dutch morphology). For a start, there is meaning specialisation in diminutives like *knietje (the act of giving someone a knee in the groin, lit. knee* DIM*),* and *briefje (note, lit. letter* DIM*).* Some diminutives are class-changing, like *groentje (greenhorn)* from the adjective *groen (green)* and *praatje (talk or lecture)* from *praten (to talk).*

There certainly are examples of diminutives that meet many of the requirements for being inflectional. A form like *huisje (little house)* is transparent in meaning and of the same word class as its base. But whether a morphological process should be categorised as inflectional or derivational should be decided on the basis of all cases. What is needed is a morphological system that can capture all diminutive forms, not only the regular ones.[3]

Method

As in the previous experiment, each prompt was presented eight times in each set, as opposed to five times in other experiments. Apart from that, the procedure was the same as in Experiment 1.

Materials Construction of materials was done according to the procedure described under Experiment 1. The three diminutive forms that were used as a basis for constructing the variable sets were: *huisje (house* DIM*), vakje (box* DIM*),* and *zuurtje (sourball* DIM*).* These words were combined with two

[3]This is why the conjugation of verbal forms is definitively inflectional. There is to my knowledge no conjugated form whose meaning has moved away from the composition of base form and affix.

Table 4.4: Mean production latencies (ms), error percentages, and preparation effects (Δ) for Experiment 3: Diminutives

	Context				
Word Type	homogeneous		heterogeneous		Δ
variable	653	*3.0%*	692	*2.7%*	39
constant	679	*6.4%*	728	*3.1%*	50

compounds that shared the first morpheme with the diminutive forms. For the constant sets, three additional such compounds were selected. See the Appendix for the full set of materials.

Results and Discussion

The mean production time latencies for this experiment are shown in Table 4.4. The overall error rate was 3.8%. In the variable condition, subjects were faster in the homogeneous sets than in the heterogeneous sets, yielding a preparation effect of 39 ms. A slightly larger preparation effect was obtained in the constant condition, 50 ms. There was a significant overall preparation effect, $F'(1, 17) = 11.45, MS_e = 29146, p = .003$. No other factors were significant. Of most importance, the preparation effect for variable and constant Word Type did not differ: $F'(4, 23) = 0.81, MS_e = 7084, p = .539$.

A morphological preparation effect was found in both conditions, implying that metrical structure has no influence on this preparation effect. The results obtained in Experiment 1 cannot be due to deviant metrical structure and must be caused by a difference in inflectional frames.

Experiment 4: Phonological overlap

In Experiment 2 and 3, two possible counterarguments to Experiment 1 were refuted. Now it is time to examine the effects found in that experiment more closely. In Experiment 1, all response words started with the same morpheme and with the same segments. In the variable sets, a mix of verbs and nouns (compounds) was used which lead to a much smaller preparation effect than in the constant condition. It was argued that the metrical frame difference introduced by mixing verbs and nouns eliminated the morphological preparation effect: Lacking an inflectional frame to store the morphemes in, none could be prepared.

In the current experiment, all response words are formed from different morphemes (and most are actually monomorphemic). There is no morphological overlap between the responses, only phonological overlap in the first syllable. The purpose of this experiment is to test whether the inflectional frame has an effect on morphological preparation effect only, or also on phonological preparation effect. Again, there will be a variable condition with a mix of verbs and nouns, and a constant condition with only nouns. The variable condition contains one past tense and two monomorphemic nouns, as in *waadde, wapen, wasem (waded, weapon, mist).* The verb and nouns share initial segments, but their inflectional frames are different. There is only *segmental* overlap between the response words, and a *phonological* preparation effect is expected. The constant condition will contain only nouns and resembles the sets for which phonological preparation was originally found (Meyer, 1990). On the basis of a basic modularity assumption, morphological differences should have no effect at the phonological level. The prediction is that the preparation effect will be same in both conditions, despite the inflectional frame difference in the variable condition.

Method

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

Materials Three past tense verb forms were selected to form the basis for the homogeneous sets (*heette, pootte, waadde,* respectively *(be) called, planted, waded*). Each verb was combined with five bisyllabic, monomorphemic nouns that shared two initial segments (1 syllable) with the verb. Three of these nouns made up the response words for the constant condition, the two remaining nouns and the verb were the response words in the variable condition.

It should be noted that the singular past tense form *waadde* is homophonous with 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 past tense. In Experiment 6, we will address another form of homophony that occurs with some of these forms.

Results and Discussion

In this experiment, 2.2% of the observations was removed because of errors. From the results (Table 4.5) it is clear that the variable condition with the verb showed a much smaller preparation effect than the comparable condition with only nouns (30 vs. 67 ms). This is supported by the statistical tests: There was a significant overall preparation effect, $F'(1,20) =$

Table 4.5: Mean production latencies (ms), error percentages, and preparation effects (Δ) for Experiment 4: Phonological overlap

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

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 $F'(1,21) = 28.53$, $MS_e = 11032$, $p = .000$

Essentially, what is obtained is an effect of the inflectional frame on the phonological encoding process: Although there still is a phonological preparation effect in the variable condition (the 30 ms is significant in the analysis of simple main effects), the effect is much smaller than in the constant condition. This result is counterintuitive and might have far reaching consequences, because it violates the modularity of the morphological encoding (the level where the inflectional frame resides) and phonological encoding (the level where phonological preparation takes place).

Therefore, a replication study was done which is reported in the next chapter (Experiment 7). This study contained mostly new words that were selected using the same constraints used here. The results of the replication were very similar to the present experiment (29 vs 91 ms preparation effect in variable and constant sets). See Chapter 5 for details.

A timing proposal for phonological preparation

How can we account for the fact that the availability of an inflectional frame seems to influence the workings of the phonological encoding component? As said, under a modular view on language production, there should preferably be no dependency between two components that are concerned with different subtasks. No breach of modularity is necessary by assuming that the true effect of phonological preparation is masked if there is no shared inflectional frame. Only in the conditions with preparable inflectional frames the production latency fully benefits from the effect of phonological preparation and this explains the difference that was found between the reaction times in the variable and constant sets.

The explanation for this experiment builds on the timing proposal given

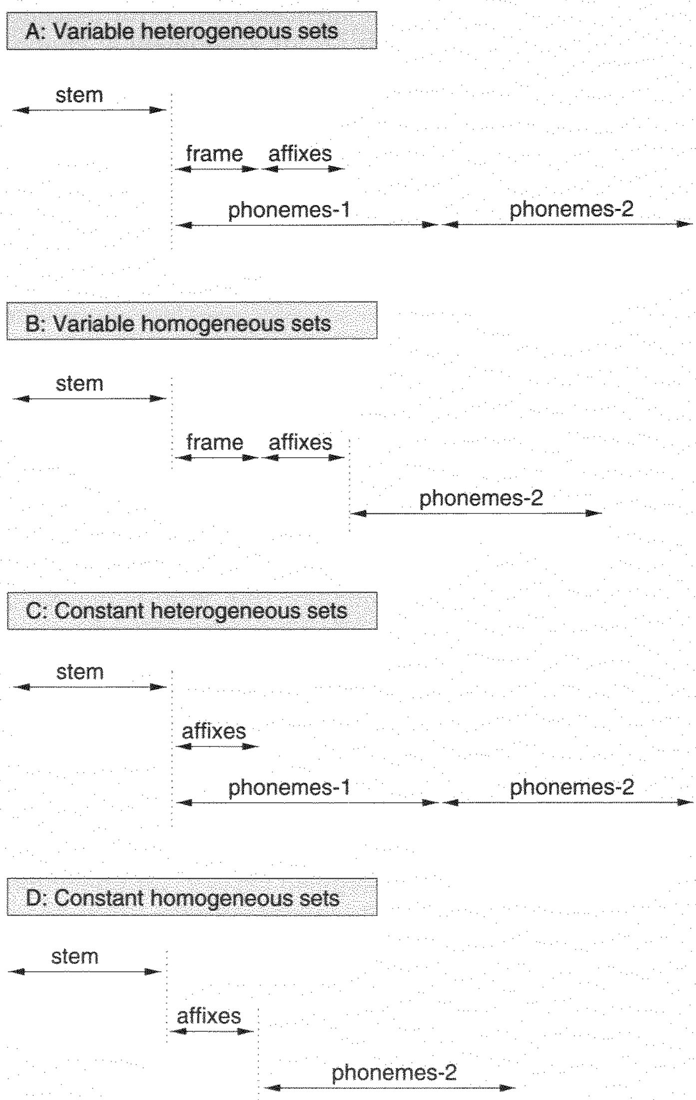


Figure 4.3: Ordering of on-line processes in Experiment 4. Phonological encoding has been split into *Phonemes-1*, the part that can be prepared in the homogeneous sets, and *Phonemes-2*, the part that always has to be done on-line.

for Experiment 1 on page 72. First, the stem (there is only one stem morpheme involved here) is retrieved (Panel A in Figure 4.3). Secondly, the inflectional frame is generated from the word class information of the lemma. As soon as the frame is available, the affixes can be retrieved. For the verbs in this experiment, a past tense affix has to be generated (eg. *te* in *heet+te*). For the nouns, there was no discernible affix, but the affix retrieval process has to be run to code 'no affix' (eg. *hekel+ø*).

Phonological encoding starts as soon the stem is available. In normal speaking situations, the affix morphemes are early enough and phonological encoding never has to wait for them. This is shown in Panel A. The part of phonological encoding that is labelled *phonemes-1* is the first part of the word, which is predictable in the homogeneous conditions. During *phonemes-2*, the second part is encoded, which differs depending on the response word uttered.[4]

The variable–homogeneous condition is shown in Panel B. In this condition, phonological encoding of the first part of the word can be prepared and hence the step *phonemes-1* is removed from Panel B. The phonological encoding of the second part of the word has to wait until the affixes are known (shown by the dashed line) and this prevents the encoding of the second part of the word *(phonemes-2)* to start right after stem retrieval. As before, affix encoding itself has to await the availability of the inflectional frame and this imposes an additional delay.

In the constant sets, the inflectional frame was predictable both in the homogeneous and heterogeneous contexts: all words were nouns. This leads to preparation of the frame in a heterogeneous context, as shown in Panel C. Because the encoding of *stem*, *phonemes-1* and *phonemes-2* determine the production time, that is, there is no effect of frame preparation per se. For constant–homogeneous sets however, *phonemes-1* can also be prepared, as shown in Panel D. Again, phonological encoding of the second syllable *(phonemes-2)* has to wait until the affixes are generated. But because of frame preparation, the affix encoding can be done at an earlier stage, leading to a larger preparation effect than that obtained in the variable sets (67 vs 30 ms in the experiment; the figure is not to scale).

Experiment 5: Plural forms

There is still another possible alternative explanation for the data we have presented. A distinction made in many theories of the comprehension of morphologically complex words is that between storage and computation. I will follow the proposal of Schreuder and Baayen (1995) here. According

[4] In this experiment, the first part of the word is also the first syllable. The wider term 'part of the word' is used because the explanation also holds for homogeneous conditions which share less or more than the first syllable.

to these theories, some complex forms are stored in the lexicon and in the process of language comprehension, their meaning is retrieved by look-up. High-frequent, opaque words like *department* are obvious candidates for storage. The meaning of other forms has to be computed on-line from their parts. Low-frequent transparent forms like *restructure* are likely to be dealt with this way.

In the first experiment, we compared compounds with inflected verbs. Compounds are widely assumed to be stored in the lexicon, and inflected verbs are supposed to be computed on-line. This means that there was a confound in the experiment.[5] The difference that was obtained might be caused by mixing a word that has to be computed on-line, with a set of stored words. One might further assume that the *odd* words used in Experiments 2 and 3 (adjectives and diminutives) are stored. Under this assumption, the absence of significant differences in these experiments can be explained by the absence of a storage–computation difference. In Experiment 4, words that had to be computed (past tense verbs) were inserted and there a difference between the variable and constant sets was found.

There are theoretical considerations on the relation between language comprehension and language production that suggest that the distinction between storage and computation, along the lines suggested above, has no crucial influence on our results and that the explanation in terms of inflectional frames has to be preferred. I will return to these considerations in the discussion of the present experiment. Theoretical arguments should, where possible, be supported by empirical data and an experiment testing the storage–computation hypothesis is easily constructed. We will make use of this opportunity to try to replicate the striking results of Experiment 4 with slightly different materials.

This experiment will mimic Experiment 4 to a large extent. Instead of singular nouns and verbs, plural forms will be used throughout. These plural forms have to be constructed on-line from the stem and a plural suffix. All words are thus computed on-line, and if the storage–computation hypothesis is correct, all sets in this experiment should behave in the same way. 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

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

[5]The storage vs. computation distinction is in many cases drawn along the same lines as the derivational vs. inflectional distinction. The interdependence of these two concepts makes it difficult, if at all possible, to avoid the one when manipulating the other.

Table 4.6: Mean production latencies (ms), error percentages, and preparation effects (Δ) for Experiment 5: Plural forms

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

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 only in the first syllable.

Three bisyllabic plural past tenses were selected: *haatten, raadden, doodden (hated, guessed, killed)*. Each verb was combined with 5 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 4.6 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 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 4 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. In the previous experiment, the number slots were empty, in this experiment, they were filled with the plural suffixes *-en* or *-s*, but no qualitative difference results from the presence of the nominal plural suffixes. The definition of the inflectional frame that was given in Chapter 3, and which stipulated a slot for every *possible* type of affix, is enforced. I will pay more attention to the theoretical side of this issue at the end of this chapter, and return to it again in Experiment 12 in the next chapter.

The theoretical reason for expecting the outcome of Experiment 5, which was hinted at above, has to do with the difference between models of language comprehension and language production. When we try to understand what is said to us, we are faced with the extremely difficult task of matching the input to precisely one of the thousands of words that are in our lexicon. To complicate matters even further, some forms might not even be in the lexicon because they are, for example, inflectional variants of other words that *are* in the lexicon. For the recognition of these inflectional variants, we need a process that can derive the meaning of a word form from its components (see McQueen & Cutler, 1998, for an overview).

The crucial point is, however, that such a mechanism is not needed for language production. When producing language, we already know what we are going to say. In the Levelt/Roelofs-model of language production, the lexicon does not have to be searched for a lemma matching a given lexical concept because a lexical concept is linked to the lemma (see also Janssen, 1998). Similarly, the form entry that matches a lemma can be easily found because the lemma contains a pointer into the form lexicon. Hence, in production it is not necessary to limit the number of candidates by, for example, decomposing lemma entries. At the form level, however, decomposition is everywhere: Even opaque forms like *apartment* are generated from their constituent morphemes *apart* and *ment*, reflecting the fact that language uses the same morphemes for many purposes (Roelofs & Baayen, submitted; Roelofs et al., submitted). Because meaning and form are fully decoupled, the production system can fully exploit the redundancy of language (Levelt et al., 1999, see also Dell, 1986, Figure 1).

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 delay on the affix encoding stage (cf. Figure 4.3). More evidence for this hypothesis will be reported in the discussion of Experiment 12.

Experiment 6: Strong past tense

Because the preparation difference found in Experiment 4 is attributed to a difference in inflectional frames, an obvious prediction is that we will not find such a preparation difference when mixing verbs and nouns that have the same inflectional frame. Ideally, the materials of Experiment 4 should be replicated as closely as possible. Given that verbal inflectional affixes were claimed to apply to all verbs, it might seem contradictory to look for verbs that have the same inflectional frame as nouns. There is, however, one class of verbs that one can suspect of having exactly this property: strong past forms.

Dutch has a small but frequently used class of strong past tense verbs (also called irregular verbs), for which the past tense is signalled by choosing a variant form of the stem (a so-called allomorph, an English example is the pair *swim, swam*). Most past tense forms were historically derived from the present tense by an *ablaut* rule (see Hare & Elman, 1995, for an interesting simulation of these historical processes), but this process is no longer productive.

Because the strong past tense forms are inherently marked for tense, there is no need for the tense affix slot that normal verbal forms possess. The strong past forms still need a number slot, and this makes their inflectional frame exactly identical to that of a noun. Thus, similar effects for nouns and strong past forms are to be expected.

(4.4)	Nouns:	Stem	+	Number		
	Strong Verbs:	Stem	+	Number		
	Regular Verbs:	Stem	+	Tense	+	Number

Of course, the assumption about the absence of a tense slot might be proven wrong. If inflectional frames are assigned solely on the basis of word class, without room for any exceptions, strong past tense forms will have a tense suffix slot too. If this is the case, the same pattern as in Experiment 4 will be obtained.

Method

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

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, the number of items in this experiment was increased.

There were 4 practice sets and 16 experimental sets. Four strong past tense forms were selected to form the basis for the experimental sets: *kreeg (got), spoot (squirted), droeg (carried),* and *sliep (slept).* For the response words in the variable condition, the past tense and three additional monomorphemic, monosyllabic nouns that shared the onset cluster with the past tense were used. For the constant condition, four additional words were selected according to the same criteria.

Some Dutch nouns are homographic with first person, present tense verbal forms. I tried to avoid using these nouns, but seven cases had to be included to complete the materials. It was ascertained that they were paired with prompts that activate the nominal reading. For 6 of them, the nominal frequency was also much higher than the verbal frequency (averages for these 6 were, respectively, 1536 and 529 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 3 prompts were presented 5 times each. In all other respects, the construction of the materials was the same as in Experiment 1.

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

Results and Discussion

Table 4.7 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, $F'(1,21) = 10.15, MS_e = 79387, p = .004$. No other effects were significant. 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. Simulations by Santa et al. (1979, Table 3) have shown that for the observed heterogeneity and model parameters, H_0 will be rejected more often (observed $\alpha = .068$). Despite this inflated α-level, we found no significant effect of Word Type.

The results are congruent with the assumptions. The equivalence of the effects for both conditions shows that, when the inflectional frame is predictable, initial segments of a set containing nouns and one verb can be prepared without difficulty. The analysis of strong 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 or morphological preparation (cf. Experiment 2). 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 in terms of verb–noun differences.

Table 4.7: Mean production latencies (ms), error percentages, and preparation effects (Δ) for Experiment 6: Strong past tense

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

Experiments 4, 5 and 6 support the following view on phonological preparation: Regardless of mixing word classes, of using plural or singular forms, or of inflectional frame differences, phonological preparation can always take place. If there is a difference in inflectional frames, however, the word's frame has to be made available on-line, and this causes affix checking to start late. This will mask part of the benefit the subjects experience from phonological preparation.

General Discussion

In line with what was expected from the speech error data, we found evidence for the existence of inflectional frames in language production in **Experiment 1**. Subjects can prepare the initial morpheme of a word, but only if they can prepare the inflectional frame also: In a condition where two compound nouns and a verb shared their initial morpheme, no *morphological* preparation was found, because the compounds and the verb differ in inflectional frames. This difference prevented the subjects from preparing the inflectional frame, which in turn made it impossible for them to prepare the morpheme: Lacking a frame to store the morpheme in, no effect of the shared morpheme could be found.

Apparently, the inflectional frame is needed to attach the shared morpheme to. There is no reason to limit this finding to the task situation; the speech error data that inspired the frame theory strongly suggest that during normal speaking an inflectional frame is required to store stems and affixes as they come available.

I will now try to tie all results together and fill in the details of the morphological encoding process. The data support a division of the morphological encoding process into two stages, as was conjectured in Chapter 3. More specifically, I now propose the further specification of the morphological encoding component into a *stem encoding component* and an *inflectional encoding component* (Figure 4.4), which deal with derivation and inflection respectively. The first subcomponent creates the morphological representation of the stem and handles compounds, derivations, and irregular forms. The second subcomponent combines the stem with the inflectional suffixes, and feeds the resultant morphemes to the phonological encoding process. I will first deal with this latter stage and with the nature of the inflectional frame.

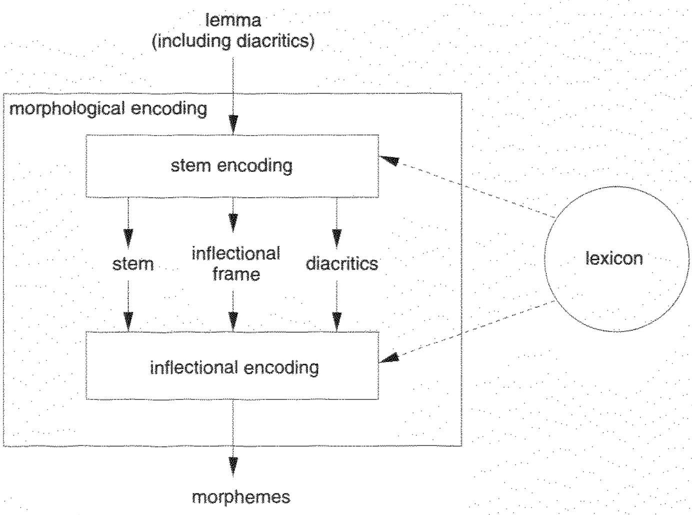


Figure 4.4: Two stages within the morphological encoding process

The nature of inflectional frames

How many slots does the inflectional frame contain for a given form? Throughout the chapter, it has become clear that our definition of the inflectional frame containing one slot for the stem and one slot for *every possible type of inflectional affix* (Chapter 3) was correct. Most telling was **Experiment 5**, where the materials of an earlier experiment with singular nouns and verbs (Experiment 4) were modified: All forms were pluralised and the experiment was run again. The same results were obtained as before, showing that the actual number of suffixes does not influence the effect of the inflectional frame on the preparation effect.

What types of slots can be specified by inflectional frames? A very minimal view on this was presented in the previous chapter, which allowed only for stem, prefix and suffix slots. **Experiment 2** supplied some evidence for this view. There, the inflectional frame for an adjective and for a noun turned out to be equivalent. Adjectives bear a gender suffix, while nouns are inflected for number. If their frames are alike, this can only be due to the fact that slots are not labelled for content, that is, type of inflection. A generic labelling in terms of stem, prefix, and suffix conforms with these data. Note that this is in accordance with the speech error data that was surveyed in Chapter 2, which points to a different status for stems and inflectional affixes.

This issue certainly has to be researched more systematically before one can make definitive claims about it. One theoretical point, however, can be addressed already: If all suffix slots are the same, how come the suffixes end up in the right order? For segments, the WEAVER model (Roelofs, 1997c) assumes numbered links that express the order in which the segments appear. A similar approach can be taken here, with the additional benefit that in Dutch, English, and other Indo-European languages, inflectional suffixes occur in a fixed order throughout the language (Anderson, 1992, p. 123). This means that, for Dutch, tense suffixes can all have number 1, and number suffixes number 2. This solves the question where the information on the labelled links comes from in a simple way.

Inflectional encoding

In the current proposal, inflectional encoding is parallel in two aspects: First, activation of all required affixes is done in parallel. Affixes are selected to express the meaning of the diacritics; the diacritic *tense = past* will be expressed as *-ed* in English. Note that we distinguish between activation of the possible candidates, which is done in parallel for all inflectional slots, and insertion of the correct candidate in a slot, which is done serially. This distinction is a characteristic of the WEAVER model of speech production.

Second, as soon as an inflectional affix is put into a slot, the phonological encoding process will start to retrieve the corresponding segments. The inflectional frame is the single device responsible for the synchronisation of inflectional encoding (affix retrieval) and phonological encoding (conversion of affix morphemes into segments). The phonological encoding process is allowed to work on morphemes inserted into the inflectional frame, up to the rightmost filled slot. It is ensured that phonological encoding proceeds from left to right and the process will be suspended if it encounters a slot that is still empty. After this slot has been filled by a morpheme, the phonological encoding process is resumed. When no more empty slots remain, phonological encoding can continue until the end of the word. The WEAVER model has a built-in *suspend–resume* mechanism that allows for precisely this kind of coordination.

A zero morpheme is needed for all inflectional processes that have 'no affix' as one of their realisations, like regular present tense in Dutch and English. The slot cannot be left open, for the phonological encoding process would never be resumed. For regular English verbs, the present tense is mapped to *-ø* and the past tense to *-ed*. This is in accordance with one of the criteria for inflectional expression formulated by Bybee: "Further, in inflectional expression, the lack of a marker must be interpreted as meaningful (as zero expression) rather than as the absence of the category." (Bybee, 1985, p. 27)

Stem encoding

The account given above cannot explain the behaviour of non-inflectional forms. Both compounds and derivations give rise to a morphological preparation effect, but their derivational structure does not have repercussions for the inflectional frames. This was shown in **Experiment 2**, where adjectival derivations and compounds gave rise to effects of equal size, and in **Experiment 3**, where the same was found for diminutives and compounds. From the similar preparation effects in the conditions of these experiments, it was concluded that there is no frame difference between derived or diminutive forms and the compounds that have been used for comparison throughout.

An alternative solution is to disregard the linguistic difference between compounds and derivations and posit a frame with two stem slots. This second stem slot would then either hold the second part of the compound, or the derivational suffix. This solution is less elegant than the one proposed earlier precisely because it neglects the difference between compounds and derivations. In this study we found experimental evidence for the distinction between inflection and derivation and it would be unwise to treat compounds and derivations on a par without the necessary empirical evidence. More fundamentally, the number of stem constituents or derivational affixes is, at least in principle, unlimited and more stem slots would need to be introduced for *railroad agency* or *interdenominational.*

Instead, I want to propose that derivations and compounds are dealt with in the first stage of morphological encoding, *stem encoding.* In the most straightforward case, stem encoding is simple because the concept to be expressed belongs to a monomorphemic word. One morpheme is retrieved from the lexicon and passed on to the inflectional encoding stage. If, however, the concept happens to be expressed by a compound in the language, the lemma points to two morphemes (eg., *work* and *bench,* see also Levelt et al., 1999) and these two morphemes are concatenated. Together they will form the basis for inflectional encoding. A similar complex stem representation is obtained when the word is expressed by a derivation. The lemma with the meaning 'the state takes over major industries (verb)' has as its form representation a list of three morphemes: *nation+al+ize.*

Next, the stem encoding process constructs the inflectional frame on the basis of the word class. Word class is present on the lemma and it can be used to compute the inflectional frame for all regular Dutch words, by applying the following simple rules:

(4.5)	Noun	→	Stem + Suffix
	Verb	→	Stem + Suffix + Suffix
	Adjective	→	Stem + Suffix

Finally, the stem encoding process passes on the diacritics from the lemma to the inflectional encoding stage, together with the stem and the inflectional frame.

The ordering of the processing steps

In the discussion of **Experiments 1 and 4**, a proposal was made for the timing of the various processing steps in morphological encoding (Figures 4.2 and 4.3). How do the processing steps in this proposal relate to the division of the morphological encoding component into a stem encoding and an inflectional encoding component? This is schematically shown in Figure 4.5: The retrieval of the stem morphemes (steps *stem-1* and *stem-2*) is done in the stem encoding component, and so is the generation of the inflectional frame (step labelled *frame*). The encoding of inflectional affixes is the sole responsibility of the inflectional encoding component. All further steps involve phonological encoding and these are carried out by the next lower component in the Levelt/Roelofs-model.

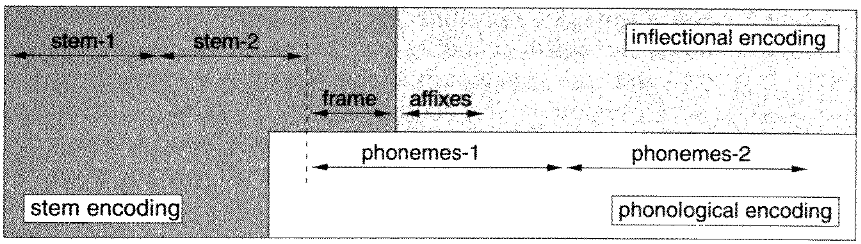


Figure 4.5: Assignment of processing steps used in Figure 4.2 to stem encoding, inflectional encoding, and phonological encoding.

Apart from concretising the course of events in producing morphologically complex words, the timing explanation was specifically developed to account for the results obtained with words that only overlapped in initial phonemes. An influence of the inflectional frame on the phonological preparation effect was found in Experiment 4, where nouns overlapping in their first syllable were mixed with a verb that shared the same initial syllable. A direct influence of a *morphological* entity like the inflectional frame on *phonological* processing would violate a module boundary in the Levelt/Roelofs-model: As was discussed in Chapter 1, separate levels in a modular model communicate only via designated interface representations. The inflectional frame is not necessary for phonological encoding and should not be passed on to or, *a fortiori*, directly influence this lower level.

It was proposed that the inflectional frame difference *masks* part of the phonological preparation effect in the variable condition. In this way, there is no direct influence of the frame on phonological encoding, but we can still account for the results. The solution is schematically shown in Figure 4.3 on page 82. The crucial dependencies are between frame generation (arrow labelled *frame*) and affix retrieval (arrow labelled *affixes*), and between affix retrieval and phonological encoding of the second syllable *(phonemes-2)*. In the homogeneous condition with variable inflectional frames (Panel B of the figure), the first syllable can be prepared and the encoding of the second syllable can start earlier than in the matching heterogeneous condition without any initial overlap (Panel A of the figure). It can, however, start no earlier than after frame generation and affix retrieval, because the affixes form part of the second syllable of these particular words. This dependency reduces the benefit of preparing the first syllable. Compare this to the homogeneous condition with constant frames (Panel D). Here, the inflectional frame is known in advance and the encoding of the second syllable is only delayed by affix retrieval. The net result is a larger preparation effect in the conditions with constant frames than in the condition with variable frames, as was empirically obtained.

There are still some open questions on the cooperation between morphological and phonological encoding and the nature of the preparation effect. In a condition with constant inflectional frames in Experiment 1, a morphological and a phonological preparation effect was found for the first morpheme of compound words. When the inflectional frame was variable, similar morphological overlap gave rise to phonological preparation alone. The occurrence of phonological preparation in this condition cannot easily be explained by the approach taken earlier, in which a preparation effect was caused by partial input: No morpheme can be prepared in the absence of a frame, but the same morpheme must form the partial input which allows for preparatory phonological encoding.

If we turn our attention to Experiment 4 and its follow-ups, the same problem arises. The response words exhibit only partial phonological overlap (eg. /kr/ in *krent* and *krant*) and we have no precise theory on how the phonological preparation effect comes about in this case. It is certainly not the case that it is mediated by a common morpheme, because all response words were monomorphemic and hence uniquely determined by their morpheme. There are many possible ways to account for the results and further experimentation has to make clear in which direction the correct solution is to be found.

Irregular forms

The process sketched above holds only for regular forms. Irregular forms, like strong verbs, require special treatment even at the time of lexical retrieval. When following the link from the lemma into the form lexicon, a complex lexical entry will be encountered for these forms. Normal lexical entries specify precisely one morpheme. This will be diagramed as shown in (a):

(4.6)

	lemma		condition	morpheme	inflectional frame
a.	CHAIR	→		**stoel**	
b.	TO SWIM	→	*present:*	**zwem**	Stem + Affix
			past:	**zwom**	Stem + Affix
c.	MUSEUM	→	*singular:*	**museum**	Stem
			plural:	**musea**	Stem

This entry has no value listed under *inflectional frame* because the frame for normal words is predictable from syntactic class and hence not listed in the lexicon. For Dutch nouns, the frame contains two slots (stem + affix). A complex entry for a strong verb specifies one morpheme for *tense = present* and one morpheme for *tense = past*, as in (b). The stem encoding process checks the diacritics and takes the correct morpheme. The diacritics used for selection of the correct stem morpheme (in this case *tense*) will not be passed to inflectional encoding. Otherwise, a double marked form like **zwomde (*swammed)* would result. This process will be called *diacritic absorption.*

It is easy to see how nominal irregular nouns, like *museum–musea*, can be dealt with in the same way. All that is needed is a lexical entry as shown under (c). The diacritic for number will be absorbed by selecting the right morpheme, precluding **museas.*

We have seen that the tense diacritic was absorbed by using it to select the correct allomorph of the stem. For the past tense, this absorption leads to the correct absence of a regular past tense suffix. For the present tense, this means that there is no diacritic left to generate the unnecessary *-ø*. But the systematic absence of a tense affix cannot go together with a tense slot in the inflectional frame: As was explained above, the inflectional encoding component will pause indefinitely when an empty slot remains. A frame without a tense slot is the solution here. It also explains the outcome of **Experiment 6**, where no frame difference between nouns (with one suffix slot) and strong verbs was found: Apparently, strong verbs have an irregular inflectional frame with only one suffix slot (for number), which is comparable to the frame for nouns.

Because it was stipulated that the inflectional frame cannot have empty slots, the frame of a strong verb must be irregular, and the usual procedure for computing the frame from word class cannot go through. Instead, the inflectional frame is retrieved from the lexicon, where it is stored in the complex lexical entry that holds the allomorphic forms. If an inflectional frame is present at the lexical entry, application of the normal rule is blocked. In (b) and (c), the irregular frame that has to be used instead is listed in the last column.

These assumptions put constraints on the regularity of derivational forms. The word form of a derivation like *nationalise* is made by combining three morphemes. By virtue of the 'right hand head rule' (Williams, 1981), the rightmost morpheme defines the properties of the combination. Indeed, *nationalise* is a verb, like all words formed with *-ise.* And if it were an irregular form, this information has to be stored at the lexical entry for the morpheme *-ise*. This would make all derivations on *-ise* irregular. Dutch has a suffix that shows that this is indeed the case: the suffix *-heid* makes an abstract noun from an adjective. If these nouns allow pluralisation, it is always irregular: *vrijheid, vrijheden (liberty, liberties; literally free+dom).*

The present take on stem encoding entails a word-based morphology (Anderson, 1992; Beard, 1995): The meaning of a derivation or compound is not derived from its parts, but stored at the word level. The meaning of a derivation is thus free to move away from its origin, which is one of the defining properties of derivation. This also allows us to rigorously use all form overlap for sharing morphemes, even for opaque compounds like *makeshift*.

Inflections, on the other hand, are restricted in their meaning. Because all inflectional variants of a word are formed on the basis of one lemma, they all have the same basic meaning, to which the inflectional ending can add only a meaning nuance. This is reflected in the proposed separation between stem encoding and inflectional encoding (Figure 4.4), following the split-morphology hypothesis.

Further findings with inflectional frames

CHAPTER 5

In the previous chapter, evidence was presented for the existence of an inflectional frame in language production. There are several questions and problems that have not been addressed yet, and in this chapter I will make a stab at resolving some of those. First, the surprising results of Experiment 4 will be further examined (in Experiment 4, variability of inflectional frames appeared to reduce the phonological preparation effect). A replication study (Experiment 7)is carried out with partly new materials, which, as we will see, confirms the outcomes of the original experiment. A variant of the same materials will be used in Experiment 8 to show that it really is the verbal frame that caused the results of Experiment 4.

Next, we will focus again on the definition of the inflectional frame. The inflectional frame has one slot for every possible type of inflectional affix, plus one for the stem. In Experiment 9, it will be examined whether it is true that there always is precisely one stem slot, even for forms with inbuilt complexity like compounds. It was also claimed that the inflectional encoding component maintains a strict left-to-right order in the encoding of the slots. This claim will be put to test by researching whether suffix slots can be coded before the stem slot.

Two further experiments (9 and 12) will extend on results obtained in the previous chapter. The experiment with strong verbs is rerun, now using the present tense forms of these verbs. This way it can be found out whether irregularity resides in the past tense forms only or whether the whole paradigm is receives special treatment. Also, in the previous chapter, a reduction of the phonological preparation effect was obtained for singular nouns and verbs, and for plural variants of the same materials. In Experiment 12 we will test whether mixing plural nouns with singular verbs will lead to a different result.

Experiment 7: Replication

In Experiment 4 in the previous chapter, the question was addressed whether a frame difference influences the phonological preparation effect. To invoke phonological preparation, words were used that overlapped in the first syllable. The frame difference was caused by mixing nouns (with 2-slot inflectional frames) and verbs (with 3-slot frames). The result of Experiment 4 was rather counterintuitive: The inflectional frame difference caused a reduction of the phonological preparation effect. Under a modular perspective on language production, there should be no influence from the morphological level (the predictability of the inflectional frame) on the phonological level (where the phonological preparation occurs). The results of that experiment were statistically very reliable, but there is no better confirmation of a finding than a replication of the experiment.

Words similar to those used in Experiment 4 will be used here. When looking for new materials, I spotted an interesting trend: It seems to be the case that all nouns that contain an initially embedded verbal stem, are also morphologically related to that stem. As an example, given the stem /wak/ (from *waken*, 'to wake') no morphologically unrelated form like **wakem* could be found in the CELEX lexical database for Dutch (Baayen et al., 1993). All nouns that start with /wak/ are derivations from *waken*: *wake (wake)*, *waker (guard)* and similarly for adjectives: *waaks (watchful)*. A little survey was carried out for the eight verbs that start with the /wa/ onset. Eight nouns were found that overlapped with, but were not (synchronically) related to any of the verbs. Of these, two were archaic nouns, two were loanwords, two were derived forms from an overlapping country name (*Walonië*, 'French speaking part of Belgium'), and only two were nouns that had an unrelated verb embedded in them.

Although the sample is small and there is no control group to compare this result to, the suggestion that language disfavours nouns that have an embedded unrelated verb stem is compelling. Language comprehension would of course benefit greatly from such a constraint. If a dual route model, like that of Schreuder and Baayen (1995), encounters the form /wadi/ (*wadi*, a loan word), it will attempt to parse the form as /wad+i/, where /wad/ is the stem of *waden (to wade)*. This particular attempt will eventually fail and so will most other misparses of the input. Still, considerable effort will be spent betting on the wrong horse. By trying to avoid nouns with embedded unrelated verb stems, the language system would reduce a possible source of confusion. Further research is necessary to find out whether these unrelated-embedded nouns are indeed less common than would be expected.

Method

The methods for this experiment were similar to those used for Experiment 4. 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 4 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 that have a stem matching the following pattern: *Initial consonant(s) – long vowel –* /t/. The final /t/ can appear as *t* or *d* in orthography, because of final devoicing in Dutch. The past tense was formed by adding *-te* or *-de*, depending on the underlying voicedness of the coda consonant. Metrically, this leads to forms like /wa-də/; the stem-final consonant appears in onset position of the second syllable. This subclass of verbs was chosen to enable the selection of nouns like /wa-pə(n)/ (*weapon*), that overlap for the complete first syllable with the verb. Verbs with other coda consonants show a different metrical pattern and it is generally impossible to find nouns overlapping with their full first syllables (see above).

Even for this subclass, only few overlapping nouns exist that are suitable for experimentation: Only two of the three sets from Experiment 4 could be replaced. The verb *heette* and its accompanying nouns on /he/ were kept in this replication, with the modification that the one proper noun in its set (*Hema*) was replaced by an adjective (*hevig*). I had to resort to using this adjective because not enough nouns starting with /he/ could be found. Experiment 2 rendered evidence for a similar inflectional frame for adjectives and nouns, and this finding suggests that using this adjective was not optimal but certainly feasible.

The two new sets were centred around the verbs *doodde (killed)* and *raadde (guessed)*. 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.

Results and Discussion

Overall, 3.8% errors were made. The results are summarised in Table 5.1. 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 4. The large effect leads to low probability values: The overall preparation effect was significant at $F'(1, 18) = 11.68$, $MS_e = 48855$, $p = .003$ 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 4 were replicated with new subjects and mostly new items. The result confirms the timing explanation that was suggested on page 82. The preparation effect in the constant condition was surprisingly large, but relatively large effects have been found with all *odd man out*

Table 5.1: Mean production latencies (ms), error percentages, and preparation effects (Δ) for Experiment 7: Replication

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

experiments, when compared to, for instance, the original experiments by Meyer (1990). See the discussion of Experiment 8 for related comparisons.

In general, comparisons of preparation effect size cannot be made between experiments. The effect size depends on many factors, and not all are systematically controlled for. The comparison to Meyer's experiments, for example, should be done with the proviso that the results she reports are averages over three repetitions of the whole experiment, while the present results are based on only one such repetition. Furthermore, items are certainly one of the more influential factors on the preparation effect, and this is why the item sets were carefully constrained. Different experiments are subject to different constraints, and are thus not always comparable.

The constraints on the items in this experiment and in Experiment 4 were exactly the same; this warrants statistical analysis of the differences between the two experiments. Two questions will be answered: First, we want to find out whether the effect in the constant set of Experiment 7 (91 ms) was larger than the constant effect in Experiment 4 (67 ms). Second, we want to compare the outcomes for the sets containing *heette* in both experiments.

To answer the first question, the constant sets from the two experiments were submitted to an ANOVA with two factors: Experiment (4 or 7), and Base within Experiment (the three possible word onsets in each experiment). The result showed 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 analyse and the factor Experiment was again non-significant at $F'(66,17) = 0.25, MS_e = 25175$ while there was an overall preparation effect, $F'(1,17) = 5.81, MS_e = 24963, p = .026$

The second question was tested by selecting the words on /he-/ from both variable conditions. The data were subjected to an ANOVA with Experiment as the only factor. A quasi F was computed that turned out to be non-significant: $F'(17,6) = 0.35, MS_e = 7434, p = .958$. This means that there was no significant difference between the two sets containing words starting with /he/.

Experiment 8: Homophones

Some of the past tense forms that were used in Experiment 4 and 7 are homophonous to nouns. The form *waadde (waded)*, for example, has the same pronunciation as the word *wade (shroud)*.[1] In the previous experiments, it could be ensured that subjects were using the verbal reading by the orthography (which reflects the inflectional rules for verbs) and by choosing a prompt that was strongly related to the verbal sense.

In the present experiment, this homophony will actually be used to check whether the results obtained in Experiment 4 and 7 were indeed due to the fact that verbs appeared as the odd men out. A combination of the materials of Experiment 4 and 7 will be tested, with one crucial modification: The orthography of the past tense items will be changed such that they will now be read as their nominal homophones. The prompts that are tied to these homophones will also be changed such that they enforce the nominal reading. For verbal *waadde (waded)*, this means the response words will be changed to *wade (shroud)* (both pronounced /wa-də/), and the prompt will be changed from *river* to *priest*.

The prediction is that a similar preparation effect will be found in both conditions of this experiment: The homophonic nouns will not be treated any different 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 4 and 7, 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)*. The last noun is a plural form. No alternatives were available among the materials used in the previous experiments, and considering the results of Experiment 5 and Experiment 12 that will be reported shortly, no possibly disturbing effect from the one plural form is expected. I decided against using

[1]The homophones are of a supposedly small class of unrelated-embedded nouns, see the introduction to Experiment 7.

Table 5.2: Mean production latencies (ms), error percentages, and preparation effects (Δ) for Experiment 8: Homophones

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

a homophone of a verb that was not used in Experiment 4 and 7, to be able to compare the same phonological strings in the two conditions.

New prompts were selected for the homophonic nouns that enforce the nominal reading of the word. This was not possible for *dode (corpse),* because it is semantically related to the past tense homophone *doodde (killed).* The complete materials are listed in the Appendix.

Results and Discussion

The results of the experiment are shown in Table 5.2. 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 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 some variance between subjects and the fact that the items starting with *po* did not show a substantial preparation effect across the board, see Figure 5.1 (the same items produced healthy preparation effects of 62, 43, and 58 ms in Experiment 4). The F' test is known to be a trifle conservative by nature and facing two source 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. Because these nominal homophones have exactly the same sound structure as their verbal counterparts, the critical difference must lie somewhere else. Word class per se has also been excluded as a factor in other experiments (eg. Experiment 2 and 6). This is taken as evidence that it was indeed the inflectional frame of

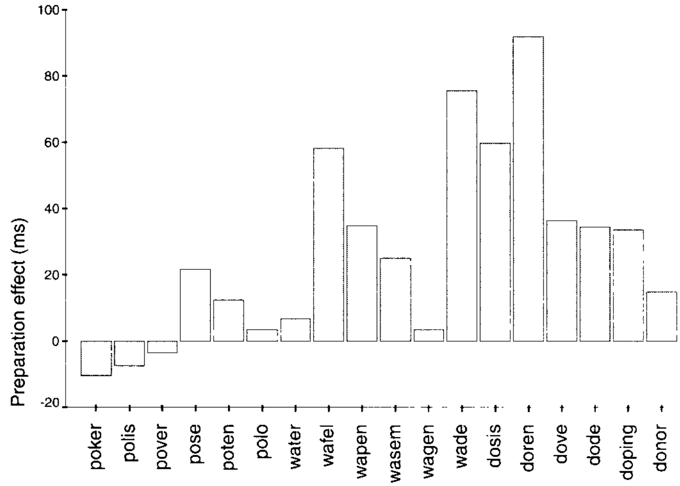


Figure 5.1: Mean preparation effect for the items in Experiment 8

the past tense forms that caused the reduction in preparation effect in Experiment 4 and 7, that was not found here.

It was claimed that the difference in inflectional frames in Experiment 4 caused the reduction of the preparation effect in its variable condition. Because there was no inflectional frame difference, no such reduction was found in the present experiment. When comparing the actual size of the preparation effects between the experiments, the effect in both conditions of this experiment is comparable to the reduced variable effect in Experiment 4 and 7. This may give the misleading impression of reduction in *both* conditions of this experiment. This cannot be right because the constant sets in this experiment consisted of items that also featured in experiments 4 or 7.

More data on this issue were gathered by looking at the results for the constant sets only. As said, similar criteria were used for selecting the nouns in this condition over the three experiments and we expect the results of these sets to be similar. As an additional constraint, only those subjects that saw the constant sets first were examined. These subjects have seen only one verb (or verb homophone) in the practice sets, and there is no a priori reason to expect them to respond differently to the constant set. The mean preparation effects over the three constant sets are listed in the last column of Table 5.3. It can clearly be seen that the average preparation effects did vary substantially.

Table 5.3: Mean CELEX response word frequency (per 42 million) and mean preparation effect (ms) for constant sets for those subjects that received the constant sets first.

Constant sets from		Mean frequency	Δ
Exp 4:	Original	2319	55
Exp 7:	Replication	327	80
Exp 8:	Homophones	2157	37

One reason for this might lie in the frequency of the response words. Roelofs (1998) has shown that high frequent response words give rise to a *smaller* preparation effect, presumably because producing the word is faster and the benefit of preparation is therefore reduced. For the construction of the replication experiment Experiment 7, additional items had to be found that met our constraints. There were not many of those and the mean frequency of the items is substantially lower than that of the original ones (column 3 of Table 5.3). Therefore, the relatively large preparation effect in the replication experiment can indeed be caused by a response word frequency difference.

The mean preparation effect obtained in Experiment 8 is, however, still smaller than that in the original Experiment 4. An explanation in terms of response word frequency cannot account for this because this measure is almost identical for Experiment 4 and 8. In Meyer's (1990) experiments, the preparation effect for one syllable overlap was of the same magnitude as the effect found in this experiment, and this suggests that the right question to ask is why the effects in Experiment 4 and 7 are relatively large. I can give no explanation for this and can only attribute this difference to noise, because hardly anything (apart from subjects) was changed between experiments 4 and 7. The effect size differences between the experiments are not large, though, and it is doubtful whether one should pursue an argument based on three data points.

Experiment 9: Suffix overlap

The original implicit priming experiments by Meyer (1990) have shown that a phonological preparation effect is obtained when all response words in a set had the same initial segments, but not when the response words shared final segments (eg. *thema–poema*, 'theme–puma'). This led Meyer to conclude that words are built up from left-to-right, and that preparation is thus restricted to initial segments. Under this hypothesis, no preparation effect should be obtained from response words containing the same suffix, because suffixes are non-initial. It has not yet been shown, though, that Meyer's findings extend beyond the encoding of the stem. Morphologically simplex words were used throughout in her experiments. Roelofs (1996, 1998) has used morphologically complex words, and found no preparation effect (phonological nor morphological) for the second morpheme of a compound noun like *deeg+rol* or a complex verb like *op+zoeken*. But we have seen before (Experiment 2) that the inflectional frame, the principal representation at the level of encoding of inflectional affixes, is not sensitive to stem complexity. Roelofs has shown that there is no preparation of non-initial morphemes of the stem, but this does not *a priori* preclude preparation of inflectional suffixes.

In this experiment, four different verbs will appear in a set that all bear the same past tense suffix. If words are built up from left-to-right, no preparation is possible. If stems and suffixes are built up from left-to-right, but are independent of each other, preparation of the segments of the suffix is possible, even though the verbal stem itself cannot be prepared. In this case, the production latencies in the homogeneous condition (with overlapping suffixes) should be shorter than those in the heterogeneous condition (without any overlap). This is the prediction that will be tested here.

To construct the necessary control conditions, three types of words will be used in the experiment: First, verbs with past tense inflectional suffixes (either *-de* or *-te*). These set will be called ME, for Morphological End overlap. Second, nouns with Phonological End overlap will form set PE (eg. *nikkel*, 'nickel' and *cirkel*, 'circle'). The third set PI, contains nouns with Phonological, word Initial overlap (eg. *kade*, 'quay' and *kano*, 'canoe'). From the literature, one expects no preparation effect for PE, and a significant preparation effect for PI. If inflectional suffixes are encoded completely separately from the stem, and their phonological encoding can start before the phonemes of the stem are known, the past tense suffixes can be prepared and condition ME is expected to show a preparation effect similar to that of PI.

Method

Participants Sixteen subjects participated in the experiment. They were randomly selected from the Max Planck subject pool and were paid for their efforts. For the three months preceding the experiment, they had not participated in any other experiment related to this thesis.

Materials Two sets of four past tense forms were selected. The first set contained verbs taking *-de* as past tense suffix, the second set contained *-te* verbs. Excluded were all verbs for which the orthography of the final stem consonant is changed in the past tense (eg. *verven* - *verfde*), all verbs that undergo the orthographic rule of consonant duplication (eg. *branden* - *brandde*), and verbs that have a complex consonant cluster in the past tense (eg. *snurkte*). For the two PE sets, nouns were selected that showed phonological end overlap in the syllables *ter* or *kel*. Lastly, four PI sets were constructed. In these sets the response nouns shared initial syllables with the phonemes *ka*, *re*, *ze*, or *boe*. All words were bisyllabic and stress initial.

The response words described above were combined with associated nouns that served as prompts. All prompts were of approximately the same length in letters, showed no initial overlap with their response word or with any other prompt in the same set, and were as dissimilar as possible. See the Appendix for the full set of materials.

Design Contrary to all other implicit priming experiments reported here, this experiment does not use the odd-man-out paradigm. The preparation effect caused by the three word types (ME, PE, PI) can be directly compared. The experiment is separated in two halves (subexperiments) because the *-de* and *-te* verbs were to be kept apart. I considered an experiment that contained both types of verbs together, but having homogeneous sets with *-de* and homogeneous sets with *-te* would lead to heterogeneous sets containing two highly similar inflectional affixes *-de* and *-te*, and it was felt that that would make the heterogeneous sets too much like the homogeneous sets.

To make the analysis and construction easier, the four sets of nouns with initial phonological overlap (PI) were arbitrarily divided into two separate groups, referred to as PI-A and PI-B. This created four types of words (ME, PE, PI-A, PI-B), with eight items for each type. Thus, all word types are administered on an equal number of items, at the expense of the generalisability of the results for initial phonological overlap. But this latter result is uninteresting for our present purposes; the pseudo-prefixed nouns are merely included to create heterogeneous sets and to prove that the experiment is working.

In each half of the experiment, one set of four past tense verbs (ME) was combined with a set of 4 end overlapping nouns (PE), and a set of four initially overlapping PI-A nouns and four PI-B nouns. In the homogeneous conditions, sets were kept together to create end overlap or begin overlap. For each heterogeneous set, one word from all four sets was taken, and combined these comprised a new set showing no overlap at all.

Procedure and Apparatus The procedure and apparatus was the same as for all other implicit priming experiments. See Experiment 1 for a detailed description.

Analysis Difference scores were computed for all item by subject cells, by averaging over the five presentations of one item in the heterogeneous con-

Table 5.4: Mean production latencies (ms), error percentages, and preparation effects (Δ) for Experiment 9: Suffix overlap

	Context				
Word Type	homogeneous		heterogeneous		Δ
ME: morph. end overlap	758	*6.6%*	753	*3.3%*	-5
PE: phon. end overlap	722	*1.4%*	717	*2.2%*	-5
PI-A: phon. initial overlap	723	*4.3%*	752	*2.5%*	29
PI-B: phon. initial overlap	691	*2.5%*	729	*2.2%*	38

dition, and subtracting the mean of the five presentations in the homogeneous condition. The item and subject factors were accounted for by computing F'. This left two factors in the statistical analysis: Half, with values first and second, and Type, with values ME, PE, PI-A, and PI-B.

Three planned orthogonal comparisons were run, contrasting the levels of the factor Type. ME was compared to PE, ME to PI, and PI-A to PI-B. All and only significant results are reported; insignificant results will be mentioned only when they are critical to the hypothesis.

Results and Discussion

Overall, 3.1% errors were made. The means for the factor Type are summarised in Table 5.4. The first two conditions, with past tense verbs and nouns showing phonological end overlap, showed no preparation effect at all. A small negative effect was found, signalling that the homogeneous condition was in fact slower than the heterogeneous condition. A clear preparation effect was obtained for the initial overlapping nouns, showing once more that subjects can prepare initial substrings of words.

In the statistical analysis, the effect for Type was the only factor that approached significance, $F'(4,59) = 2.46, MS_e = 13202, p = .055$. Contrasts were computed to evaluate where the effect of Type originated. It turned out that verbs (ME) were no different from PE nouns with phonological end overlap: $t_1 = -0.02, p = .981, t_2 = -0.05, p = .963$. Verbs were different from the average PI noun with initial overlap in one analysis but not the other: $t_1 = 1.46, p = .146, t_2 = 2.79, p = .006$. The two arbitrary subsets of PI were the statistically the same: $t_1 = -0.28, p = .778, t_2 = 0.54, p = .589$.

The effect for Type and the planned comparisons show that we have obtained a preparation effect only for initial phonological overlap. No effect of past tense suffix preparation was found. Meyer's claim about the left-to-right planning of segments also holds for the combination of a stem and affixes: The stem has to be segmentally encoded first, after this the affixes are encoded one by one. This process can be controlled by the inflectional frame: As was proposed earlier, the frame feeds the contents of the slots to phonological encoding up to the first (leftmost) empty slot. If an empty slot is encountered, the process is suspended. Resumption takes place at the moment the empty slot is filled. When the end of the frame is reached, the stem and the affixes have all been handed over to phonological encoding and, as far the morphology is concerned, the word is ready to be pronounced.

The strict left-to-right encoding of the full inflectional word is confirmed by linguistic observation: There is, at least in Dutch, affix variation that depends on the phonology of the stem. One example we have encountered frequently in this thesis is the past tense suffix that surfaces as either /də/ or /tə/, depending on the voicing of the stem coda. Phonological encoding of the suffix before phonological encoding of the stem is problematic for this case, because the phonological encoding component needs to know the voicing of the stem coda before the initial stop of the suffix can be encoded.

Experiment 10: Present tenses of strong verbs

In Experiment 6 in the previous chapter, it was shown that strong past tense verbs have a deviant inflectional frame. These verbs encode tense by selection of one of two stem allomorphs, and no regular inflectional tense affix is needed. A proposal was made about how the stem encoding process might encounter a complex form entry when looking up the morphemes for the lemma of a strong past tense verb. The proposed lexical entry for the Dutch strong verb *zwem–zwom (swim–swam)* was given in 4.6 and is repeated here:

(5.1)

lemma		condition	morpheme	inflectional frame
TO SWIM	—	*present:*	**zwem**	Stem + Affix
		past:	**zwom**	Stem + Affix

Because the specification for *tense* is already used at the stem encoding level, it is not copied to the inflectional encoding component. The *tense* diacritic is said to be absorbed by stem encoding. The past tense form should get an irregular inflectional frame to reflect the fact that there will be no tense affix, otherwise the *tense* slot will stay empty and the inflectional encoding process will be suspended indefinitely (see the Discussion section of Chapter 4 for more explanation of this). The inflectional frame for a strong past tense form thus contains a stem slot and a number slot. Evidence for a 2-slot frame for strong past tense verbs was indeed what was obtained in Experiment 6.

An intriguing question is what the frame for the present tense form of a strong past tense verb will look like. In the previous chapter, it was silently assumed that both present and past tense forms of a strong verb take irregular frames. With a complex form entry proposed above, the tense specification has to be taken into account to access either stem allomorph: To obtain the present tense allomorph the condition 'tense = present' has to be satisfied. In keeping with what was proposed for the past tense, satisfying this condition does not only render the present tense allomorph, but absorbs the tense diacritic too. This should be reflected in the inflectional frame, for the reasons outlined above, and an irregular frame with only stem and number slots should be generated. The net result of this is that the present tense form, which is not irregular in itself, is treated in the same way as the irregular past tense form.

An alternative approach is to use a default or *elsewhere* condition. This latter term is commonly used in morphology, after Kiparsky (1973). The formal definition of the phenomenon need not worry us here, the gist is that the elsewhere rule (in this case: the present tense allomorph) is taken if and only if none of the others apply. Under this approach, the 'irregularity' of the present tense allomorph is removed.

The form entry for TO SWIM with an *elsewhere* clause looks as follows:

(5.2)

lemma		condition	morpheme	inflectional frame
TO SWIM	→	*past:*	**zwom**	Stem + Affix
		<elsewhere>	**zwem**	

Encoding of the past tense form is done as before, but encoding of a present tense form is slightly different: When the complex form entry is encountered, a test is done whether the special case 'past' applies. This is not the case and the elsewhere condition is taken. Because no positive decisions were made on the basis of the tense diacritic, it is not absorbed. Instead, the diacritic is copied down to the inflectional encoding level, together with a regular 3-slot frame (no irregular frame is listed for the elsewhere condition) and the present tense allomorph *zwem*. At the inflectional encoding level, the regular present tense affix, which happens to be -ø, is put into the tense slot. In sum, the present tense form of a strong verb is treated like any other regular form under this alternative proposal.

The *elsewhere* approach implies that there is a subtle difference between the regular and the irregular forms of TO SWIM. The regular present tense form bears an inflectional tense suffix and an inflectional number suffix, and its forms are thus *zwem+ø+ø (swim* SG*)* and *zwem+ø+en (swim* PL*)*. The irregular past tense form has no inflectional tense suffix (as was shown in Experiment 6) and is generated as *zwom+ø (swam* SG*)* and *zwom+en (swam* PL*)*. Under the representation proposed in 5.1, however, the past tense is the same but the present tense also lacks an inflectional tense suffix and is generated as *zwem+ø (swim* SG*)* and *zwem+en (swim* PL*)*.

In the following experiment, decisive evidence for either of these two positions will be looked for. The prediction that will be tested is that the present tense form of strong verbs takes an irregular frame, because it does not bear the regular present tense suffix *-ø*. The frame of the present tense form will be compared to the 2-slot frame of a noun. If the present tense takes an irregular verbal frame with two slots, no frame difference will be caused by combining present tense and nouns and the preparation effects in both conditions should be similar. If, on the other hand, a regular 3-slot frame is generated for a present tense, the resulting frame difference should reduce the preparation effect in the variable condition.

Method

The experiment was carried out in much the same way as all other implicit priming experiments (see Experiment 1 for a detailed description). It closely resembles Experiment 6, which investigated the frame of the past tense form of strong verbs. In contrast to the other experiments (except Experiment 6), the sets of this experiment contained four words. This was done to compensate for the fact that the initial overlap in this experiment comprised only the first two segments, and not the full first syllable as in all other experiments.

Participants Sixteen subjects participated in this experiment. This is four more than in other experiments to enable full counterbalancing of the larger sets.

Materials The materials for this experiment were similar to those used in Experiment 6 on strong past tense forms. The verbs from Experiment 6 could not be reused because their present tense forms were sometimes ambiguous between a verbal and a nominal reading. New verbs were selected that were monomorphemic, monosyllabic, and did not have a competing nominal reading. 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 6 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 phonologically and semantically unrelated to all other responses and prompts.

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 materials. For heterogeneous

Table 5.5: Mean production latencies (ms), error percentages, and preparation effects (Δ) for Experiment 10: Present tenses of strong verbs

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

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 5.5). 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. The difference between the variable and constant sets was not significant, the F' for the factor Word Type was $F'(2, 26) = 1.37, MS_e = 27291, p = .266$.

The condition with the present tense forms of strong verbs showed a preparation effect that was slightly, but insignificantly, larger than that in the constant condition. There is no trace of the reduction of the preparation effect that was to be expected under the assumption of a 3-slot inflectional frame for the present tense of strong verbs. From this, the conclusion is drawn that the frame of the present tense forms must contain two slots. An inflectional frame without a tense slot implies that there is no regular present tense inflectional suffix present. This is evidence that the present tense of a strong verb is not dealt with by an *elsewhere* case. Instead, it has a specific representation that is marked 'tense = present'. Accessing that representation absorbs the tense diacritic.

It might come as a surprise that the language system has chosen for 'extra irregularity' in the representation of present tense forms. It is possible to produce these forms along the same lines as all other regular forms with an *elsewhere* condition, and this would honour the fact that present tenses of strong verbs have no inherent irregular properties themselves. The result of this experiment indicates that the language system ignores the resemblance between present tense forms of strong and regular verbs, and treats all forms of a strong verb as irregular forms. In linguistic terms, the paradigmatic dimension is seen as more important than the syntagmatic dimension. This approach also has a processing advantage: The only tense affix that occurs with the present tense allomorph of a strong verb is the null morpheme -*ø*. By using an irregular inflectional frame, the redundancy of a slot that always contains *ø* is removed.

Experiment 11: Compounds

The theory outlined in Chapter 3 claims that no slots for derivational suffixes occur in the inflectional frame. A similar claim has been made for compounds: Although there are two or more clearly discernible elements to a compound stem, only one stem slot is reserved in the inflectional frame. This hypothesis is put to test in this experiment.

The hypothesis will be tested by comparing nouns and compounds that overlap in onset and nucleus of the first syllable. Example items are the compound *potgrond (potting soil)* and the noun *polka (polka)*. The initial /po/ overlap will give rise to a phonological preparation effect, which can be reduced by a difference in inflectional frames. If compound stems do indeed occupy only one stem slot in the inflectional frame, no frame difference between (monomorphemic) nouns and compounds is expected and the full phonological preparation effect should be obtained.

Method

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

Participants Nine subjects participated in this experiment. None of them had participated in another experiment related to this thesis in the three months prior to experimentation. Subjects were paid for their efforts.

Materials Three compounds were selected that were bisyllabic, stress-initial, and contained two parts that were easily discernible, familiar words. To maximise the chance of obtaining an effect of the complex stem, only transparent compounds were used. The three compound words were *windvlaag (gust of wind, lit. wind+gust), hakblok (chopping block, lit. chop+blok), potgrond (potting soil, lit. pot+soil).*

Matching monomorphemic nouns were taken from the CELEX database. For each compound, five nouns with that overlapped in onset and nucleus (CV) were taken. The variable condition contained two such nouns and the compound, the constant condition contained the remaining three nouns.

Table 5.6: Mean production latencies (ms), error percentages, and preparation effects (Δ) for Experiment 11: Compounds

	Context				
Word Type	homogeneous		heterogeneous		Δ
variable	650	*2.8%*	696	*2.8%*	46
constant	669	*4.8%*	698	*2.2%*	30

Results and Discussion

Overall, 3.0% errors were made. There was a significant overall preparation effect $F'(1, 18) = 7.37, MS_e = 18214, p = .013$. There were no other significant effects. Crucially, the effect of Word Type (constant vs. variable sets) was not significant at $F'(3, 17) = 0.98, MS_e = 3586, p = .414$. Levene's test for heterogeneity was significant, $F(5, 12) = 3.70, p = .029$, signalling that there was a considerable difference in variance between the six subsets defined by Word Type and Base. The effect for Base was not significant, however. From the plot of the item means in Figure 5.2, one can see that the heterogeneity stems from the items *halster, halte, handel*. Again, the variance in the data was compared to the simulations of Santa et al. (1979). For variance parameters that are almost identical to the ones observed and similar heterogeneity between groups, their simulations lead to $\alpha = .073$. This means that our test is actually too liberal, and the absence of a significant difference between the conditions is statistically very reliable.

With the precaution that there was some variance between items, it can be concluded from this experiment that there is indeed no inflectional frame difference between compounds and simple nouns. The presence of a compound did not give rise to a systematic reduction of the preparation effect, as was found in the experiments where there was an inflectional frame difference.

Experiment 12: Mixing singulars and plurals

The last experiment in this chapter considers one more property of the inflectional frame. The definition of an inflectional frame states that *it contains a slot for every possible type of inflectional affix, and one slot for the stem.* Are the slots for affixes that do not surface in the uttered word really present in the inflectional frame? To answer this question two classes of words will be examined that differ in their number of possible affixes, but

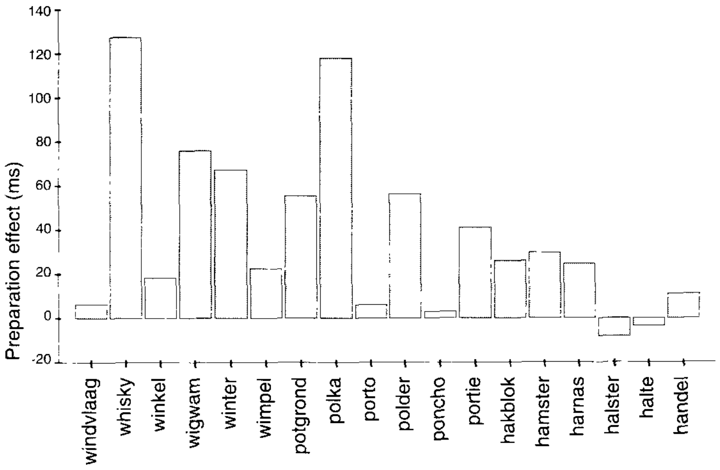


Figure 5.2: Mean preparation effects for the items in Experiment 11

are equal in the number of actual affixes (not counting -ø). The well-known contrast between 2-slot nouns and 3-slot verbs will be used again. In Experiment 5, nouns and past tense verbs in their respective plural forms were used and the expected frame difference (2 vs 3 slots) was obtained. These materials will be reused in this experiment. Singular verbs are used here instead of plural ones, and the number of actual suffixes on the verb will be one (tense). Plural nouns have the same number of actual suffixes, but a different number of hypothesised slots in the inflectional frame. If our proposal for the number of slots in the inflectional frame is correct, an inflectional frame difference is predicted when singular past tense forms and plural nouns are mixed. Under the alternative hypothesis, the number of actual suffixes is constant, and no frame difference is expected.

Method

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

Participants Twelve subjects participated in this experiment. None of them had been a subject in Experiment 5 or in any other implicit priming experiment for the three months prior to testing.

Table 5.7: Mean production latencies (ms), error percentages, and preparation effects (Δ) for Experiment 12: Mixing singulars and plurals

	Context				
Word Type	homogeneous		heterogeneous		Δ
variable	705	*4.6%*	705	*3.2%*	0
constant	693	*6.1%*	740	*5.9%*	47

Materials The materials from Experiment 5 were used again with one modification: The three plural past tense forms were changed to singular past tense forms. Their prompts stayed the same and so did all other prompts and responses.

Results and Discussion

A dramatic difference between the two conditions showed up: In the variable condition (with the singular past tense) there was no preparation effect at all, and in the constant condition (with only plural nouns) 47 ms of preparation was obtained (see Table 5.7). This difference was statistically significant, $F'(1, 21) = 6.41, MS_e = 13097, p = .017$. Probably due to the zero preparation effect in the variable condition, the overall preparation effect failed to reach significance $F'(1, 17) = 4.18, MS_e = 23472, p = .052$. The problem lies in the variability between subjects, as is apparent from the classical tests: $F_1(1, 11) = 5.13, MS_e = 23472, p = .045; F_2(1, 12) = 16.73, MS_e = 7204, p = .002$. Notice that the F' just fails to reach significance, while F_1 and F_2 are within the standard $\alpha < .05$ limits. With some precaution, I will continue to discuss the results as if the overall effect is fully reliable. An analysis of simple main effect showed that the preparation effect in the constant condition is significantly larger than zero: $F'(1, 19) = 9.84, MS_e = 18284, p = .005$.

The preparation effect obtained in the constant condition was reduced in the variable condition by a deviant inflectional frame. The constant condition contained plural nouns; the variable condition contained plural nouns and one singular past tense form. Both bear one discernible inflectional suffix. This refutes the hypothesis that it is the number of actually present inflectional affixes that determines the inflectional frame. Instead, the number of slots in the inflectional frame equals the number of possible types of inflectional affixes, plus one for the stem.

Having answered the question that motivated this experiment, we are left with the problem why there is no preparation effect in the variable condition. Careful inspection of subject and item means showed that the effect is consistent across both and it is not an artefact caused by the singular verbs alone. If one compares the results of this experiment to the results of Experiment 5 (Table 4.6 on page 85), it is clear that the RTs in both the homogeneous conditions are a good 20 ms slower in this experiment. The heterogeneous conditions are virtually identical, and this reduces the preparation effects of Experiment 5, 22 and 70 ms for variable and constant conditions, to 0 and 47 ms for this experiment.

In an ANOVA comparing the preparation effects of this experiment with those of Experiment 5, the 20 ms difference did not come out significantly: $F'(2,34) = 2.08, MS_e = 12236, p = .150$. But one should keep in mind that this is a between-subjects comparison, and inspection of the F_1 and F_2 revealed that it is indeed only the F_1 that is non-significant: $F_1(1,22) = 1.17, MS_e = 28433, p = .290; F_2(1,12) = 13.83, MS_e = 4649, p = .003$. In other words, the 20 ms difference between the two experiments might well be a real effect.

What then is the cause of this reduced preparation effect? The proposed interpretation is that the phonological preparation effect due to shared initial syllables was masked by an extra processing step that was of shorter duration (or not necessary at all) in Experiment 5. The timing diagram that was proposed for Experiment 4 is copied in Figure 5.3, with the modification that the affix generation step (bold arrow labelled *affixes*) is now slightly lengthened. The phonological encoding of the second part of the word depends on the availability of the affix morphemes, because these influence the second part of the word. As shown in the Figure, the slight lengthening of the affix selection step delays the encoding of the second syllable in the homogeneous condition. For the variable sets, this happens to mask the benefit of preparing *phonemes-1* and no difference between the homogeneous and heterogeneous condition is predicted (zero preparation effect). For the constant sets, it merely reduces the preparation effect.

This tentative timing diagram has shifted our problem from explaining the absence of a preparation effect to explaining why affix selection takes longer in this experiment. The unpredictability of the affixes in this experiment is an obvious possible source of the delay: Throughout the experiment, the language production system had to choose between two past tense affixes, *-te* and *-de*, and two plural suffixes, *-en* and *-s*. Compare this to the other experiments, in which the choice was always between two possible inflectional suffixes (and often between a real affix and *ø*). Given more affixes to choose from, the affix selection process took more time in this experiment.

The explanation may seem to go against the results of Experiment 9, where no preparation effect was obtained for consistent suffixation with the past tense affix. Absence of a preparation effect for words that only overlapped in their affixes, however, does not necessarily imply that there can be no inhibitory effect from the forced on-line choice between a larger range

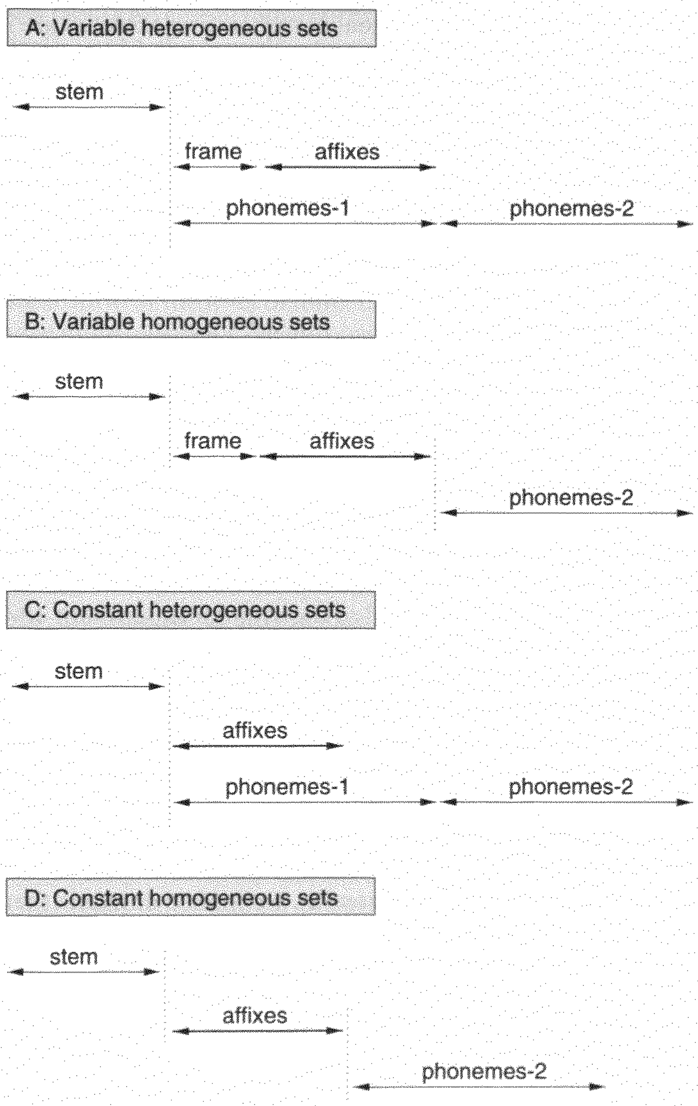


Figure 5.3: Tentative timing diagram for Experiment 12

of affixes.

The conclusion remains that the inflectional frame contains a slot for every possible type of inflectional affix. This provides a check on the presence of the inflectional affixes: The word is not complete before every inflectional affix is coded by either a real morpheme or ø. Affixes are inserted into the frame by converting the diacritics into morphemes, either by rule or by look-up. The same diacritic (eg. number) can be expressed by different affixes (eg. *-s* and *-en* in Dutch), and this has to be taken into account. The strong point of this proposal is that both affix lookup and affix checking are extremely simple processes. This simplicity more than compensates for the extra effort that has to be put into generating two null morphemes for a form like *ik zeil+ø+ø (I sail).*

General Discussion

The experiments reported in this chapter help us to understand further the inflectional frame and the processes working on it. First of all, the most counterintuitive result obtained in the previous chapter was replicated in **Experiment 7**, and it was again shown that when the response words in an implicit priming experiment have different inflectional frames, this reduces the *phonological* preparation effect. The explanation given for this in Chapter 4 was supported.

The results of this replication and of the original Experiment 4, were put to a critical test by replacing the verbs in the materials by homophonous nouns. It was predicted that the reduction of the phonological preparation effect would disappear because now all response words are nominal and the inflectional frame is consistent across the set of possible responses. This was indeed what was obtained in **Experiment 8**. It means that the phonological preparation effect due to predictable initial segments will only be fully obtained if the inflectional frame is the same for all response words and therefore predictable.

In the next experiment (**Experiment 9**), we looked at words that shared the same inflectional suffix. All other experiments reported in this chapter contained words that showed initial phonological overlap, and gave rise to a phonological preparation effect. In this experiment, it was questioned whether words with different onsets but the same inflectional suffix would also give rise to a preparation effect. No evidence for this claim was found, and the conclusion was that the phonological representation of an inflectional form is built up from left-to-right: The stem has to be completed before phonological encoding of the affixes can start.

To broaden our knowledge of what happens to irregular forms, the verbs with strong past tenses from Experiment 6 were revisited. In the previous chapter, it was shown that an irregular inflectional frame accompanied the past tenses of a strong verb. The same question was asked for the present tense forms of these verbs in **Experiment 10.** Although the present tense itself is not irregular at all, it turned out that it is treated in the same way as the past tense form. The merits of this approach were discussed. A processing advantage of this representation is that it removes the need to insert a fully predictable and empty tense suffix (-ø) for all present tense forms.

In **Experiment 11**, the complex stem of compounds was investigated. The stem of a compound word contains at least two easily discernible parts (in Dutch: *onder+deel*), but it turns out that it is stored in one slot of the inflectional frame. The same was found for derivational adjectives and diminutives in the previous chapter (Experiment 2 and 3). The conclusion that can be drawn is that the inflectional frame is not sensitive to the derivational make-up of the word at all: Whether the stem contains derivational suffixes or is created by combining two morphemes with word status, it is treated as a monolithic form at the level of inflectional encoding.

This once more pleads for the proposed subdivision of morphological encoding. The stem is created in the stem encoding subcomponent, and together with the inflectional frame and the (remaining) diacritics, it is fed to inflectional encoding. As is usual for the communication between components, the possible complexity of the stem is not relevant for further processing and it stays internal to the stem encoding component. In linguistics, the same principle has been formulated as the Bracket Erasure Convention (Kiparsky, 1982).

Experiment 12 addressed the question whether the inflectional frame is affected by the number of affixes that are actually present at the word. It seems slightly circumlocutory to have a frame with slots for all possible affixes, and then fill all of them with ø for a form like *ik zeil+ø+ø (I sail).* In the experiment, singular past tense forms (with 2 possible suffixes, of which one is realised in this form) were compared to plural nouns (with one possible and actual affix). It turned out that the past tense words carried different frames, even though the number of actual affixes was the same. Clearly, the inflectional frame takes the circumlocutory approach. The reason for this might be that only in this way can the inflectional frame be used to check whether all required inflectional affixes are present.

Experimental considerations

The experiments presented in this and the previous chapter make a strong case for the existence of an inflectional frame and for a split morphology approach to the language production process. One should keep in mind, however, that we have used only one methodology so far, the implicit priming paradigm. This paradigm has proven to be extremely useful for researching language production. The task does not require the subjects to read the word they are about to utter (cf. word naming) and does not impose 'picturability' constraints on the response words (cf. picture naming). Compared to the picture word interference paradigm, the task is confined to language production proper and does not involve a concurrent comprehension task.

The key differences between casual speaking and uttering words in an implicit priming task concern the context of utterance and the occurrence of a preparation effect. In contrast to realistic situation, the implicit priming task puts the subject in a one-word utterance mode. This is a reasonable simplifying assumption for the production of nouns, but its merits for the production of inflected verbs are less self-evident: An inflected verb always occurs in the context of a sentence and this sentential context is removed in a one-word utterance. When subjects were questioned after the experiments, they never complained about the task being unnatural or hard to accomplish and this argues against a possibly disturbing effect of the one-word utterances. Next to that, in a series of experiments not reported here, I used lead-in sentences to create a sentential context for a second person singular present tense form:

(5.3)

lead-in		prompt		response
dat	*jij*	*een*	*plant*	*begiet*
that	you	a	plant	water.2/3SG

This lead-in sentence did make the task more natural for the subjects, but the results were fully comparable to another experiment, which addressed a similar question without making use of lead-ins.

The one fact that is principally task-specific about the implicit priming paradigm is the occurrence of a preparation effect. It is via this effect that we gain insight in the workings of the production process, but the effect is induced by a context in which the participant knows she is about to utter a word from a restricted set of alternatives. Clearly, this is not an everyday situation, although there is a parallel with the (much more common) situation that the speaker has prepared a complete utterance.

There currently is no precise enough theory of how the preparation effect comes about. The suspend/resume mechanism can account for some but not all of the effects: To explain the basic finding of phonological preparation of shared initial segments, some extra assumptions have to be made. Systematic research into the properties of the task should be carried out to make clear what the mechanisms underlying these effects are. It should be stressed, however, that a precise theory is lacking for many other experimental paradigms. A better theory of what implicit priming is, is still of greatest importance, but given the consistency of the data, one should not shy away from drawing more than preliminary conclusions from what has been found so far.

A similar precaution applies to the materials used in the experiments. The implicit priming task forces us to use rather small sets of materials and the question of the generalisability of the results thus arises. In the one case where an experiment was replicated (Experiment 4 and Experiment 7) an identical data pattern was obtained with different materials and this strengthens our assumption of generalisability.

Related to this, there are many factors that might influence (morphological) processing that have not been addressed in this study. The fact that we are unable to predict raw production latencies indicates that there are possible influences from factors like association strength, morphological opacity, morpheme frequency, prompt reading time, etc. By using a heterogeneous control condition and working on difference scores, the effects of these factors have mostly been eliminated. This does not mean that they are uninteresting and a complete theory of language production should certainly address these factors too.

The TTT-account

The account presented in this thesis is of course not the only possible way to explain the data. Various other descriptions of the results can be made, some of which may give rise to concurrent theoretical interpretations. I will examine one such account here.

If, for the ease of exposition, one abstracts from the actual outcomes and inspects which experiments did and which did not give rise to an effect of Word Type (variable sets were different from constant sets), the following pattern emerges: All and only those experiments in which an overt inflectional tense suffix (*-te* or *-de*) was present on the odd man out gave rise to an effect of Word Type (see the overview of the experiments in Table 5.8).

What are the possible theoretical underpinnings of this description? One possibility is to make a distinction between tense inflection and number inflection and propose distinct processes for the two. To explain the reduced preparation effect in the variable conditions, we assume that the process that expresses overt tense inflections *-te* and *-de* takes a considerable amount of time (I will refer to the explanation developed here as the tense-takes-time or TTT-account). In the constant condition, the absence of any tense inflection allows this component to be bypassed and the preparation effects appears in full bloom. The same holds for the variable conditions of experiments 2, 3, 6, 8, 10, and 11, in which no past tense suffix was present: Again, the presumably lengthy tense expression process could be bypassed and the same preparation effects as in the constant condition was obtained.

The strong point of the TTT-account is that it is extremely simple. One of its weaker points is that it does not capture the data in such detail as the inflectional frame hypothesis does. Instead of only predicting the presence of a difference between the two levels of Word Type, we were able to account for the magnitude of the morphological preparation effect in Experiment 1 and for the size of the reduced phonological preparation effects in Experiments 4, 6, and 7. The absence of any preparation effect in the variable condition of Experiment 12 was also accounted for.

Experiment	Word Type	Example response words	Δ	Significance of Word Type
1	variable	*bouwde, bouwjaar*	36	**Yes**
	constant	*bouwkunst*	122	
2	variabel	*brandbaar, branddeur*	92	No
	constant	*brandmerk*	77	
3	variable	*huisje, huisvrouw*	39	No
	constant	*huiswerk*	50	
4	variable	*waadde, wagen*	30	**Yes**
	constant	*wafel*	67	
5	variable	*haatten, hanen*	22	**Yes**
	constant	*hazen*	70	
6	variable	*kreeg, krent*	44	No
	constant	*krant*	43	
7	variable	*raadde, radar*	27	**Yes**
	constant	*ratel*	91	
8	variable	*wade, wagen*	25	No
	constant	*wafel*	30	
9	Suffix overlap experiment *nieste–jeukte* vs. *ruiter–kater*			
10	variable	*kruip, krent*	49	No
	constant	*krant*	29	
11	variable	*potgrond, polka*	46	No
	constant	*polder*	30	
12	variable	*haatte, hanen*	0	**Yes**
	constant	*hazen*	47	

Table 5.8: Overview of the experiments

The distinction between tense inflection and number inflection made by the TTT-account is also problematic. There is, to my mind, little independent reason for assuming that the two are generated by different processes. Both types of inflection produce a number of predictable forms from a base, and both processes can be preempted by irregular lexical forms (eg. *sheep, swam*). Booij (1993, 1996) has proposed to divide inflection into two subclasses, inherent and contextual inflection. Apart from the fact that his division can be disputed on a number of grounds,[2] number inflection on nouns

[2]Part of Booij's argumentation rests on a very strict interpretation of the split morphology hypothesis, such that no inflectionally complex forms are allowed in the lexicon. This certainly

and tense inflection on verbs are both cases of inherent inflection in Booij's proposal and this would instead argue for grouping verbal tense and nominal number together in one process.

In his papers, Booij gives several arguments for a distinction between inherent and contextual inflection. On closer examination, it appears that two of his most important arguments apply only to number marking on nouns, and not to tense marking on verbs. These two arguments concern the possibility for inflected forms to acquire a meaning different from their stem (as in *affair–affairs*) and the appearance of inflected forms in derivations or compounds (as in *botenhuis*, lit. 'boats house'). Both arguments apply mainly to nouns an if they apply to inflected verbs (the example *gezond*, 'healthy' lit. 'sunbathed' was given in Chapter 1) it is to participles and infinitives and never to tense or person inflected forms.

Can this provide us with a justification for the distinction between inflections on nouns and verbs? I want to argue that there is no principled reason why forms like *schrijft (writes*+2/3SG*)* cannot acquire a specialised meaning or appear in compounds. There is, however, usually little reason or use for creating such a form because the language user is rarely in need for the associated meaning nuance of specificity to the addressee or a third person. Hence, *?sterftscène* (dies.2/3SG scene) is unlikely, but *Ceasar-sterft-scène* (Ceasar-dies-scene) is definitely possible. Past tense inflected forms have found some less marginal applications. Meaning specialisation of the past tense, for example, has taken place for the in itself outlandish combination of the past tense and the future auxiliary *zou (would)*, which has come to mean *irrealis*. Similarly, imperative inflection appears in composites:

(5.4)	*een*	*laat-*	*maar-*	*waaien-*	*houding*
	a	let.IMP!	particle	be.blown	attitude
	a laisez-faire attitude				

Notice that the English and French translations also use an imperative inflection.

Slightly more speculative is the analysis of the past tense *dronk (drunk)* in the compound *dronkaard (drunkard)*. If the word were **drinkaard (*drinkard)*, the invited meaning would be 'someone who is likely to drink a lot'. The past tense enforces a more appropriate resultative reading, 'someone who is likely to have drunk a lot'. This analysis is speculative (and slightly tongue-in-cheek) but serves here to demonstrate that there are, to my mind, no cases of meaning oppositions due to the choice of inflectional constituent. Returning to compounds with plural nouns: Next to *botenhuis (boats house)*, there is no **boothuis*; *dakenzee (sea of roofs)* has no counterpart in **dakzee*; and if *?vaderenmoord (lit. forefather murder)* exists next

creates problems for words like *clothes*, but the position attacked seems rather artificial. Also, number marking on verbs is a case of contextual inflection because it is dictated by agreement with the subject noun phrase. Does this make verbal number marking a case of inherent inflection when the subject is omitted in a pro-drop language?

to *vadermoord (patricide)*, it is because of the meaning difference between *vader* and its former plural *vaderen*.

Summarising the point made, plural nouns are more likely to acquire independent meaning or to appear in compounds as inflected verbs because they have wider semantic content. In Bybee's terminology, number inflection has more *relevance* to nouns than to verbs. By the same token, infinitives and participles are more likely to lexicalise than tense inflected forms. For all forms, there seems to be a generalisation that irregular forms, which are already stored in the lexicon, are more likely to specialise and occur more often in compounds than regular forms. Because irregular forms are also of relative high frequency, it is currently unclear whether this is an effect of lexicalisation or only of frequency.

There are, apart from the results of the experiments, no sound reasons for assuming a difference between nominal inflection and verbal inflection. The TTT-account, which assumed a processing difference between the two to account for the presence of a Word Type effect, is extremely simple and sparse but has little to recommend for it in terms of precise data fit and independent evidence for its assumptions.

Comparison with other models

How do the other models of morphology that were discussed in Chapter 2 fare in accounting for findings reported in this thesis? Because the process of inflectional affixation is not covered in detail by any model, I will limit the discussion to the coverage of the following, more general, findings:

- There is a processing difference between inflection and derivation;
- Regular inflected verbs are different from nouns: A reduced preparation effect is obtained when verbs and nouns are mixed;
- Both present and past tense of strong verbs are similar to nouns in this respect;
- Evidence was obtained for inflectional frames and for the presence of null morphemes in this frame.

The connectionist models are currently not very outspoken on the difference between inflection and derivation. The difference between regular and irregular forms is explicitly denied by these models and it is not easy to see how the reduction of the preparation effect caused by the regular verbs but not by the irregular verbs can be modelled. The line of research that has sprung from the work by Rumelhart and McClelland (1986) considers past tense generation to be a mapping problem from a phonological input *(swim)* to a phonological output *(swam)*. In such an approach, null morphemes can play no role at all. Connectionist models deny the existence of symbolic representations in cognitive processes and hence there is no natural place for a frame either. In all, most of the findings in this thesis do not agree with the assumptions made by connectionist models.

Bybee's theory of inflection is better suited to account for our findings. Although Bybee stresses the fact that inflection and derivation are two points on a scale, it is still possible that there is a processing difference between the two. Similarly for the difference between regular and irregular verbs: Although both are full listed in the model, the 'association density' for regular verbs must be higher and this can form the basis for a difference between the two. Especially if the associations that express inflectional morphology are compulsory and can be distinguished from associations that express meaning, form overlap, etc., a frame-like explanation for the effects of number of possible affixes can be implemented.

The production model proposed by Dell does already differentiate between inflection and derivation. Frames play an important role at all levels of speech production and an inflectional frame would be the natural associate of a new layer in the model that specialised in inflectional encoding. This specific layer may also be the key to distinguishing between regular and irregular verbs, similar to the proposed treatment of this difference in the Levelt/Roelofs-model. One necessary addition to the theory would be that of the inflectional null morpheme.

None of the existing models of language production or comprehension addresses the issue of inflectional affixation in sufficient detail to derive predictions for the experiments presented in this thesis. A modification to the Levelt/Roelofs-model was proposed to capture the findings with this model. Similar modification to other theories, or —even more interesting— completely different modifications that can account for the same data, should be made before one can evaluate the parsimony and elegance of the current proposal. The models of Bybee and Dell are particularly well suited for this.

Implications and Extensions

CHAPTER 6

6.1 The inflectional null morpheme

Zero or null elements have a long history in linguistics. Especially the morphological theories of the Item and Arrangement-type (Hockett, 1958) have to make extensive use of null morphemes because almost all morphological processes that do not result in the affixation of a discernible morpheme have to be accounted for by a null morpheme. Take for example the process of noun-to-verb conversion in English: Assuming that the grammatical class of a word is always determined by the rightmost constituent, a noun like *the plan*$_N$ can only be converted into the verb *to plan* by suffixation with a derivational null morpheme, as in *the [plan*$_N$ + $ø_V]_V$. Pesetsky (1995) has argued for the existence of derivational null morphemes in this and many other places.

Next to phonologically empty derivations, a much favoured null unit in morphology is the boundary element (or juncture), which appears as '#' or '+' in most theories. One of the first analyses to use this boundary element was proposed by Bloomfield (1930) and concerns the classical problem of German allophony between /ç/ and /x/. In almost all cases the choice between these allophones can be predicted from the phonological context, but some apparent violations to this rule occur when the German diminutive suffix /-çən/ is added to a noun:

(6.1)	a.	*Kuhchen*	/kuːçən/	cow DIM
	b.	*Autochen*	/autoːçən/	car DIM

Both examples contain /ç/ after a long back vowel, something which is prohibited by the allophonic rule. Bloomfield proposed that a phonological null element be introduced at the end of the stem. This results in [*Kuh+ø*]+*chen* and the boundary element would block the allophony rule that would otherwise change the /-çen/ into a /*-xen/.

This analysis in terms of boundary elements is essentially maintained in modern-day linguistics, but it has recently been challenged on observational and psycholinguistic grounds by A. Weber (work in progress, MPI): It seems that German speakers (nowadays) avoid the use of *Kuhchen* and similar forms that are based on monosyllabic stems. Planned experiments on the recognition of these sounds in various contexts will hopefully shed more light on this issue. Similarly, the necessity of the derivational null morphemes, like -$ø_V$, can be questioned in a linguistic model of morphology that is process-based and does not require that every morphological rule leaves a discernible trace (eg. Anderson, 1992). The Levelt/Roelofs-model of language production goes one step further and does not even require an on-line process that derives *to plan*$_V$ from *the plan*$_N$ or vice versa. The two forms corresponding to the concepts A PLAN and TO PLAN are arrived at by accessing two different lemmas, which so happen to both point to the same morpheme *plan*.

A principal question is how the listener can infer that phonologically empty boundaries or phonologically empty derivational affixes are present in the input. A first language learner is facing the same problem, with the added twist that there must be a compelling reason for assuming the existence of null elements in the first place (other than a stopgap innateness account). Because there is no phonological evidence for the null morpheme, the cue to its presence must come from a structural source. Of course, this is also what Bloomfield's account appeals to: Given the German rule that changes /ç/ to /x/ after a long back vowel, the failure of this rule to apply in the case of *Kuhchen* can be taken as a structural cue to the presence of an invisible disturbing element. However, this cue is only structural under the assumption of the aforementioned rule.

The difference between the types of null morphemes discussed so far and the proposed inflectional null morphemes lies precisely in the fact that there is reliable structural evidence for inflectional null morphemes. This evidence is not bound to any rule or theory, but apparent from the paradigms in which the inflections occur. The case for English may not be immediately convincing, but the case for Dutch certainly is. In Table 1.1 (on page 7) the finite tenses of Dutch were listed and the pattern of affixation is summarised again in Table 6.1. As can be clearly seen from this summary, the Dutch finite verb almost always bears an inflectional affix, with the first person singular form of the present tense as the exception. The null morpheme posited for this form is based on the fact that, within this paradigm, the absence of any suffix is salient enough to provide the listener with an unambiguous structural clue. Put differently, in this paradigm the absence of an inflectional affix deserves its own representation: ø.

No such reasoning can make a case for the derivational null morphemes. There is no structural clue that indicates the presence of a derivational null suffix in [*plan*$_N$ + $ø_V$]$_V$: Although some verbs do contain a derivational suffix (ie. *null+ify*), this is not true for all verbs of English. Similarly, although other forms of the stem *plan* bear one derivational suffix (*plan+ner*), this is no necessity either. The only reason for assuming the null morpheme is that a trace of the alleged noun-to-verb conversion is required by that specific type of theory. There is no structural evidence for the reality of ø, like there

Table 6.1: Inflectional paradigm for finite verbs in Dutch and in English

A: Dutch

		Number	
Tense	Person	Singular	Plural
present	1		en
	2	t	en
	3	t	en
past	1	de	den
	2	de	den
	3	de	den

B: English

		Number	
Tense	Person	Singular	Plural
present	1		
	2		
	3	s	
past	1	ed	ed
	2	ed	ed
	3	ed	ed

is for the case of Dutch verbal inflections.

The strict reasoning pursued above grants Dutch verbs with an inflectional null suffix, while denying the reality of derivational null morphemes. As said, the case for English is less clearcut: Many cells of the paradigm are empty (Table 6.1). I want to claim that English nevertheless has inflectional null morphemes, based on the observation that one of the properties of inflectional affixes is that their presence is required (Greenberg, 1954; Matthews, 1974; Bybee, 1985). Although an analysis of the English present tense without null morphemes is possible (at least in principle), assuming a null morpheme for the empty cells of Table 6.1 accords with the requiredness of inflectional markers. To definitively solve the issue, experiments similar to those reported in this thesis should be run in English.

Once again, the criteria for null morphemes formulated here deviate strongly from the criteria that are used in linguistics. We have denounced the use of derivational null morphemes, which are most popular in linguistics to maintain a strict meaning-to-form mapping. Inflectional null morphemes, on the other hand, are essential to the presented theory, but are absent from most linguistically inspired theories. The processing perspective taken here requires a null symbol for signalling that the underlying process (of affix encoding) has successfully run but generated no output. This is entirely different from the approach taken by linguistic theories, where null morphemes are introduced to account for forms that do not fit the proposed rules and regularities.

There is, to my knowledge, no independent experimental evidence for the presence of inflectional null morphemes in the psycholinguistic literature (see Pesetsky, 1995 for some linguistic arguments). The experiments reported in this thesis provided evidence for the existence of an inflectional frame, with slots for every type of inflectional affix. Indirectly, this is evidence for inflectional null morphemes because these slots must always be filled. When no overt affix is available, a null morpheme must stand in to satisfy the proposed mechanism that checks whether all slots of the inflectional frame are filled.

For Dutch verbs there is an additional complication: Two inflectional markers are present and both can be zero at the same time, like in *ik zeil+ø +ø*. From a processing perspective, this certainly makes sense because the two zeros occupy the places of two overt affixes that would be present in other contexts (past tense *-de* and plural *-en*). By linguistic criteria, the two null morphemes are hard to justify because they do not serve any linguistically relevant purpose, nor do they accord with the linguistic constraints on zero elements.

It might, most crucially, also be hard to see how *two* null morphemes could be structurally unambiguous in the sense proposed above. Inspection of the Dutch paradigm in Table 6.1 does indeed lead to the assumption of one null morpheme for the first person, singular, present tense case. But even though the null morphemes must be structurally recognisable in a word, their ultimate use is to serve the language production process. If this process produces one portmanteau suffix for the combination of tense, number, and person, one zero morpheme suffices and the set of possible affixes comprises *-t, -en, -de, -den,* and *-ø*. If the production system observes that tense can be separated form the other affixes, it is left with two simpler choices: One between *-de* and *-ø*, driven by tense, and one between *-t, -en,* and *-ø*, which reflects the combination of person and number.

6.2 Semi-morphs

In the section on speech errors (in Chapter 2), it was observed that there is a remaining class of speech errors that cannot be explained by most theories of language production:

(6.2) Lackner and Goldstein → Goldner and Lackstein (Garrett, 1975)

In these errors, the exchanging elements resemble morphemes but cannot be granted that status because there is no linguistic rule or productive process of *-ner* and *-stein* suffixation. These strings are also not viable inputs for compound formation, because this process combines only existing words. An explanation in terms of exchanging syllables is at odds with the generally observed absence of the syllable as an error unit (Garrett, 1988; Levelt, 1989) and it would not explain the morpheme-like quality of the exchanging elements.

Under a strictly word-based approach towards morphology, however, these examples are explained in natural way. What is required is that one cuts all ties with semantics and productivity. In this rigorous interpretation of split morphology, there are no longer any constraints on how far the system can go in exploiting the redundancy of language by decomposing it into morphemes. To return to an earlier example, there is actually little that will withhold the language learner from representing *capable* as a combination of *cape* and *able*, and *capability* as *cape* and *ability*. No harm is done by the fact that *cape* has a nautical or garment reading when standing alone: This situation regularly occurs for components of opaque compounds, like *sore* in *eyesore*.

Rigorous application of morphological analysis means that, given a certain number of words ending in *-ner*, the language production system may decide that it can represent *ner* as independent unit and treat all such words as combinations of two morphemic units: *Lack+ner*. The resulting unit *ner* is not distinguishable from other morphemes and will be called a *semi-morph*. Because the Levelt/Roelofs-model represents each use of a morpheme by a pointer at the lemma level, a tremendous increase in efficiency of storage can be reached by representing semi-morphs in this way. As said, this requires us to cut all ties with semantics and semi-morphs do not fall under the traditional definition of the morpheme as the smallest meaningful part of a word. Below, semi-morphs will be compared to morphemes in opaque combinations and to morphemes functioning in cranberry compounds. The morphemes in these combinations do not adhere to the definition of the morpheme as the smallest meaningful form element either and this is taken as an argument in favour of the current analysis.

Whereas the representation for real morphemes and semi-morphs is the same, they come about for different reasons and behave differently under word formation. Productive morphemes, like *-able*, have a meaning component associated with them (something like "can be the subject or product of V"). This meaning component is used to create and understand neologisms and is still (partly) present in many words formed with *-able*. This resulting semantic and formal regularity is captured by linguistic descriptions, which relate base forms to their derivational counterparts. Semi-morphs, on the other hand, are devoid of meaning. They cannot be used in neologisms and many uses of semi-morphs resist assigning them a base or affix status (as in *Lack+ner*). The reason for independently representing a semi-morph is purely a matter of efficiency of storage and there is not necessarily a systematic meaning relation between base forms and derived forms created by affixation with a semi-morph.

Despite the differences sketched above, there are certain classes of traditionally recognised morphemes that bear resemblance to semi-morphs. First, consider the morphemes that figure in opaque compounds. The morpheme *eye* as used in *eyesore* has hardly any relation to its meaning "organ of sight". Neologisms with the *eyesore*-meaning of *eye* are not readily created. The most compelling reason for analysing *eye+sore* into two morphemes is the appearance of these morphemes in other combinations and as independent words. Similar observations can be made for opaque derivations like *depart+ment.*

A second example is provided by the cranberry morphs (eg. *cran*). These morphemes occur in only one compound, have no independent meaning, and are recognised as separate morphemes to salvage the morphological status of *berry* in *cranberry, raspberry,* and *loganberry.* Some semi-morphs are recognised for the same reason.[1] Consider the following examples:

(6.3) a. My shoulders are frozen → My frozers are shoulden (Garrett, 1975)
b. I've got a load of chicken cooked → I've got a load of cooken chicked (Garrett, 1980)

The identification of the semi-morph /ʃɔld/ (from shoulders) in the first example is due to the pull for separating the suffixes *-er* and *-s.* This case is very similar to that of *cran+berry.* The speech error shows that morphemes and semi-morphs can freely interact: The constituents of *froz+en* do form morphemes in the traditional sense and *froz* is exchanged with the semi-morph /ʃɔld/.

Example (6.3b) shows that the language production system cannot only ignore semantic constraints on morphemehood, but also the formal requirements. The suffix *-en* is usually (but not productively) attached to participle forms. In the case of *chicken,* the production system has apparently disregarded this requirement and represented *chicken* as *chick+en.* This opened the possibility for exchanging this morpheme with the inflectional suffix of *cooked.*

Next to the aforementioned cases of real morphemes resembling semi-morphs, semi-morphs can also acquire partly morphological status. Many examples exist of a salient part of a word which is subsequently analysed and extended by analogy, cf. the analysis of *workaholic* by Anshen and Aronoff (1988) presented in Chapter 2. The *-aholic* suffix appears in only one or two forms. Much more widespread is the semi-morph *Mac* or *Mc.* The case made for the analysis of *Lack+ner* is very compelling for that of *Mc+Queen,* given the frequency of this type of name in English. This semi-morph has since long had an independent meaning component "of

[1] Maybe these should be called *semi-berries.*

Scotch origin", as is witnessed by the comic book personality *Scrooge McDuck* (coined by C. Barks in 1947). Recently, a more commercial use of the semi-prefix has regrettably taken over. [2]

The proposed basis for the analysis of morphemes and semi-morphs is similar in spirit to a proposal for word recognition made by Taft (1994). His unimplemented interactive-activation network was made to isolate stems from the input in order to provide a processing explanation for the prefix-stripping hypothesis (Taft & Forster, 1975). No semantic rationale for isolating stems is necessary in Taft's model. This allows for the independent representation of bound and meaningless stems, like *mit* in *remit*. To explain his experimental results, Taft proposes that the pull towards isolating *im-* in *impeccable* is strong enough to store *peccable* as a bound stem, even though that stem occurs only in the word *impeccable.* This is similar to the isolation of /ʃɔld/ from *shoulders,* that was suggested by the speech error above.[3]

Inflectional and derivational morphemes

A remaining question is whether derivation can make use of the same affix inventory as inflection. In Chapter 1, the problematic German participle *singend (singing)* was introduced, which can function both as a participle and as an adjective. In its adjectival guise, it must be constructed before the inflectional encoding component, for it changes the class of the base verb into an adjective. Still, the adjective is identical in form to the undoubtedly inflectional participle.

An elegant way to describe these twin forms is to have inflection and derivation produce the same form from the same elements by different means: When the inflectional process encounters the stem *sing* with the appropriate tense diacritic, it adds the *-end* suffix to mark it as a participle. The derivational process, on the other hand, has once created a lemma entry in the lexicon. This lemma was made to point to the two morphemes that express this word: *sing* and *end.* In this way, there are now two ways to arrive at the same form. One involves an on-line process of *-end* suffixation caused by a tense diacritic, the other involves the lexical concatenation of the two morphemes that make up a stem. There need not be a more principled relation between the two processes than that they involve the same affix.

This proposal can be tested experimentally: Given the difference between the two hypothesised processes that give rise to *singend,* it must be possible to find evidence for the psychological reality of this analysis. From a lin-

[2]Historically, the form evolved from a Scots Gaelic morpheme meaning "son of". One of my committee members pointed out that a *Chicken McNugget* is thus by some strange genetic engineering the offspring of a nugget.

[3]Taft has reintroduced semantics as a factor that influences morphological analysis in later models (eg. Taft, Liu, & Zhu, in prep).

guistic point of view, this proposal solves the problem of inflectional forms occurring in the lexicon at the price of moving the inflection–derivation distinction towards the psycholinguistic domain: The only difference between inflectional and derivational processes can now lie in the way the form was created. Also, an account of why and when derivation can use the same morphemes as inflection has to be given. This is still an open issue, but certainly not an unsolvable one.

6.3 Language and thought

This section will not cover the full width and depth of this much debated and extremely interesting issue, but will focus on the role that inflectional morphology can play in delineating the relation between language and the conceptual system. We will see that inflectional forms make an interesting case for a certain mapping between the words in the language and the concepts in the mind, which holds that lexical concepts exist for all (head)words of the language and not for its inflectional forms.

Inflections and conceptual classification nodes

In Chapter 2, the debate on the holistic or decomposed nature of lexical concepts was briefly mentioned in the context of the Levelt/Roelofs-model, which chooses the holistic concepts approach (Roelofs, 1992a, 1993, 1997d). Many theorists have argued that lexical concepts are decomposed, and a word like *table* is produced whenever the defining primitives (or features) <has four legs>, <to eat at>, <furniture>, etc. are activated (eg. Bierwisch & Schreuder, 1992; Jackendoff, 1987). The appeal of decomposed lexical concepts is the affirmative answer it gives to the question whether language and thought are different: In the decomposed view, thoughts are expressed in far more basic terms than language. We are forced to decide whether a certain object is referred to by word A or word B in our language, but in thinking one can abstract away from such details.

There are many serious problems with this approach (Fodor, Garrett, Walker, & Parkes, 1980; Fodor, 1998; Roelofs, 1997d). One is how to find out which primitives underly thinking. Is <furniture> indeed a primitive of our thoughts, as assumed above for the description of *table*? And if so, how can we ever prove it is? When trying to define everyday items such as *mountain bike* or *jib*, the need for an increasing number of primitives like <means of transport>, <type of sail> will arise. Also, the items will tend to depend heavily on one another: One way to avoid <type of sail> is to define *jib* as "cloth used for driving a *sail boat*, for which see", but this means we have to define *sail boat* without referring to *jib* or *sail*. Is it possible to define all concepts in a limited set of words? This question is put to the test by second language learner dictionaries like that of Hornby (1995). This dictionary claims to define 65000 terms with a vocabulary of about 3750 words. But the dictionary definitions are usually not very restrictive, and by no means do they sum up the necessary and sufficient conditions for the use of a word, as can be seen from the following examples:

(6.4) a. **mountain bike** a bicycle with a strong frame, large wheels and many gears, intended esp. for use on rough ground.
b. **jib** (nautical) a small sail in front of the MAINSAIL, picture at YACHT.
c. **mainsail** the principal sail on a ship, picture at YACHT
Adapted from Hornby (1995).

The entry for *jib* contains *mainsail*, a word that is not in the defining vocabulary. Looking up *mainsail* does not help much in the understanding of *jib* and in the end, the distinction between the sails can only be made by means of a picture. Clearly, defining the words of English in a restricted vocabulary is an impossible task. Fodor (1998) and Roelofs (1997d) discuss a range of similar arguments against the inclusion of a decomposed conceptual system in a model of language production.

The alternative way to define lexical concepts is to assume that they are holistic (or non-decompositional, in Roelofs' terminology). There is a concept TABLE in the mind, which cannot be analysed any further. The concept thus stands in a one-to-one relationship to the word *table*. This approach is also taken in the study of formal semantics, where it is held that the meaning of most words is actually quite uninteresting. In the PTQ grammar (Montague, 1973), for example, the meaning of *table* is simply **table**′ and only a strictly limited set of words has a more complex meaning (notably quantifiers, adverbs of time, and the like).

The holistic approach to lexical concepts does not completely deny the role of features. The model by Collins and Loftus (1975) contains a layer that is used to relate the concepts with each other. The features that appear here are neither primitive nor defining, but simply express properties of the concepts. It is this layer that explains why people can produce a list of furniture items and which encodes that *boat* and *motorbike* have more in common than *boat* and *strawberry*.

In general, the holistic view on lexical concepts stresses the fact that there is not much difference between words and concepts, apart from their modality. Thinking is given no privileged status above speaking; both processes have to work with a similarly limited set of building blocks. Of course, there are certain ways of non-linguistic thinking that do not suffer from this restriction and this might explain the common feeling that there are thoughts that cannot be expressed in words.

Inflection and concepts

Under a decomposed view of the conceptual system, inflectional forms do not pose much of a problem. The difference between *table* and *tables* can be accounted for by a number primitive, <±plural>. Primitives have to be added for every inflectional category, like number, tense, person, etc. The discussion centres around which primitives are chosen (<±singular> vs. <±plural>) and which values are positive, unmarked, or default (Jakobson, 1939).

For the holistic approach to lexical concepts, however, an interesting problem arises when considering the representation of inflectional forms at the conceptual level: Do concepts map one-to-one to the headwords of the language or to all inflectional forms of the language? Under the latter and most strictly holistic approach, every inflectional form is granted its own lexical concept. As was described in Chapter 2, the Levelt/Roelofs-model takes a different perspective and produces inflectional forms by the combined activation of a lexical concept node (eg. ANCHOR) and a concept classification node (eg. PLURAL). By using conceptual classification nodes to express inflectional variation, the number of lexical concepts is of course drastically reduced.

This means that thought is again an abstraction of language. This time, thought is not claimed to be phrased in more primitive terms than language, but some of the frills and peculiarities of language are not expressed by the conceptual system. Under this assumption, language is a richer and more elaborate expression of our thoughts and this gives this view the intuitive appeal of separating language and thought while avoiding the pitfalls of the decomposed view explained above.

A question that has to be answered is what constitutes the fundamental difference between PLURAL and <has four legs>. The holistic approach refutes decomposition of a word like *table* into anything else but the concept TABLE, while the word *tables* is decomposed into TABLE and PLURAL. Why is the plural primitive granted separate conceptual status that 'have four legs' and other defining features of *tables* do not get? Two parts of the solution to this problem will be discussed in turn.

Efficiency of storage

The first aspect to the solution is a formalisation of the observation made above: Without concept classification nodes, the number of lexical concepts in the system would be more than doubled. It seems logical to introduce a force towards efficiency of storage that propagates reducing the number of conceptual nodes by reusing and combining them. The obvious limitation on reduction is that all words must be unambiguously expressed by one or more conceptual nodes. As we have seen from the dictionary example above, basic expressions like *mountain bike* resist expression by a combination of conceptual nodes, because the definition never matches the true extent of the word.

Variation due to inflection is in a different ball game: Inflectional suffixes add little or no meaning to the stem and they combine with stems in extremely predictable and systematic ways (cf. Chapter 1). The plural *-s* suffix always adds a meaning nuance "more than one" to the concept expressed by the stem. Finally, almost all combinations of stems and applicable inflectional suffixes exist in the language, and this makes the translation from concept classification nodes to inflectional suffixes reliable. Together, transparency of inflectional forms makes it fruitful to decompose them at the conceptual level, while treating all other words holistically.

The same reasoning holds, with even more force, for the properties of language that are confined to the linguistic level, like gender marking or classifier systems. Again, the lexical concepts abstract from these properties to adhere to the economy of storage constraint at the conceptual level. For gender and classifiers, no conceptual classification nodes are necessary at all because they have no conceptual relevance. In the sketch of the Levelt/Roelofs-model given in Chapter 2, gender and classifiers are represented by inherent properties of lemmas. This puts them on a par with word class and similar language specific information.

What cannot be explained by efficiency of storage alone is why other, non-inflectional, transparent combinations of morphemes are not decomposed. Consider *board game*, a transparent compound which, under the abovementioned antagonism between efficiency of storage and transparency, should be represented by the combination of concepts BOARD and GAME and not by a holistic concept BOARDGAME. Recall that BOARDGAME is the representation of choice for any holistic approach. The force of transparency is too strong and efficiency cannot be the only basis for a distinction between PLURAL and <has four legs>.

Using language as a yardstick

As an answer to the problem with distinguishing PLURAL from <has four legs>, one might try to define transparency in such a way that it includes inflection but excludes transparent compounding (*board game*) and transparent derivations. This, however, amounts to defining inflection, an endeavour that has so far been fruitless and harbours a great many pitfalls (see Scalise, 1988; Anderson, 1992, and the discussion in Chapter 1).

Instead of trying to indirectly define inflection, we can also opt for making conceptual structure dependent on whether something is inflectionally expressed in the language: Concept classification nodes are used exclusively to describe the set of inflectional variants each lexical concept covers. Under this view, the stem *table* is expressed by one holistic lexical concept and both the singular *table+ø* and the plural *tables* are expressed by the lexical concept TABLE and a concept classification node, SINGULAR or PLURAL.

Returning to the difference between PLURAL and <has four legs>, we can now say that the existence of PLURAL is backed up by a language that systematically distinguishes between singular and plural word forms. The property of having four legs is not expressed by the language, and there certainly is no productive way to specify whether animals, furniture items, or regattas have four legs or not.

The obvious drawback to this move is that, since inflection is not precisely defined, we also have no practical way to decide whether something is expressed by a concept *cum* classification node or not. This should not bother us, for the validity of a theory does not depend on whether its underlying variables have already been given an operational definition or not. Even without a strict definition, it is clear that, for example, most tenses in English and Dutch are of inflectional origin and a concept classification node for tense, or for present and past, can thus safely be assumed.

Differences between languages

The sketched relationship between lexical concepts and words depends heavily on the facts of the language. One of the fringe benefits of this approach is that it can graciously deal with the differences between languages.[4] The lexical concepts correspond to stems like *table*, *green*, and *to jump:* Those meaning elements that can be used independent of the target language. In formal semantics, these concepts are translated into the constants and predicates of the logic. The specific linguistic expression of lexical concepts may vary in many respects: Obviously, different languages use different stems (*signifiants*) for a particular lexical concept. If no stem is available, languages can choose for a derivation, a compound, or a grammatical construction to express the concept (cf. Bybee, 1985). The particular choice of form does not affect the conceptual system, however, just as it does not affect the constants and predicates of a rendering in predicate logic. Similarly, the conceptual system is free to generate messages that include concepts that require intricate circumlocution in the speaker's language. In the Levelt/Roelofs-model, there is no feedback from the word form level to the conceptual system and to avoid excessively wordy utterances, the speaker must monitor her own output. According to Levelt (1989) and Levelt et al. (1999), both internal speech and overt speech are monitored.

Of course, lexical concepts in different languages are not perfectly identical. It is often claimed that the word *vegetables* does not perfectly translate to Dutch *groenten*, because *groenten* does not include potatoes whereas *vegetables* does. Similarly, some languages have words for concepts that are not of any use in other languages, one might think of words like Dutch *gezelligheid* and (American) English *tenure-track*. However fascinating these differences might be, they reflect differences in culture rather than differences in languages. In The Netherlands, one sees a slow but steady increase of meals that do not adhere to the traditional *aardappelen–groenten–*

[4]This property formed the incentive to formulate the proposal made in this chapter and resulted from valuable discussions with Michael Cysouw.

vlees (potatoes–'vegetables'–meat) structure. This change gives way to a new interpretation of *groenten* that does include potatoes. It would be misguided to claim that a new word *groenten*$_b$ is replacing the old one, instead the culturally determined meaning of the concept GROENTEN is changing.

Next to assigning different words to different lexical concepts, languages differ in the inflectional 'decoration' that words receive. We can define this decoration tentatively as that part of language for which there exists no obvious place in a predicate logic rendering. There are two principal sources of information that are used to create inflectional decoration (as was discussed in the context of agreement in Chapter 2): Information from conceptual classification nodes and from inherent lemma properties. The former gives rise to tense and number inflection in Dutch, and the latter drives gender marking in Dutch and classifiers in Kilivila (Senft, 1996). Because of their conceptual origin, tense and number marking are considered meaningful, while gender marking and classifiers are considered redundant. This is a misleading classification because the relevance of, for example, tense marking is highly language specific. Yukatek Maya (Bohnemeyer, 1999) does not systematically express tense and for a Yukatek speaker, the ubiquitous Germanic tense inflections may seem redundant. Which they technically are, considering that speakers of Germanic languages inflect all the main verbs of a story to mark the same or a similar tense, while it can be easily inferred from the first past tense that the whole story is situated in the past.

The difference between the production of an inflection stemming from the conceptual level and the production of an inflection stemming from the lemma level is small. At the level of the inflectional encoding component, the source of the information is irrelevant. To add a classifier system to a language, all the lemmas of the language have to be assigned a classification. To add a tense system, the conceptual system must be trained to always overtly specify tense in the shape of a concept classification node (Levelt, 1989). What differs is the involvement of the conceptual system in the language specific 'decoration', but the communicative core of the message and the lexical concepts are never affected.

Does a speaker of a tenseless language have a concept classification node for tense? According to the approach defended here, this is not the case.[5] Concept classification nodes arise by requirement of the speaker's language, as do inherent lemma properties. Only speakers of classifier or gender languages have marked their lemmas for masculine or feminine, in possibly most intricate ways. Dutch and German but not English speakers are required to express the social status of the addressee in a concept classification node of social distance that rules the choice between *jij* and *u* or *du* and *Sie* (the familiar and polite forms of *you*, singular).

This limits the dependency of the preverbal message on the target lan-

[5]This assumes that the language has no other forms that are marked for tense.

guage to one very specific domain, that of the concept classification nodes. The set of lexical concepts that form the core of the message can be successfully expressed in any language, given a large enough lexicon. Language specific details are added by the inherent properties of the lemmas and by the additional conceptual information carried by the concept classification nodes. In a comparative study of German and Yukatek Maya, Bohnemeyer (1999) has found that these two languages show little or no difference in their capabilities for expressing event order. This occurs even though German requires tense inflection while Yukatek Maya does not allow it and does not systematically express temporal relationships at all. Also, German and Yukatek renderings of event sequences had the same pragmatic force and the same communicative impact for explaining the order of events to an addressee. Clearly, the presence of certain concept classification nodes and the absence of others does not fundamentally affect the conceptual system.

Required conceptual classifications

One dependency remains: In addition to generating lexical concepts, the conceptual system must supply all the required conceptual classifications that are needed by the target language. In a system with limited feedback from the language level to the conceptual level, there is no immediately obvious solution to this problem. Levelt (1989) has proposed that the conceptual system has learned to foresee the needs of the language system by monitoring its output. Of course, the main objective of the conceptual system is to get the message across. In the light of this objective, the suggested preventive and adaptive strategy of supplying all information that has proven to be essential for the language system is certainly a viable option.

If concept classification nodes are used only for conceptual specification of inflectional affixes, as defended here, the adaptive burden on the conceptual system is further simplified because grammatical classes do require the same conceptual specifications across the board: In Dutch, nouns always require number, and verbs require tense and number (the inflectional frame in fact rests on the same bedrock). Also, young children and certain types of aphasic patients often fail to supply the necessary concept classifications, resulting in speech that contains the intended words, but expresses the message in an uninflected form (De Roo, 1999). For children, there is evidence indicating that inflectional affixes and, in the present theory, concept classifications are acquired in a certain order (Brown, 1973, see Tager-Flusberg, 1997, for a review).

6.4 The status of derivational rules

When we evaluate the system described so far, it does not fundamentally deviate from received linguistic insights, and not from the proposals by linguists that subscribe to the split morphology hypothesis (eg. Beard, 1995 and Anderson, 1992, see also Spencer, 1991). Some extensions were made, most notably in the form of semi-morphs, but otherwise the workings of the inflectional encoding component agree to a large extent with that of inflectional rules in linguistic theory with split morphology. This resemblance I consider a virtue, but it cannot be maintained for derivational morphology.

Not much experimental work in this thesis bears on derivations, and the account given here must remain speculative. It is clear, however, that there is little room for on-line rules creating derivational forms in the Levelt/Roelofs-model. Rather, derivational forms and compounds are generated from a lemma that points to two (or more) morphemes. A simple concatenation device combines these morphemes into the complex stem.[6] The consequences of this are that there is no difference between the various word formation processes because the lexicon lists only that a form consists of some parts and does not represent by which rule this form was created from the parts. Put more concretely, if *readable* is stored in the lexicon, what is stored is that there is a form that consists of the morphemes *read* and *-able*. Although the form was once created by a rule of *-able* suffixation, no connection to this rule has to be made when using the form.

The same processes can apply to the results of the historic noun formation rule Adjective + *-th* → Noun, like *warmth* and *width*. This rule is no longer productive, but that does not have consequences at the lexical level: Without knowing the historical rule, the lexical system will pick up the relationship between these forms and, provided there are enough forms in which it occurs, the system will represent *-th* as a separate morpheme. As was argued above, the same may happen to parts of the language that have no linguistic status at all.

What is the status of derivational rules in this system? First of all, they are called in when new combinations are created. This process might be more widespread than the creation of neologisms, if one assumes that some low-frequent and transparent combinations are not stored in the lexicon, but generated time and again (cf. Schreuder & Baayen, 1995). Whether the generation of these 'new' forms is based on rules or on analogy is an open question, but a sophisticated system is needed to account for the many (sub)regularities that language users can take advantage of when wanting to express something: Consider "they drove by at breakneck speed.". Here, *breakneck* means "in a way that makes it likely that one might brake his neck (and all other bone in the body)". This is a very rare pattern for compounds, that usually exhibit a modifier–head noun pattern, as in *breakpoint* or *bottle-neck*.

Second, derivational rules form a good basis for the analysis of forms into

[6]It is only this latter device that must be more sophisticated for non-concatenative languages, the basic representation with pointers to the constituting morphemes can remain the same.

parts. When encountering the word *entertainer* for the first time, the derivational rule of agentive *-er* suffixation may lead the language system directly to an analysis in terms of *entertain* and *-er*. When encountering a form containing a non-productive suffixes, like *breadth*, the form might first be stored in full and only a later analysis of its neighbouring forms will reveal that the word can be described more parsimoniously as a results of opaque *-th* suffixation of the base *broad*.[7] The larger the inventory of derivational rules, the more lexical items can be analysed into parts without having to revert to a brute-force search for formal regularities in the lexicon.

6.5 Inflectional frames in inflectionally rich languages

In this short section, the four major themes that we have touched upon in this chapter will be tied together in considering the elaborate inflectional systems that have been found in languages around the world. Would the inflectional frame for such languages contain a large number of slots, one for each such dimension? This is not necessarily the case. The gist of the answer is that some inflectional dimensions will use the inflectional frame, and have a slot reserved for them, while others might not.

For acquiring a place in the inflectional frame, an inflectional dimension must be *required*, in the sense that there must always be an affix present that expresses this dimension. Although one may allow for null morphemes in some places, these should be introduced with care. This has two repercussions for inflectionally rich systems: First, the inflectional status of fused categories is unclear. When two inflectional dimensions are systematically expressed by one affix, only one slot needs to be reserved in the inflectional frame for both categories. But when two or more categories are sometimes expressed independently and sometimes together (as is the case for Latin verbal inflections), the number of slots in the resultant inflectional frame is indeterminate. Various solutions are possible and only experimentation can decide which one is preferred by speakers of that particular language.

Second, dimensions that are not always expressed but are more or less optional can be treated like derivations by the system. This does not make these affixes derivational in the linguistic sense, they can still have limited content, no room for opaque semantics, and be applicable across the board. It does not necessitate lexical storage of every variant either, because transparent forms can be created by rule in the lexicon. Compared to extra-lexical inflections, this type of inflection would be more likely to change in the direction of derivation, because it is already formed in the lexicon. The

[7] It might be the case that some speakers have stored *-th* suffixation as a non-productive rule or association. For now, I assume that non-productive implies that there is no derivational regularity describing the relation.

case of Dutch diminutive formation fits this description perfectly: Although the average diminutive is highly predictable and almost inflectional in nature (eg. *boot+je* boat+DIM, meaning 'little boat'), there is room for exceptional forms that are not transparent or do not adhere to the formal criteria for diminutive formation (eg. *groen+tje* green+DIM, 'greenhorn'). This type of irregularity cannot be accounted for in the inflectional encoding component.

The proposed overlap between the inflectional and derivational morpheme inventory may bridge the gap between the behaviour of inflections and derivations even further. One fundamental difference remains: Derivational forms were claimed to be represented by separate lexical concepts and inflectional forms by the combination of a lexical concept and a concept classification node. There are substantial repercussions for the holistic conceptual system of treating a form as derivational instead of inflectional: The number of lexical concepts is inflated and some concept classification nodes might become superfluous. Careful inspection of a range of languages can reveal whether this is a viable option. For cases like the Dutch diminutive system, a separate concept node for the irregular form *groentje* is always necessary and the tendency of diminutive forms to carry an aspect of cuteness (*liefje*, 'love', *reusje*, lit. 'little giant') or relativity (*botsinkje*, 'collision DIM') strengthens the view that separate concepts are indeed at play.

6.6 Inflection and derivation

How does the current proposal relate to Anderson's definition of inflection? Recall that Anderson's definition was cast in terms of the availability to syntax:

(6.5) Inflection is that part of morphology that is accessible to and/or manipulated by the rules of syntax (after Anderson, 1992, p. 83).

In the current proposal, this still holds. Inflectional encoding is always based on the expression of diacritics and these are manipulated by the agreement rules of the grammar. The grammar interfaces between the conceptual system and the lexical system and as such, it has access to both the concept classification nodes and the inherent lemma properties that drive inflectional encoding. Whether a diacritic originates from the conceptual level or from the lemma level should not matter to the grammar.

It was suggested above that in inflectionally rich languages, the phenomena that are considered inflectional by linguists may be divided into a group of extra-lexical inflections (that require a slot in the inflectional frame and are conceptually represented by concept classification nodes) and a group of forms that are of lexical origin (signalled by the fact that they do not occupy a slot and are represented by their own lexical concepts). If evidence for this division would be found, this would have interesting consequences for the inflection–derivation distinction: Because the inflectional frame contains frames for extra-lexical inflections only, the frame could be used as a test of this distinction. Of course, validation of the task over a range of languages is necessary before we can embark on such an approach.

6.7 Conclusions

In this chapter, I first laid out some requirements on when inflectional null morphemes should be assumed. A null morpheme should only be posited in places where affixation is mandatory. The stem should give an unambiguous structural cue to the presence of an affix, such that absence of an affix can also be meaningful.

Next, some extensions and modifications to the Levelt/Roelofs-model were proposed, which can summarised as follows:

▷ The theory should be word-based in the most rigorous sense: No semantic considerations or productivity constraints must be imposed on the morphological system.

▷ Next to real morphemes, bound morphemes, and cranberry morphs, the theory should therefore also posit semi-morphs. These morphs have no independent meaning, can occur in only a limited number of forms, and are not productive. Their contribution to the model is that they allow for more efficiency in storage, because recurring strings can be treated as instances of the same semi-morph.

▷ Inflection and derivation make use of the same morpheme inventory. The problematic German word *sing+end* (from Chapter 1) is ambiguous between a derivational form (the adjective) and an inflectional form (the participle). These two forms come about by different processes, but use the same morphemes.

▷ The boundary between decomposition and non-decomposition at the conceptual level coincides with the distinction between inflection and derivation in the target language. The practical consequences of this are that the conceptual system should contain lexical concepts for each headword. Part of the inflectional variation is steered by the concept classification nodes that express properties of concepts and form the decompositional part of the conceptual system.

▷ When considering the production of various languages, the core of the conceptual system is assumed to be invariant among languages after cultural differences have been accounted for. Of course, lexical concepts will be mapped onto other stems, or combinations of stems, but this does not affect their use in preverbal messages. What does fundamentally differ are the concept classification nodes that are required with certain concepts.

- ▷ Most derivational rules have no role in the everyday production of derivationally complex words. The lexicon simply harbours an entry for every complex word, with a specification of the morphemes that constitute this form. An exception might be made for low-frequent, transparent forms that can be created anew each time they are put to use.

Lastly, the experimentally-demonstrated existence of an inflectional frame reflects which phenomena are really inflectional and which are not. After ample validation across languages, this may offer a way out of the thorny problem of defining inflection.

Summary and conclusions

CHAPTER 7

The split morphology hypothesis

In the **first chapter** of this thesis, the notions inflection and derivation were outlined and the split morphology hypothesis was presented. This hypothesis holds that inflection and derivation are two very different processes and are situated at two different stages of the language production process. Specifically, inflection works on the output of the mental lexicon, adding affixes to the lexical stems. Derivation, on the other hand, creates new forms in the lexicon by recombining existing stems with affixes. These new forms become lexical stems in their own right, with the possibility for meaning specialisation and for renewed application of derivational rules.

Some linguistic problems with the split morphology hypothesis were discussed and possible solutions were outlined. Next, experimental evidence from various sources was presented, which indicates that there is a sharp distinction between the processing of inflection and derivation and between the processing of regular and irregular morphology. Together, these findings make a case for the viability of the split morphology hypothesis in a language production theory.

Psycholinguistic models

In **Chapter 2**, the psycholinguistic literature on the role of morphology in language production was reviewed. Speech errors have been used as a major source of empirical evidence since the early days of psycholinguistics. The error corpora indicate that stems and (inflectional) affixes are treated differently in language production. The hypothesis of an inflectional frame, entertained in this thesis, is based on notions introduced for explaining speech errors.

Several models of language production that include a layer of morphological processing were reviewed. The various approaches stress different aspects of morphological processing and are based on different lines of empirical evidence: Speech errors, production of strong past tense forms, the relation between inflection and derivation, (simplified) language acquisition data, and experimental evidence. A fully satisfactory model of morphological processing should probably incorporate all these considerations. Next, the race models of morphological complexity in language comprehension were taken as examples of how the traditional dichotomies between storage and computation and between parsing and full-form access can be interpreted in a more flexible way.

Without wanting to claim full coverage of all aspects of morphology that were outlined earlier, the Levelt/Roelofs-model of language production was taken as the theoretical framework for this thesis. The split morphology hypothesis finds a natural place in this model, as does an inflectional frame guiding the generation of inflectional affixes. Relevant for our purposes here are the claims that the Levelt/Roelofs-model makes with respect to the representations at the conceptual level, the existence of a lemma level, and the role diacritics play in specifying inflectional forms. The lemma level contains abstract word representations that contain all formal properties of a word (its word class, its gender, etc.) except for its phonological or morphological form. Every lemma contains diacritical parameters (or diacritics) that act as stores for additional conceptual information (like number or politeness) or for agreement information (like structural case). All inflectional marking of a word is steered by information stored in the word's diacritics or by properties that are inherent to the word.

Predictions

The predictions for the empirical part of this thesis were formulated in **Chapter 3**. I outlined the properties of the inflectional frame and sketched its role in the production of morphologically complex words. The benefits of using an inflectional frame lie in its capabilities for checking the presence of all required inflectional affixes. As is evident from its definition, the inflectional frame is not sensitive to derivational morphology:

> **The inflectional frame** is used to guide the process of combining the stem with the inflectional affixes. The frame contains one slot for the stem and one slot for every possible type of inflectional suffix that occurs with this stem. The stem may contain derivational affixes and may be a compound, but will always occupy exactly one slot of the inflectional frame. [repeated from page 61]

The slots that figure in this frame come in three types: Stem slots, prefix slots, and suffix slots. This predicts that different types of affixes (eg. number, person, or tense affixes) can be exchanged, as is borne out by the speech error data. The stipulations for the number of slots result in a 2-slot frame for Dutch nouns (stem + number), a 2-slot frame for Dutch adjectives (stem + gender/number/definiteness), and a 3-slot frame for Dutch verbs (stem + tense + number).

Evidence for inflectional frames

The evidence for the inflectional frame was presented in **Chapter 4** and **Chapter 5**. The task used was the implicit priming task, which is concisely explained in Section 2.5 on page 55. In this task, subjects have to produce words from a restricted set of alternatives. When the words in the set show overlap, subjects are faster to produce the individual words because they can prepare part of their utterance in advance. This is reflected by the preparation effect.

The experimental findings are summarised at the end of the chapters, the major point of theoretical interest being that Dutch verbs do indeed differ from nouns in their inflectional frame. Within a set of N–N compounds, like *werkplaats*, 'workshop' and *werkbij*, 'worker bee', a large preparation effect was observed due to the fact that all items begin with the morpheme *werk*. The presence of a verbal form *werkte (worked)* reduced this effect. This reduction was explained by the fact that in a condition with nouns only, the inflectional frame is constant and can be prepared. In a condition with nouns and one verb, the frame cannot be prepared because nouns require 2-slot frames and verbs require 3-slot frames. On-line access of the inflectional frame leads to partial reduction of the preparation effect.

No reduction was obtained when derived adjectives and diminutives were compared to compounds, indicating that the inflectional frame is sensitive to inflectional morphology only and does not contain slots for derivational affixes. This confirms the last clause of the definition of the inflectional frame given above. A further backup of this claim comes from an experiment in Chapter 5, which mixed N–N compounds with morphologically simple nouns to find no difference between the two.

In sum, these findings plead for a split morphology approach to language production. I suggest that the morphological encoding component of the Levelt/Roelofs-model should be divided into two subcomponents, as was shown in Figure 4.4 on page 90. The *stem encoding component* takes care of the retrieval of derivational forms and compounds from the lexicon, whereas the *inflectional encoding component* adorns these forms with the required inflectional affixes. Stems that are morphologically complex are not stored as a whole. Instead, the lemma contains multiple, numbered pointers to the form (morpheme) lexicon and a morphologically complex form like *starfish* is created by concatenating the morphemes *star* and *fish*. The stem encoding component takes care of this concatenation process. Next to that, it supplies the appropriate inflectional frame on the basis of the word class information on the lemma.

A reduced preparation effect was also found for verbs and monomorphemic nouns that share initial segments (*waadde* and *wagen*, 'waded' and 'wagon'), showing that the effect of the inflectional frame is not restricted to conditions with morphological overlap. A timing proposal was made that explains this effect without positing a direct influence of morphological encoding (inflectional frames) on phonological encoding (see Figure 4.3 on page 82). The crucial assumption underlying the timing proposal is that encoding of the second syllable of *wagen* /waː-ɣə/ must await encoding of the first syllable *and* generation of the inflectional affixes. The generation of affixes is guided by and depends on the availability of the inflectional frame, which can be prepared in a condition with only nouns, but not in a condition with nouns and verbs. Hence, mixing nouns and verbs leads to on-line access to the inflectional frame, which delays generation of the inflectional affixes, which in turn masks part of the benefit of being able to prepare the first syllable.

Two control experiments were run that asserted the finding of an effect of the inflectional frame on nouns and verbs sharing initial segments only. Crucially, the reduction of the preparation was no longer obtained when the inflected verbs were replaced by homophonic nouns, that is, when *waadde (waded)* was replaced by *wade (shroud)*. This shows that the reduction of the preparation effect is solely due to the 3-slot inflectional frame that accompanies *waadde*, and not to its phonological form.

Irregular forms

The inflectional frame hypothesis is corroborated by an experiment using strong past tense verbs (like *zwem–zwom*, 'swim–swam'). Because the past tense of these verbs is not marked by a regular inflectional affix, no tense slot is necessary in their inflectional frame. Indeed, no reduction of the preparation effect was found when strong past tense verbs were mixed with monomorphemic nouns, indicating that both nouns and strong verbs require a 2-slot frame containing slots for the stem and for number.

The behaviour of strong verbs was once more put to the test in another experiment, which used the present tense form of these verbs. The result of this experiment indicates that the present tense form of a strong verb, although in no respect irregular in itself, does associate with an inflectional frame that contains no slot for tense. Apparently, the language production system optimises for the consistent absence of a discernible tense suffix with the present tense stem of a strong verb. The inflectional frame for both present and past forms of strong verbs therefore contains 2 slots: stem and number.

Clearly, the stem alternation between *swim* and *swam* has to be accounted for by the lexicon. In this respect, the results on strong verbs agree with the exception that is made for irregular forms in Perlmutter's formulation of the split morphology hypothesis (cf. 1.4 on page 8). A representation for dealing with irregular forms was proposed: An irregular form has a complex lexical entry in the form lexicon, comparable to what is shown in (7.1b-c). A regular entry is shown for comparison in (7.1a) [repeated from page 94]:

(7.1)		lemma		condition	morpheme	inflectional frame
	a.	CHAIR	→		**stoel**	
	b.	TO SWIM	→	*present:*	**zwem**	Stem + Affix
				past:	**zwom**	Stem + Affix
	c.	MUSEUM	→	*singular:*	**museum**	Stem
				plural:	**musea**	Stem

When the stem encoding component encounters a complex lexical entry, it uses the values of the diacritics to resolve which subentry has to be used. Here, *tense* is used and that diacritic will not be passed on to the inflectional encoding component , because its value has already found linguistical expression. In line with the error-checking redundancy that inflectional frames supply (cf. Chapter 3), an irregular inflectional frame is also specified at the complex lexical entry. For MUSEUM in (7.1c), this irregular frame does not contain a slot for a number affix to ensure that no double marked form **museas* occurs. The listed inflectional frame overrides the regular specification of a 2-slot inflectional frame for nouns.

Number of slots

The definition of the inflectional frame in terms of number of *possible* affixes was backed up by two experiments. One was a replication of the experiment with a verb and monomorphemic nouns, but now run with plural forms throughout (*haatten* and *havens*, 'hated PLURAL' and 'harbours'). The same reduced preparation effect as in the original experiments was obtained. A second experiment mixed plural nouns with singular past tense forms (*haatte* and *havens*, 'hated SINGULAR' and 'harbours'). Both forms bear one actual inflectional affix (*viz.* for number and tense), but a reduced preparation effect due to different inflectional frames was nonetheless obtained.

This finding argues strongly for the use of an inflectional null morpheme: When producing a form like *haatte (hated)*, the stem slot of the inflectional frame is filled with *haat*, the first suffix slot with the tense suffix *-te*, and the second suffix slot with *-ø*. The presence of this latter form shows the inflectional encoding component that all required suffixes have been generated and phonological encoding can start. If no *-ø* suffix were inserted in the slot, the inflectional encoding component could not easily check for the presence of all required affixes. Error checking capabilities were among the reasons that motivated the existence of the inflectional frame.

Implications and extensions

In **Chapter 6**, some extrapolations of the obtained results were made and the linguistic consequences of the findings were considered. First of all, the use of null morphemes was discussed. In general, null morphemes should be avoided because their existence is hard to prove. Some restrictions on the use of *ø* were proposed, which do allow for the appearance of a null suffix in regular inflectional paradigms where one of the alternatives is signalled by no inflectional marker. For production, this means that *haatte (hated)* is generated as *haat+te+ø (hate+ed+*SG*)* and similarly, *anker (anchor)* is generated as *anker+ø (anchor+*SG*)*.

In Chapter 1, a word based morphology was ascribed to the Levelt/Roelofs-model, which means that in the model the meaning of a complex word is not derived from the combination of its morphemes, but from the word as a whole. In essence, morphemes have no meaning in every day use and function only as a rule of thumb for predicting the meaning of a new or low-frequent complex word. Such a radical departure from the traditional semantic definition of morpheme allows us to account for the puzzling speech errors in which non-morphemic parts of words are exchanged:

(7.2) Lackner and Goldstein → Goldner and Lackstein (Garrett, 1975)

A model is proposed in which the language production system uses all redundancy in the lexicon to arrive at the most efficient representation in terms of storage. In such a system, the semi-morphs *-ner* and *-stein* might be isolated to efficiently account for the group of names that follow the same pattern. This leads to the identification of *lack* in *Lackner* and, to use another example featured in the chapter, of *cape* in *capable.* Neither morpheme contributes to the meaning of the whole, but it is argued that this is no different from the generally accepted recognition of *way* in *wayward* or *cran* in *cranberry.*

To account for some of the linguistic problems with the split morphology hypothesis that were mentioned in Chapter 1, it is proposed that inflection and derivation make use of the same morpheme inventory. Thus there are two ways to arrive at the German word *singend (singing, present participle or adjective)*: The first one is inflectional and involves on-line suffixation of the stem *sing* with *-end* by virtue of the tense diacritic. The second one is derivational and involves the concatenation of the two morphemes that the lemma SINGEND points to: *sing* and *end.* The same form can have two origins that correspond to its two possible, and quite different, uses.

A large section of Chapter 6 is dedicated to the relation between the conceptual system and the inflectional system. The Levelt/Roelofs-model holds that the conceptual system is non-decomposed, that is, the word *table* is linked to a lexical concept TABLE, and not to a decomposed conceptual representation of the form ‹has four legs›, ‹to eat at›, ‹furniture›, etc. To account for the language-specific encoding of number on nouns, tense on verbs, etc., I propose to add *concept classification nodes* that specify such traits at the conceptual level. These classification nodes add a level of decomposition to the conceptual system that precisely coincides with the subset of conceptually relevant inflectional categories in the language. The classification nodes are therefore highly language specific. The remaining part of inflectional variation in a language, which is not steered by conceptual specification, reflects linguistic properties of words or properties of syntactic structure, like gender and structural case. This type of inflection has no bearing on the conceptual system. The same is true for derivation and compounding: Whether a lexical concept is expressed by a monomorphemic word or by a combination of morphemes is not relevant to the conceptual system.

A further claim that is made in this chapter is that all language variation is in essence variation in concept classification nodes. Such variation is peripheral to the conceptual system, whose core comprises concepts and lexical concepts. This means that the influence of a language on the conceptual system of its speakers is very limited, especially when compared to the influence of the cultural background.

The status of derivational rules in the resulting model is rather marginal from a linguistic perspective. Whereas inflectional rules are used to create inflected forms on the fly, derivational rules are only used to introduce new forms. Because low-frequent, semantically transparent words may be created anew on every use, the derivational rules play a slightly less marginal role in everyday language use than the incidence of neologisms suggests. Secondly, derivational rules are extremely reliable clues for the decomposition of lexical items into parts. When the system contains a rule of *-able* suffixation, this means that any word formed from an appropriate morpheme and *-able* can be successfully analysed into two parts, rendering a sparse representation without having to scan the complete lexicon for recurrent form elements that can be optimised for.

Finale

In all, the two linguistic hypotheses of split morphology and word based morphology can be successfully transfered to the psycholinguistic domain. The empirical results presented in Chapters 4 and 5 together with the theoretical considerations in Chapter 6 make out a strong case for a production model of morphology that assumes a radical distinction between inflection and derivation, with inflection as an extralexical process and with derivation describing regularity inside the lexicon. The Levelt/Roelofs-model contained a word-based morphology to start with, which was modified to ignore the remaining semantic constraints on morphemehood. This allows for the recognition of semi-morphs and for an unrestricted analysis of words into opaque combinations of (possibly unrelated) morphemes. A conceptual strand to the theory of inflection was included, which tentatively relates inflection to that part of the conceptual system that is language dependent.

Obviously, the study of the role of morphology in psycholinguistics owes a great deal to the results of several decades of linguistic thinking. Hopefully, the psycholinguistic measures might one day be able to help decide on those issues that have so far resisted linguistic analysis, like the distinction between inflection and derivation.

Don't loof for a solution at the end of this. I'm too busy trying to stay alive to help you out. What you've heard isn't true, it's not always a party for the geniuses. Women leave us for other guys and we get thrown out for not paying the rent and people steal things from us, too. I love you.

Al Hansen, Performance/Live art notes

Bibliography

Allerton, D. J. (1992). Disagreement about agreement. *English Studies, 5*, 458–470.

Anderson, S. R. (1982). Where's morphology? *Linguistic Inquiry, 13*, 571–612.

Anderson, S. R. (1988). Inflection. In M. Hammond & M. Noonan (Eds.), *Theoretical morphology: approaches in modern linguistics* (pp. 23–43). San Diego, CA: Academic Press.

Anderson, S. R. (1992). *A-morphous morphology.* Cambridge: Cambridge University Press.

Anshen, F., & Aronoff, M. (1988). Producing morphologically complex words. *Linguistics, 26*, 641–655.

Aronoff, M. (1976). *Word formation in generative grammar.* Cambridge, MA: MIT Press.

Baayen, H., Burani, C., & Schreuder, R. (1997). Effects of semantic markedness in the processing of regular nominal singulars and plurals in Italian. In G. Booij & J. van Marle (Eds.), *Yearbook of morphology 1996* (pp. 13–33). Dordrecht, The Netherlands: Kluwer.

Baayen, R. H., Dijkstra, T., & Schreuder, R. (1997). Singulars and plurals in Dutch: Evidence for a parallel dual-route model. *Journal of Memory and Language, 37*, 94–117.

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). *The* CELEX *lexical database* [CD-ROM]. Philadelphia: Linguistic Data Consortium.

Badecker, W., & Caramazza, A. (1989). A lexical distinction between inflection and derivation. *Linguistic Inquiry, 20*, 108–116.

Beard, R. (1995). *Lexeme-morpheme base morphology: A general theory of inflection and word formation.* New York: SUNY Press.

Berg, T. (1987). The case against accomodation: evidence from German speech error data. *Journal of Memory and Language, 26*, 277-299.

Berko, J. (1958). The child's learning of English morphology. *Word, 14*, 150–177.

Bierwisch, M., & Schreuder, R. (1992). From concepts to lexical items. *Cognition, 42*, 23–60.

Bloomfield, L. (1930). German ç and x. *Le maître phonétique, 20,* 27–28.

Bock, K., & Miller, C. A. (1991). Broken agreement. *Cognitive Psychology, 23,* 45–93.

Bohnemeyer, J. (1999). *Time relations in discourse: evidence from a comparative approach to Yukatek Maya.* Unpublished doctoral dissertation, Tilburg University.

Booij, G. (1993). Against split morphology. In G. Booij & J. van Marle (Eds.), *Yearbook of morphology 1993* (pp. 27–49). Dordrecht, The Netherlands: Kluwer.

Booij, G. (1996). Inherent versus contextual inflection and the split morphology hypothesis. In G. Booij & J. van Marle (Eds.), *Yearbook of morphology 1995* (pp. 1–16). Dordrecht, The Netherlands: Kluwer.

Booij, G., & van Santen, A. (1998). *Morfologie: de woordstructuur van het Nederlands* [Morphology: the word structure of Dutch] (2nd ed.). Amsterdam: Amsterdam University Press.

Brown, R. (1973). *A first language.* Cambridge, MA: Harvard University Press.

Bybee, J. L. (1985). *Morphology: A study of the relation between meaning and form.* Amsterdam/Philadelphia: John Benjamins.

Bybee, J. L. (1988). Morphology as lexical organization. In M. Hammond & M. Noonan (Eds.), *Theoretical morphology: approaches in modern linguistics* (pp. 119–141). San Diego, CA: Academic Press.

Bybee, J. L. (1995). Diachronic and typological properties of morphology and their implications for representation. In L. B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 225–246). Hillsdale, NJ: Lawrence Erlbaum Associates.

Bybee, J. L., & Slobin, D. I. (1982). Rules and schemas in the development and use of the English past tense. *Language, 58,* 265–289.

Caramazza, A., Laudanna, A., & Romani, C. (1988). Lexical access and inflectional morphology. *Cognition, 28,* 297–332.

Chialant, D., & Caramazza, A. (1995). Where is morphology and how is it processed? The case of written word recognition. In L. B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 55–76). Hillsdale, NJ: Erlbaum.

Chomsky, N., & Halle, M. (1968). *The sound pattern of English.* New York: Harper and Row.

Clahsen, H. (in press). Lexical entries and rules of language: a multidisciplinary study of German inflection. *Behavioral and Brain Sciences.*

Clahsen, H., & Almazán, M. (1998). Syntax and morphology in children with William's syndrome. *Cognition, 68,* 167–198.

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.

Coleman, E. B. (1964). Generalizing to a language population. *Psychological Reports, 14,* 219–226.

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

Cutler, A. (1980). Errors of stress and intonation. In V. A. Fromkin (Ed.), *Errors in linguistic performance: Slips of the tongue, ear, pen, and hand* (pp. 67–80). New York: Academic Press.

Cutler, A. (1982). The reliability of speech error data. In A. Cutler (Ed.), *Slips of the tongue and language production* (pp. 7–28). Berlin: Mouton. (also appeared as *Linguistics – special issue, 19,* 1982)

Cutler, A., Hawkins, J. A., & Gilligan, G. (1985). The suffixing preference: A processing explanation. *Linguistics, 23,* 723–758.

Daugherty, K. G., & Seidenberg, M. S. (1994). Beyond rules and exceptions: A connectionist approach to inflectional morphology. In S. D. Lima, R. L. Corrigan, & G. K. Iverson (Eds.), *The reality of linguistic rules.* Amsterdam: Benjamins.

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., Ferreira, V. S., & Bock, K. (1999). Binding, attention, and exchanges. Commentary to "A theory of lexical access in speech production" by Levelt, Roelofs and Meyer. *Behavioral and Brain Sciences, 22,* 41–42.

Derwing, B. L., & Skousen, R. (1994). Productivity and the English past tense: Testing Skousen's analogy model. In S. D. Lima, R. L. Corrigan, & G. K. Iverson (Eds.), *The reality of linguistic rules* (pp. 193–218). Amsterdam/Philadelphia: John Benjamins.

Eberhard, K. M. (1997). The marked effect of number on subject–verb agreement. *Journal of Memory and Language, 36,* 147–164.

Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development.* Cambridge, MA: MIT Press.

Fay, D. (1980). Transformational errors. In V. A. Fromkin (Ed.), *Errors in linguistic performance: Slips of the tongue, ear, pen, and hand.* New York: Academic Press.

Feldman, L. B. (1994). Beyond orthography and phonology: Differences between inflections and derivations. *Journal of Memory and Language, 33,* 442–470.

Feldman, L. B., Frost, R., & Pnini, T. (1995). Decomposing words into their constituent morphemes: Evidence from English and Hebrew. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 947–960.

Fodor, J. A. (1998). *Concepts: where cognitive science went wrong.* Oxford: Clarendon Press.

Fodor, J. A., Garrett, M. F., Walker, E. C. T., & Parkes, C. H. (1980). Against definitions. *Cognition, 8*, 263–367.

Frauenfelder, U. H., & Schreuder, R. (1992). Constraining psycholinguistic models of morphological processing and representation: The role of productivity. In G. Booij & J. van Marle (Eds.), *Yearbook of morphology 1991* (pp. 165–183). Dordrecht, The Netherlands: Kluwer.

Fromkin, V. A. (1973). *Speech errors as linguistic evidence.* The Hague: Mouton.

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.

Garrett, M. F. (1988). Processes in language production. In F. J. Newmeyer (Ed.), *Linguistics: The Cambridge survey. Vol. III. Biological and psychological aspects of language.* Cambridge, MA: Harvard University Press.

Goldberg, E., & Obler, L. K. (1997). Affixational morphology in jargonaphasia. *Brain and Language, 60*, 89–93.

Greenberg, J. (1954). A quantitave approach to the morphological typology of language. In R. F. Spencer (Ed.), *Method and perspective in anthropology* (pp. 192–220). Minneapolis: University of Minnesota. (Reprinted in International Journal of American Languages, 1960, *26*, 178–194.)

Haarman, H. J., & Kolk, H. H. J. (1992). The production of grammatical morphology in Broca's and Wernicke's aphasics: Speed and accuracy factors. *Cortex, 28*, 97–112.

Halle, M. (1973). Prolegomena to a theory of word-formation. *Linguistic Inquiry, 4*, 3–16.

Hare, M., & Elman, J. L. (1995). Learning and morphological change. *Cognition, 56*, 61–98.

Hare, M., Elman, J. L., & Daugherty, K. G. (1995). Default generalisation in connectionist networks. *Language and Cognitive Processes, 10*, 601–630.

Haspelmath, M. (1996). Word-class-changing infection and morphological theory. In G. Booij & J. van Marle (Eds.), *Yearbook of morphology 1995* (pp.

43–66). Dordrecht, The Netherlands: Kluwer.

Hays, W. L. (1994). *Statistics* (Fifth ed.). Fort Worth, TX: Harcourt Brace College Publishers.

Henderson, L. (1989). On mental representation of morphology and its diagnosis by measures of visual access speed. In W. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 357–391). Cambridge, MA: MIT Press.

Hockett, C. (1958). Two models of grammatical description. In M. Joos (Ed.), *Readings in linguistics* (Second ed.). Chicago: University of Chicago Press.

Hooper, J. B. (1976). *An introduction to natural generative phonology.* New York: Academic Press.

Hornby, A. S. (1995). *Oxford advanced learner's dictionary of current English* (Fifth ed.). Oxford: Oxford University Press.

Huyhn, H. S., & Feldt, L. S. (1976). Estimation of the Box correction for degrees of freedom from sample data in randomized block and split plot designs. *Journal of Educational Statistics, 1,* 69–82.

Jackendoff, R. (1975). Morphological and semantic regularities in the lexicon. *Language, 51,* 639–671.

Jackendoff, R. (1987). *Consciousness and the computational mind.* Cambridge, MA: MIT Press.

Jaeger, J. J., Lockwood, A. H., Kemmerer, D. L., Vanvalin, R. D., Murphy, B. W., & Khalak, H. G. (1996). A positron emission tomographic study of regular and irregular verb morphology in English. *Language, 72,* 451–497.

Jakobson, R. (1939). Signe zéro. In *Roman Jakobson, selected writings III.* The Hague: Mouton.

Janssen, D. P. (1998). *The relativity of storage versus computation in producing language.* Poster presented at the OTS Conference on Storage and Computation in Linguistics, Utrecht, The Netherlands.

Jescheniak, J.-D., & Levelt, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 824–843.

Kempen, G. (1999). *Human grammatical coding.* (Manuscript, Leiden University)

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

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

Kim, J. J., Pinker, S., Prince, A., & Prasada, S. (1991). Why no mere mortal has ever flown out to center field. *Cognitive Science, 15,* 173–218.

Kiparsky, P. (1973). 'Elsewhere' in phonology. In S. R. Anderson & P. Kiparsky (Eds.), *A Festschrift for Morris Halle* (pp. 93–106). New York: Holt, Rinehart, and Winston.

Kiparsky, P. (1982). *Explanation in phonology.* Dordrecht, The Netherlands: Foris.

Kirk, R. E. (1982). *Experimental design: Procedures for the behavioral sciences* (Second ed.). Pacific Grove, CA: Brooks/Cole.

Kuczaj, S. A., II. (1977). The acquisition of regular and irregular past tense forms. *Journal of Verbal Learning and Verbal Behavior, 16,* 589–600.

Kuczaj, S. A., II. (1978). Children's judgements of grammatical and ungrammatical irregular past-tense verbs. *Child Development, 49,* 319–326.

Kuryłowicz, J. (1964). *The inflectional categories of Indo-European.* Heidelberg, Germany: Carl Winter.

Laine, M., Niemi, J., Koivuselkä-Sallinen, P., & Hyönä, J. (1995). Morphological processing of polymorphemic nouns in a highly inflecting language. *Cognitive Neuropsychology, 12,* 457–502.

Laudanna, A., Badecker, W., & Caramazza, A. (1992). Processing inflectional and derivational morphology. *Journal of Memory and Language, 31,* 333–348.

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.

Ling, C. X., & Marinov, M. (1993). Answering the connectionist challenge: a symbolic model of learning the past tenses of English verbs. *Cognition, 49,* 235–290.

MacKay, D. G. (1976). On the retrieval and lexical structure of verbs. *Journal of Verbal Learning and Verbal Behavior, 15,* 169–182.

MacKay, D. G. (1979). Lexical insertion, inflection, and derivation: Creative processes in word production. *Journal of Psycholinguistic Research, 8,* 477-498.

MacWhinney, B., & Leinbach, J. (1991). Implementations are not conceptualizations: Revising the verb learning model. *Cognition, 40,* 121–157.

Marcus, G. F., Pinker, S., Ullman, M., Hollander, M., Rosen, T. J., & Xu, F. (1992). *Overregularization in language acquisition.* University of Chicago Press.

Marslen-Wilson, W., & Tyler, L. K. (1998). Rules, representations, and the English past tense. *Trends in cognitive sciences, 2,* 428–435.

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

33.

Marslen-Wilson, W. D., & Tyler, L. K. (1997). Dissociating types of mental computation. *Nature, 387,* 592–594.

Matthews, P. (1974). *Morphology.* Cambridge: Cambridge University Press.

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.

Meyer, A. S. (1992). Investigation of phonological encoding through speech error analyses: Achievements, limitations, and alternatives. *Cognition, 42,* 181–211.

Miceli, G., & Caramazza, A. (1988). Dissociation of inflectional and derivational morphology. *Brain and Language, 35,* 24–65.

Monsell, S. (1985). Repetition and the lexicon. In A. W. Ellis (Ed.), *Progress in the psychology of language* (Vol. 2). London: Lawrence Erlbaum.

Montague, R. (1973). The proper treatment of quantification in ordinary English. In J. Hintikka, J. Moravcsik, & P. Suppes (Eds.), *Approaches to natural language: Proceedings of the 1970 Stanford workshop on grammar and semantics* (pp. 221–242). Dordrecht, The Netherlands: D. Reidel.

Niemi, J., Laine, M., & Tuominen, J. (1994). Cognitive morphology in Finnish: Foundations of a new model. *Language and Cognitive Processes, 9,* 423–446.

Penke, M., Weyerts, H., Gross, M., Zander, E., Münte, T. F., & Clahsen, H. (1997). How the brain processes complex words: An event-related potential study of German verb inflections. *Cognitive Brain Research, 6,* 37–52.

Perlmutter, D. M. (1988). The split morphology hypothesis. In M. Hammond & M. Noonan (Eds.), *Theoretical morphology: approaches in modern linguistics* (pp. 79–100). San Diego, CA: Academic Press.

Pesetsky, D. (1995). *Zero syntax: experiencers and cascades.* Cambridge, MA: MIT Press.

Pinker, S. (1991). Rules of language. *Science, 253,* 530–535.

Pinker, S., & Prince, A. (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition, 28,* 73–193.

Plank, F. (1994). Inflection and derivation. In R. E. Asher (Ed.), *The encyclopedia of language and linguistics* (Vol. 3, pp. 1671–1678). Oxford: Pergamon Press.

Plunkett, K., & Marchman, V. A. (1991). U-shaped learning and frequency effects in a multi-layered perceptron. *Cognition, 39,* 43–102.

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

Rietveld, T., & van Hout, R. (1993). *Statistical techniques for the study of language and language behaviour.* Berlin: Mouton de Gruyter.

Roelofs, A. (1992a). *Lemma retrieval in speaking: A theory, computer simulations, and empirical data.* Unpublished doctoral dissertation, Nijmegen University, The Netherlands. (NICI Technical Report 92-08)

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

Roelofs, A. (1993). Testing a non-decompositional theory of lemma retrieval in speaking: Retrieval of verbs. *Cognition, 47,* 59–87.

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

Roelofs, A. (1997a). Morpheme frequency in speech production: Testing WEAVER. In G. Booij & J. van Marle (Eds.), *Yearbook of morphology 1996* (pp. 135–154). Dordrecht, The Netherlands: Kluwer.

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

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

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

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., & Baayen, H. (submitted). Morphology by itself in planning the production of words.

Roelofs, A., Baayen, H., & van den Brink, D. (submitted). Semantic transparency in producing polymorphemic words.

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.

De Roo, E. M. (1999). *Agrammatic grammar: functional categories in agrammatic speech.* Unpublished doctoral dissertation, Holland Institute of Linguistics.

Rumelhart, D., & McClelland, J. (1986). On learning the past tenses of English verbs. In J. McClelland, D. Rumelhart, & The PDP Research Group (Eds.), *Parallel distributed processing* (Vol. 2, pp. 217–271). Cambridge, MA: MIT Press.

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.

Scalise, S. (1988). Inflection and derivation. *Linguistics, 26*, 561–581.

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.

Schreuder, R., & Baayen, R. H. (1997). How complex simplex words can be. *Journal of Memory and Language, 37*, 118–139.

Schriefers, H., Friederici, A., & Graetz, P. (1992). Inflectional and derivational morphology in the mental lexicon: Symmetries and assymmetries in repetition priming. *The Quarterly Journal of Experimental Psychology, 44A*, 373–390.

Seidenberg, M. S. (1995). Visual word recognition: An overview. In J. L. Miller & P. D. Eimas (Eds.), *Speech, language, and communication* (Vol. 11, Second ed., pp. 137–179). San Diego, CA: Academic Press.

Senft, G. (1996). *Classificatory particles in Kilivila.* New York: Oxford University Press.

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

Skousen, R. (1989). *Analogical modeling of language.* Dordrecht, The Netherlands: Kluwer.

Spencer, A. (1991). *Morphological theory: an introduction to word structure in generative grammar.* Oxford: Blackwell.

Stemberger, J. P. (1985). An interactive activation model of language production. In A. Ellis (Ed.), *Progress in the psychology of language* (Vol. 1, pp. 143–186). London: Lawrence Erlbaum.

Stemberger, J. P. (1994). Rule-less morphology at the phonology–lexicon interface. In S. D. Lima, R. L. Corrigan, & G. K. Iverson (Eds.), *The reality of linguistic rules* (pp. 147–169). Amsterdam/Philadelphia: John Benjamins.

Stemberger, J. P., & MacWhinney, B. (1986). Form-oriented inflectional errors in language processing. *Cognitive Psychology, 18*, 329–354.

Taft, M. (1994). Interactive-activation as a framework for understanding morphological processing. *Language and Cognitive Processes, 9*, 271–294.

Taft, M., & Forster, K. I. (1975). Lexical storage and the retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior, 14*, 638–647.

Taft, M., Liu, Y., & Zhu, X. (in prep). Morphemic processing in reading Chinese. In A. Inhoff, J. Wang, & H.-C. Chen (Eds.), *Reading Chinese script: A cognitive analysis.* Hillsdale, NJ: Erlbaum.

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

Van Turennout, M., Hagoort, P., & Brown, C. M. (1997). Electrophysiological evidence on the time course of semantic and phonological processes in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 787–806.

Vigliocco, G., Butterworth, B., & Garrett, M. F. (1996). Subject-verb agreement in Spanish and English: Differences in the role of conceptual factors. *Cognition, 61*, 261–298.

Vigliocco, G., Hartsuiker, R. J., Jarema, G., & Kolk, H. H. J. (1996). One or more labels on the bottles? Notional concord in Dutch and French. *Language and Cognitive Processes, 11*, 407–442.

Vigliocco, G., & Nicol, J. (1998). Separating hierarchical relations and word order in language production: Is proximity concord syntactic or linear? *Cognition, 68*, 13–29.

De Vries, J. (1971). *Nederlands etymologisch woordenboek.* Leiden: E J Brill.

Williams, E. (1981). On the notions 'lexically related' and 'head of a word'. *Linguistic Inquiry, 12*, 245–274.

Materials

APPENDIX A

Materials for Experiment 1

Prompts and response words for the homogeneous sets. Heterogeneous sets were created by recombining pairs from different homogeneous sets.

variable sets

prompt – response		translation	
huis	– *bouwde*	house	– build PAST
auto	– *bouwjaar*	car	– year of construction
lego	– *bouwdoos*	Lego	– box of building blocks
misdaad	– *strafte*	crime	– punish PAST
voetbal	– *strafschop*	soccer	– penalty
score	– *strafpunt*	score	– penalty point
arbeid	– *werkte*	labour	– work PAST
stress	– *werkdruk*	stress	– workload
opdracht	– *werkboek*	exercise	– workbook

constant sets

prompt – response		translation	
geld	– *bouwfonds*	money	– building society
stijl	– *bouwkunst*	style	– architecture
trap	– *bouwlift*	stairs	– lift used for construction work
advocaat	– *strafrecht*	lawyer	– criminal law
brief	– *strafport*	letter	– surcharge
sovjet	– *strafkamp*	Soviet	– prison camp
honing	– *werkbij*	honey	– worker bee
overleg	– *werkgroep*	consultation	– work group
garage	– *werkplaats*	garage	– workshop

Materials for Experiment 2

Prompts and response words for the homogeneous sets. Heterogeneous sets were created by recombining pairs from different homogeneous sets.

variable sets

prompt – response		translation	
handschrift	– *leesbaar*	hand writing	– readable
buro	– *leeslamp*	desk	– reading lamp
voordracht	– *leesbeurt*	lecture	– lecture
droogte	– *brandbaar*	dryness	– combustible
focus	– *brandpunt*	focus	– focus
uitgang	– *branddeur*	exit	– fire exit
stoel	– *draaibaar*	chair	– revolving
telefoon	– *draaischijf*	telephone	– dial
zee	– *draaikolk*	sea	– whirlpool

constant sets

prompt – response		translation	
bieb	– *leeszaal*	library	– reading room
vakantie	– *leesvoer*	holiday	– reading matter
ogen	– *leesbril*	eyes	– reading glasses
kazerne	– *brandweer*	barracks	– fire brigade
stempel	– *brandmerk*	stamp	– brand
juwelen	– *brandkast*	jewels	– safe
winkel	– *draaideur*	shop	– revolving door
film	– *draaiboek*	film	– script
speelgoed	– *draaitol*	toys	– top

Materials for Experiment 3

Prompts and response words for the homogeneous sets. Heterogeneous sets were created by recombining pairs from different homogeneous sets.

variable sets

prompt – response		translation	
bungalow	– *huisje*	bungalow	– house DIM
dokter	– *huisarts*	doctor	– general practioner
moeder	– *huisvrouw*	mother	– house wife
ruimte	– *vakje*	room	– box DIM
staking	– *vakbond*	strike	– union
ambacht	– *vakman*	trade	– expert
snoep	– *zuurtje*	candy	– sourball DIM
chagrijn	– *zuurpruim*	chagrin	– sourpuss
lucht	– *zuurstof*	air	– oxygen

constant sets

prompt – response		translation	
school	– *huiswerk*	school	– home work
rotzooi	– *huisvuil*	mess	– garbage
butler	– *huisknecht*	butler	– domestic servant
faculteit	– *vakgroep*	faculty	– department
jargon	– *vakterm*	terminology	– jargon
krant	– *vakblad*	newspaper	– professional journal
worst	– *zuurkool*	sausage	– sauerkraut
azijn	– *zuurgraad*	vinegar	– acidity
kermis	– *zuurstok*	fair	– stick of rock

Materials for Experiment 4

Prompts and response words for the homogeneous sets. Heterogeneous sets were created by recombining pairs from different homogeneous sets.

variable sets

prompt – response		translation	
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

constant sets

prompt – response		translation	
gokspel	– *poker*	gambling game	– poker
contract	– *polis*	contract	– insurance
zuinig	– *pover*	thrifty	– meager, poor
oase	– *water*	oasis	– water
gebak	– *wafel*	cakes	– waffle
pistool	– *wapen*	pistol	– weapon
toekomst	– *heden*	future	– present
afkeer	– *hekel*	dislike	– aversion
slang	– *hevel*	tube	– siphon

Materials for Experiment 5

Prompts and response words for the homogeneous sets. Heterogeneous sets were created by recombining pairs from different homogeneous sets.

variable sets

prompt – response		translation	
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

constant sets

prompt – response		translation	
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

Materials for Experiment 6

Prompts and response words for the homogeneous sets. Heterogeneous sets were created by recombining pairs from different homogeneous sets.

variable sets

prompt – response		translation	
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

constant sets

prompt – response		translation	
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	– *sport*	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

Materials for Experiment 7

Prompts and response words for the homogeneous sets. Heterogeneous sets were created by recombining pairs from different homogeneous sets.

variable sets

prompt – response		translation	
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

constant sets

prompt – response		translation	
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

Materials for Experiment 8

Prompts and response words for the homogeneous sets. Heterogeneous sets were created by recombining pairs from different homogeneous sets.

variable sets

prompt – response		translation	
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

constant sets

prompt – response		translation	
gokspel	– *poker*	gambling game	– poker
contract	– *polis*	contract	– insurance
zuinig	– *pover*	thrifty	– meager, poor
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

Materials for Experiment 9

Prompts and response words for the homogeneous sets. Heterogeneous sets were created by recombining pairs from different homogeneous sets.

first subexperiment

Type	prompt – response		translation	
ME	*zakdoek*	– *nieste*	handkerchief	– sneezed
	kriebel	– *jeukte*	itching	– itched
	hond	– *blafte*	dog	– barked
	koning	– *heerste*	king	– ruled
PE	*paard*	– *ruiter*	horse	– horseman
	drank	– *kater*	drinks	– hangover
	olie	– *boter*	oil	– butter
	rivier	– *water*	river	– water
PI-A	*perceel*	– *kavel*	lot	– plot
	snoer	– *kabel*	wire	– cable
	haven	– *kade*	harbour	– quay
	peddel	– *kano*	paddle	– canoe
PI-B	*toespraak*	– *rede*	speech	– speech
	puzzel	– *rebus*	puzzle	– rebus
	storm	– *regen*	storm	– rain
	wet	– *regel*	law	– rule

second subexperiment

Type	prompt – response		translation	
ME	*geld*	– *spaarde*	money	– saved
	traan	– *huilde*	tear	– cried
	kind	– *groeide*	child	– grew
	huis	– *bouwde*	house	– built
PE	*knie*	– *enkel*	knee	– ankle
	zilver	– *nikkel*	silver	– nickel
	vierkant	– *cirkel*	square	– circle
	afkeer	– *hekel*	dislike	– aversion
PI-A	*negen*	– *zeven*	nine	– seven
	spier	– *zenuw*	muscle	– nerve
	giraffe	– *zebra*	giraffe	– zebra
	winnaar	– *zege*	winner	– victory
PI-B	*corset*	– *boezem*	corset	– bosom
	stoptrein	– *boemel*	slow train	– slow train
	straf	– *boete*	punishment	– penalty
	god	– *boeddha*	god	– buddha

Materials for Experiment 10

Prompts and response words for the homogeneous sets. Heterogeneous sets were created by recombining pairs from different homogeneous sets.

variable sets

prompt – response	translation
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

constant sets

prompt – response	translation
stoel – *kruk*	chair – stool
peuter – *creche*	todler – 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

Materials for Experiment 11

Prompts and response words for the homogeneous sets. Heterogeneous sets were created by recombining pairs from different homogeneous sets.

variable sets

prompt – response		translation	
storm	– *windvlaag*	gale	– squall
jenever	– *whisky*	gin	– whiskey
bedrijf	– *winkel*	business	– shop
aarde	– *potgrond*	soil	– potting soil
tango	– *polka*	tango	– polka
brief	– *porto*	letter	– postage
slager	– *hakblok*	butcher	– chopping-block
cavia	– *hamster*	cavia	– hamster
ridder	– *harnas*	knight	– harness

constant sets

prompt – response		translation	
vlag	– *wimpel*	flag	– streamer
zomer	– *winter*	summer	– winter
tent	– *wigwam*	tent	– wigwam
regencape	– *poncho*	raincoat	– poncho
maaltijd	– *portie*	meail	– serving
dijk	– *polder*	dike	– polder
zadel	– *halster*	saddle	– halter
markt	– *handel*	market	– trade
station	– *halte*	station	– (bus)stop

Materials for Experiment 12

Prompts and response words for the homogeneous sets. Heterogeneous sets were created by recombining pairs from different homogeneous sets.

variable sets

prompt –	response	translation	
vijand	– *<u>haatte</u>*	enemy	– hated
kip	– *hanen*	chicken	– roosters
spijker	– *hamers*	nail	– hamers
gok	– *<u>raadde</u>*	gamble	– betted
trend	– *rages*	trend	– hypes
scheur	– *rafels*	tear	– loose ends
moord	– *<u>doodde</u>*	murder	– killed
nier	– *donors*	kidney	– donors
ei	– *dooiers*	egg	– yolks

constant sets

prompt –	response	translation	
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

On the use of F'

APPENDIX B

Throughout this thesis, the value of F' is reported in favour of the more traditional analysis in terms of F_1 and F_2. The reason for this is explained in this appendix.

Random factors

A factor in an experimental design is called a fixed factor when it contains all the relevant levels of that factor in reality. If we, for example, compare the stability of three types of dinghies, we are not interested in all possible other types of boats. Our experiment captures all relevant levels of the factor at hand and we do not want to extrapolate the results beyond the levels (actual dinghies) tested.

The opposite case is formed by random factors: Here, only a limited and random selection of the possible levels is tested, but we want to extend the conclusions of our research beyond the levels included in our experiment. Subjects are the prototypical random factor because one usually tests only a random selection of all possible subjects and yet we claim that our findings will hold for the population at large.[1]

Coleman (1964) and Clark (1973) have convincingly argued that in psycholinguistic research one should not only treat subjects as a random factor, but also the items (words). Their reasoning is simple and compelling: Only a (random) sample of the words of the language is used in our experiments, but our claims are supposed to hold for the language as a whole, or for all words of the same type as the words in the experiment (eg. all monomorphemic nouns).

Statistical analyses with two random factors were hard to run in 1973. For this reason, the field has converged on the use of two tests for every factor at hand. One test, F_1, is the classical test with subjects as the only random factor. This test thus ignores any effect of the randomness of items. The second test, F_2, treats items as random but subjects as fixed. This test makes a

[1]Textbooks are not always clear on whether a random factor needs to have all three mentioned aspects (sample of existing levels, randomness, extrapolation of claims), or whether having only one suffices, see Hays (1994) and Rietveld and van Hout (1993) for discussion.

simplifying assumption that is precisely opposite to the one underlying the F_1.

The net result is that psycholinguistic experiments are evaluated by the statistical counterparts of the lame and the blind: F_1 was denounced by Clark because it ignores the random factor items. F_2 is used as its support but it ignores the principal random factor subjects. A test that accounts for the randomness of both factors is what is needed. F' does precisely this.

Random factors and the model

The assumption that items are random changes the model that underlies the ANOVA quite drastically. Assume a very simple experiment with one treatment factor, items, and subjects. In the (hypothetical) event that we recognise no random factors the equation describing each observation would be:

(B.1) $$X_{ijk} = \mu + \alpha_i + \varepsilon_{ijk}$$

This states that every observation is equal to the grand mean μ, plus a factor for the level of the factor α_i, plus a noise factor that is specific to this level–subject–item combination, ε_{ijk}.

If only subjects are random, the observation for level i, subject j, and item k is mathematically described by including a correction for subject variance, π_j:

(B.2) $$X_{ijk} = \mu + \alpha_i + \pi_j + \varepsilon_{ijk}$$

When items are treated as a random factor too, the equation changes to:

(B.3) $$X_{ijk} = \mu + \alpha_i + \pi_j + \rho_k + \varepsilon_{ijk}$$

An additional factor ρ_k is introduced to account for the item-specific deviation from the grand mean.

These equations describe the modelling of each individual case. The effect on a larger scale can be appreciated from inspection of the expected mean squares. Theoretically, one can divide all sources of variation in an experiment into independent factors, like subjects, treatments, and items. In practice, these effect cannot be independently isolated from the data because of the random factors. When only subjects are considered random, what is called the expected mean square for treatments, EMS(T), is not only influenced by the variation due to treatments, σ_T^2, but also by the variation due to treatments and subjects, σ_{TS}^2, and by error variation, σ_e^2. This gets even more complicated when Items is also treated as a random factor. This will be demonstrated by the analyses actually used in this thesis.

Expected mean squares for reported experiments

The experiments reported in this thesis contain two factors that were experimentally manipulated:[2] The first factor is Word Type, which can be either variable or constant. This is the principal treatment factor. Next there is Base, which is either of the three (morphological) bases that was used to create the words. Items are nested under the combination of Word Type and Base, because three new items were selected for each level of Word Type. The experiment is completely within-subjects, ie. each subject is presented with all items and all conditions.

The model with only subjects as a random factor is shown in the following table. To make the table more readable, a shorthand notation is used in which W stands for $p\sigma^2_W$, or a certain number p times the variance due to W. [3]

(B.4)

	effect	Variance components ($\sigma^2_{..}$)	df
Word Type	EMS(W)	W + WS + e	1
Base	EMS(B)	B + BS + e	2
	EMS(WB)	WB + WBS + e	2
Item	EMS(I)	I + IS + e	12
Subjects	EMS(S)	S + e	11
	EMS(WS)	WS + e	11
	EMS(BS)	BS + e	22
	EMS(WBS)	WBS + e	22
	EMS(IS)	IS + e	132

If we want to test the effect of Word Type, we see that the expected mean square for W depends on the variance of Word Type, the variance of Word Type by Subjects and the error variance. To compensate for those last two terms, the correct statistical test (under the assumption that Subjects is the only random factor) is:

(B.5)
$$F_1 = \frac{\text{EMS(W)}}{\text{EMS(WS)}} = \frac{\sigma^2_W + \sigma^2_{W\times S} + \sigma^2_e}{\sigma^2_{W\times S} + \sigma^2_e}$$

In the equation, the variance components above and below the line cancel each other out until only σ^2_W remains, which is the desired unbiased estimate of the variance due to Word Type.

Under the assumption that both subjects and items are random, the table of expected mean squares is cluttered up with extra variance components for the new random factor I and for the interaction of the two random factors IS:

---

[2] There were two additional factors, Repetition and Context, that were summarised over before the statistical analyses were done.

[3] The exact value of p is only of computational use and should not concern us here, it depends on the number of levels of W.

(B.6)

	effect	Variance components ($\sigma^2_{...}$)	df
Word Type	EMS(W)	W + WS + I + IS + e	1
Base	EMS(B)	B + BS + I + IS + e	2
	EMS(WB)	WB + WBS + I + IS + e	2
Item	EMS(I)	I + IS + e	12
Subjects	EMS(S)	S + IS + e	11
	EMS(WS)	WS + IS + e	11
	EMS(BS)	BS + IS + e	22
	EMS(WBS)	WBS + IS + e	22
	EMS(IS)	IS + e	132

The essence of Clark's argument is that under these assumptions for mean squares, the F_1 test is no longer correct:

(B.7) $$F_1 = \frac{\text{EMS(W)}}{\text{EMS(WS)}} = \frac{\sigma^2_W + \sigma^2_{W\times S} + \sigma^2_I + \sigma^2_{I\times S} + \sigma^2_e}{\sigma^2_{W\times S} + \sigma^2_{I\times S} + \sigma^2_e}$$

As can be seen, the F_1 now estimates the summed effect of the desired variance component σ^2_W and the spurious variance component σ^2_I. This makes the F_1 test positively biased: Any spurious variance will inflate the numerator of the fraction, which leads to higher F-values and lower probability levels.

A similar argument can be made for the F_2, which tests the expected mean square for Word Type against the expected mean square for items. When only items are treated as a random factor, the equation for this test is:

(B.8) $$F_2 = \frac{\text{EMS(W)}}{\text{EMS(I)}} = \frac{\sigma^2_W + \sigma^2_I + \sigma^2_e}{\sigma^2_I + \sigma^2_e}$$

Under the more realistic model with two random factors, this test becomes:

(B.9) $$F_2 = \frac{\text{EMS(W)}}{\text{EMS(I)}} = \frac{\sigma^2_W + \sigma^2_{W\times S} + \sigma^2_I + \sigma^2_{I\times S} + \sigma^2_e}{\sigma^2_I + \sigma^2_{I\times S} + \sigma^2_e}$$

Clearly, this test is biased by the spurious variance component $\sigma^2_{W\times S}$.

A principled approach: The F'

When one takes Clark's argument about Items being a random factor seriously, both the test for F_1 and for F_2 are biased. There is no easy solution for this, because there is not one expected mean square that might be used to compensate for all four additional variance components in the EMS(W). The solution is to use a quasi-F ratio, which combines four expected mean squares. For the factor W, the F' looks as follows:

(B.10) $$F'_W = \frac{\text{EMS(W)} + \text{EMS(IS)}}{\text{EMS(WS)} + \text{EMS(I)}}$$

$$= \frac{\sigma_W^2 + \sigma_{W\times S}^2 + \sigma_I^2 + \sigma_{I\times S}^2 + \sigma_e^2 \quad + \quad \sigma_{I\times S}^2 + \sigma_e^2}{\sigma_{W\times S}^2 + \sigma_{I\times S}^2 + \sigma_e^2 \quad + \quad \sigma_I^2 + \sigma_{I\times S}^2 + \sigma_e^2}$$

As can be seen, in this equation all variance components above and below the line cancel each other until only σ_W^2 is left. This means that the F' is an unbiased estimate of the variance due to W.

Similar equations can be constructed for the other factors of interest:

(B.11) $$F'_B = \frac{\text{EMS(B)} + \text{EMS(IS)}}{\text{EMS(BS)} + \text{EMS(I)}}$$

(B.12) $$F'_{W\times B} = \frac{\text{EMS(WB)} + \text{EMS(IS)}}{\text{EMS(WBS)} + \text{EMS(I)}}$$

The structure of the F' is always the same: To the factor of interest, add the interaction of both random factors. Divide this sum by the sum of the two factors that appear as denominators in the F_1 and F_2: The interaction of the factor at hand with subjects and the factor items.

With modern-day computer power, computing the F' is not complicated at all. The required variance estimates can be obtained with any decent statistical package. The computation of the F' itself requires an external program, because this test is, to my mind, not implemented in any of the major packages. Computing the correct degrees of freedom is also relatively complicated: The degrees of freedom of F' depend also on the mean squares used. This is so because all F-tables are constructed to evaluate the result of a division of only two factors. Despite this complexity, the computation of an F'-value requires no more than 25 lines of LISP code, that is, it is trivial to program.

Comparing F' to F$_1$ and F$_2$

The F' is slightly more conservative than the combination of F_1 and F_2 and this fact should worry researchers, because cherished experiments might suddenly not be significant any more. It should immediately be noted, though, that the differences between F' and the classical tests are small. Only in some boundary cases, do the tests differ. This is necessarily the case, because F' aims to remedy the fact that F_1 and F_2 suffer from spurious variance components in the numerator, which can inflate the F-value. It is not that F' is too conservative; F_1 and F_2 can be too liberal.

Data from an unpublished experiment demonstrate the point I want to make here: In this experiment, the interaction of Word Type and Base was significant by-subjects, $F_1(2, 22) = 5.16, MS_e = 39266, p = .014$, and not by-items, $F_2(2, 12) = 3.61, MS_e = 10870, p = .059$. In a traditional analysis, one might be tempted see this as an indication of a real effect, because the low probability level of the F_1 "compensates" for the marginal significance of the F_2. F' shows that the indications for a real effect are not as strong as one would believe, $F'(3, 27) = 2.50, MS_e = 39266, p = .084$.

Two practical considerations

There are two practical consideration that might lead one to not use the F', even though its use seems better justified that that of F_1 and F_2.

First, no systematic comparison of the values of F', F_1, and F_2 have been done. The alleged conservativeness of the test should be addressed in a systematic study, which compares the outcomes of the three F-tests and tries to trace back differences to spurious variances in F_1 and F_2.

Secondly, although trivially programmed, the F' can still not be computed with tools that are generally available. The same holds for the analysis of the estimated mean squares in an experiment. Although indispensable for sound statistical understanding of the analysis, these tables can currently be generated by hand (see Kirk, 1982, p. 390, for instructions) or by BMDP and by MANOVET. The latter is a stand-alone program for DOS computers, written by H. Quené of Utrecht University, The Netherlands. Both programs do not live up to modern-day standards of user-centered design and a better solution should be provided to make F' the statistic of choice among psycholinguists.

In all, the use of F' is preferable from a statistical viewpoint and it is more principled than the use of F_1 and F_2. Alas, the use of the latter is the received statistical wisdom and this might be hard to change. More research into the properties of the test and making the tools for computing F' generally available will help to promote its use.

Samenvatting

Algemeen

De morfologie is de tak van de linguïstiek die zich bezighoudt met de interne structuur van woorden. Veel woorden zijn opgebouwd uit morfemen, zoals bijvoorbeeld te zien is bij woorden als *boeg+spriet*, *wend+baar* en *zeil+de* (waarbij de '+' een morfeemgrens aanduidt). De morfologie probeert regels te formuleren waaraan de combinaties van morfemen moeten voldoen, zodat verklaard kan worden waarom **paard+baar* geen goed woord is en *paard+achtig* wel.

De psycholinguïstiek is de tak van de psychologie die zich bezig houdt met taal. Welke psychologische processen en representaties liggen ten grondslag aan ons taalgedrag? In dit proefschrift heb ik me bijna uitsluitend bezig gehouden met taalproductie, waarbij de meer specifieke vraag is hoe een spreker een idee, bijvoorbeeld "ik heb dorst", omzet in een serie klanken, zoals *mag ik misschien een glaasje droge witte wijn van u.*

Eén onderdeel van dat taalproductieproces is verantwoordelijk voor het produceren van morfologisch complexe woorden, zoals *zeilde*. De vraag die in dit proefschrift behandeld wordt is hoe dit onderdeel werkt. Zijn morfologisch complexe woorden in hun geheel opgeslagen, of worden ze iedere keer samengesteld uit hun respectievelijke delen? Is er een aanpak van dit probleem die geldt voor *boegspriet*, *wendbaar* en *zeilde*, of zijn er verschillen tussen deze typen woorden? En hoe zit het met onregelmatige werkwoorden, zoals *omslaan–omsloeg?*

Theoretisch deel

In het **eerste hoofdstuk** worden diverse linguïstische aspecten van de productie van morfologisch complexe woorden behandeld. Cruciaal voor dit proefschrift is het onderscheid tussen inflectie en derivatie. Onder inflectie valt onder andere de vervoeging van werkwoorden in vormen voor diverse personen en tijden, zoals in *ik zeil vandaag* versus *Marion en Dries zeilden gisteren*. Derivatie omvat het maken van nieuwe woorden of varianten van woorden, waarbij van het werkwoord *wenden* bijvoorbeeld *wendbaar* gemaakt wordt, of *paarsig* van *paars*. Een proces dat verwant is aan derivatie is samenstelling, waarbij twee bestaande woorden gecombineerd worden tot een nieuw woord, zoals in het voornoemde *boegspriet*. Samenstelling kan ook gebeuren op basis van complexe woorden, zoals in *zeepaardje.*

Het onderscheid tussen derivatie en inflectie is vrij duidelijk, maar blijkt niet scherp gedefiniëerd te kunnen worden. Dit is een probleem voor de linguïstiek en vooral voor de aanhangers van de *split morphology hypothesis.* Deze hypothese behelst dat *wenden* en *wendbaar* verschillende woorden zijn, terwijl *zeilen* en *zeilde* verschillende vormen van een en hetzelfde woord zijn. In linguïstische termen: Derivatie is een proces dat plaatsvindt in het lexicon, terwijl inflectie de lexicale vormen voorziet van verdere morfologische markering.

In het eerste hoofdstuk wordt een aantal andere problemen met de split morphology hypothesis behandeld. Deze komen erop neer dat inflectie soms een aantal eigenschappen heeft die het op derivatie doen lijken. Ik stel een alternatieve interpretatie voor waardoor deze problemen grotendeels verklaard kunnen worden. Daarna bespreek ik het Levelt/Roelofs-model voor taalproductie, dat het theoretische raamwerk vormt voor dit proefschrift. De linguïstische eigenschappen van het model worden vergeleken met wat er elders in de literatuur wordt verondersteld. De split morphology hypothesis blijkt een natuurlijke plaats te vinden in dit model.

In **Hoofdstuk 2** worden de psycholinguïstische aspecten van de productie van complexe woorden behandeld. Allereerst wordt er ingegaan op de gegevens uit versprekingen, die aangeven dat er inderdaad een rol is weggelegd voor morfologie en voor het verschil tussen inflectie en derivatie, getuige versprekingen als:

(1)
a. Fancy getting your nose remodeled → Fancy getting your model renosed. (Garrett, 1980)
b. Lots of moisture and protection → Lots of moisture and protecture (Shattuck-Hufnagel, 1979)
c. De eigenaar van de koffers *worden* verzocht...

In voorbeeld (a) zijn de stammen van de woorden verwisseld, terwijl het verledentijdssuffix *-ed* en het voorvoegsel *re-* op hun plaats bleven. Naast stammen kunnen ook de uitgangen van plaats veranderen. Behalve verwisselingen zijn er ook gevallen waarbij een uitgang eerder of later nog een keer gebruikt wordt, zie (b). Als laatste wordt soms een geheel verkeerde uitgang toegepast, zoals te zien is in (c).

Daarna worden diverse modellen van taalproductie behandeld. Een eerste grote klasse is die van modellen gebaseerd op analogie. In de literatuur zijn hiervan vooral de connectionistische modellen (Daugherty & Seidenberg, 1994; MacWhinney & Leinbach, 1991; Plunkett & Marchman, 1991; Rumelhart & McClelland, 1986) veel besproken, omdat ze de taaldata proberen te beschrijven zonder expliciete regels te formuleren. Binnen de connectionistische wereld is het probleem van de Engelse sterke verleden tijd van groot belang, hoewel dit vanuit een morfologisch oogpunt een probleem met een nogal beperkte reikwijdte is.

Een tweede klasse modellen bevat de *dual route* modellen van Frauenfelder and Schreuder (1992), Schreuder and Baayen (1995) en Caramazza et al. (1988). Deze modellen beschrijven het begrijpen van taal, maar de resultaten kunnen ook toegepast worden op de taalproductie. Het fundamentele inzicht in het werk van Schreuder, Frauenfelder en Baayen is dat hoogfrequente woorden zich anders kunnen gedragen dan laagfrequente. Voor *zeilde* nemen we aan dat de vorm herkend wordt via zijn onderdelen *zeil* en *-de*, maar of dat ook zo is voor een hoogfrequente vorm als *maakte*, is maar de vraag. De combinatie *maak* en *-te* wordt zo vaak gemaakt, dat het handiger zou zijn de vorm als geheel op te slaan.

Vervolgens bespreek ik de eigenschappen van het Levelt/Roelofs-model die vanuit morfologisch oogpunt van belang zijn. Dit model maakt onderscheid tussen een conceptueel niveau, een niveau van abstracte woorden (lemma's) en een niveau van woordvormen (morfemen). De lemma's bevatten alle talige informatie over het woord behalve de vorm. Het lemma van het Franse woord *fille* (meisje) bevat onder andere de woordsoort (zelfstandig naamwoord) en het geslacht (vrouwelijk). Deze laatste waarde wordt door middel van een *agreement* proces doorgegeven aan het bijvoegelijk naamwoord in *la fille heureuse* (het vrolijke meisje). Het lemma van *heureux* bevat een *diacritic* voor geslacht, waarnaar de waarde 'vrouwelijk' gekopiëerd kan worden. Volgens de theorie is alle inflectie de talige specificatie van waardes die zich in *diacritics* bevinden of van waardes die inherent zijn aan een bepaald woord.

Empirisch deel

Hoofdstuk 3 is een kort hoofdstuk dat de voornaamste predicties op een rij zet. Omdat inflectionele suffixen heel voorspelbaar optreden, cruciaal zijn voor de interpretatie, en bij ieder gebruik terplekke aan de stam moeten worden toegevoegd, voorspel ik dat er een *inflectioneel frame* bestaat dat dit proces coördineert en controleert. Daartoe bevat dit frame een aantal open posities of *slots*, een voor de stam van het woord en een voor iedere mogelijk inflectioneel affix. In kort bestek wordt uitgelegd waarom derivationele suffixen waarschijnlijk geen plaats in het frame zullen vinden en wat de regel is die bepaalt hoeveel slots (posities) het inflectionele frame bevat: Zelfstandige en bijvoegelijke naamwoorden dragen maximaal één inflectioneel affix met zich mee, zoals te zien is als je *een aardig woord* vergelijkt met *de aardig+e woord+en*. Deze woordsoorten hebben dus een frame met een slot voor de stam en een slot voor één mogelijk affix. Werkwoorden kunnen twee inflectionele affixen bij zich hebben, zoals te zien is in *wij zeil+de+n*. Samen met een slot voor de stam brengt dat het aantal slots voor werkwoorden op drie.

In **Hoofdstuk 4** worden de eerste zes experimenten gepresenteerd. De belangrijkste evidentie voor inflectionele frames kan worden gevonden in **Experiment 1** en een aantal vervolgexperimenten daarop. In Experiment 1 worden samengestelde woorden *(werkbij)* vergeleken met verledentijdsvormen *(werkte)* in een implicit priming experiment. In dit soort experimenten worden woorden in kleine setjes tegelijk aangeboden en daarna een voor een door de proefpersoon uitgesproken. Hier zijn twee voorbeelden van dergelijke setjes:

(2)	a.	werkbij	b.	werkte
		werkgroep		werkdruk
		werkplaats		werkboek

Als alle woorden in een setje met hetzelfde morfeem (hier *werk*) beginnen, kunnen proefpersonen de woorden bijzonder snel uitspreken. Hoeveel sneller proefpersonen zijn ten opzichte van een normale conditie (waarin de woorden geen overlap vertonen) wordt uitgedrukt door het *preparatie-effect.* Voor het setje als dat in (2a) is het preparatie-effect gemiddeld 122 ms. Als echter een van de samenstellingen vervangen wordt door een werkwoord *(werkte)*, zoals in (2b) gebeurd is, zakt het preparatie-effect naar 36 ms.

Er is een aantal mogelijke verklaringen voor dit effect, bijvoorbeeld in termen van woordsoort of klinkerkwaliteit. Deze worden weerlegd in de volgende twee experimenten, **Experiment 2 en 3**, waarin bijvoegelijke naamwoorden *(werkbaar)* en verkleinwoordjes *(werkje)* vergeleken worden met samenstellingen. Bij deze woorden vinden we geen reductie van het preparatie-effect, zoals dat gevonden werd in Experiment 1. Het enige verschil tussen *werkte* en *werkbij* dat nog in aanmerking komt om de reductie van het preparatie-effect tot 36 ms te verklaren ligt dan in de inflectioneleframes: Zoals gezegd hebben zelfstandige naamwoorden een frame met twee slots en werkwoorden een frame met drie slots en dit maakt voorbereiding van het inflectionele frame mogelijk in conditie (2a) maar niet in (2b). Het niet van te voren beschikbaar zijn van het inflectionele frame vertraagt de productie zodanig dat een deel van het effect van beginoverlap teniet wordt gedaan.

Experiment 4 testte de generaliseerbaarheid van de bevindingen van Experiment 1. In dit experiment werden woorden gebruikt die niet het hele eerste morfeem gemeenschappelijk hadden (*werkte* en *werkdruk* in Experiment 1), maar slechts de eerste paar klanken: *heette* en *hemel.* De preparatie-effecten waren over het algemeen kleiner in dit experiment, maar de aanwezigheid van een werkwoord veroorzaakte nog steeds een significante reductie van het preparatie-effect.

In **Experiment 5** werd gekeken of dezelfde resultaten ook gevonden kunnen worden met woorden in het meervoud: *heetten* en *hemels.* Nu moeten alle woorden in het experiment voorzien worden van een affix en niet alleen de werkwoordsvormen. Dit bleek echter niet van invloed op de resultaten te zijn, de conditie met het werkwoord *heetten* erin vertoonde wederom een veel kleiner preparatie-effect dan de conditie met enkel zelfstandige naamwoorden.

In het laatste experiment van dit hoofdstuk werden sterke werkwoorden onderzocht. De verleden tijd van *krijgen* is *kreeg*, waarbij er dus geen sprake is van een inflectioneel affix. Feitelijk is er maar één inflectioneel affix mogelijk bij deze verleden tijdsvorm *(kreeg+en)* en dit pleit voor een inflectioneel frame met twee slots, dat dus gelijk is aan het frame voor zelfstandige naamwoorden. Inderdaad werd er in Experiment 6 geen reductie van het preparatie-effect gevonden voor de conditie met het sterke werkwoord. Ik doe een voorstel voor de representatie van onregelmatige vormen in het lexicon, dat deze resultaten kan verklaren.

Hoofdstuk 5 bevat een zestal experimenten die de resultaten uit het vorige hoofdstuk repliceren of verder uitbreiden. **Experiment 7** is een herhaling van Experiment 4 met grotendeels ander materiaal. Dit beïnvloedt het resultaat gelukkig niet, wederom vinden we een reductie van het preparatie-effect door het afwijkende inflectionele frame van het werkwoord. In **Experiment 8** worden de materialen van Experiment 4 en 7 nogmaals gebruikt, maar nu worden de werkwoorden vervangen door gelijkluidende zelfstandige naamwoorden: Waar eerder de verleden tijd *waadde* werd gebruikt, wordt die nu vervangen door het zelfstandig naamwoord *wade*. De andere woorden in het experiment blijven gelijk en het resultaat is dat de reductie van het preparatie-effect verdwijnt en hetzelfde resultaat behaald wordt in beide condities van het experiment.

In alle experimenten tot nu toe werd er gekeken naar beginoverlap tussen de woorden in de set. In **Experiment 9** werden werkwoorden die overeenkwamen in inflectie (*nieste* en *blafte*) vergeleken met woorden die eindoverlap vertoonden (*ruiter* en *kater*). In overeenstemming met eerdere resultaten van Meyer (1990) en Roelofs (1996) werd geen effect gevonden voor eindoverlap, zelfs niet als dit aparte morfemen waren.

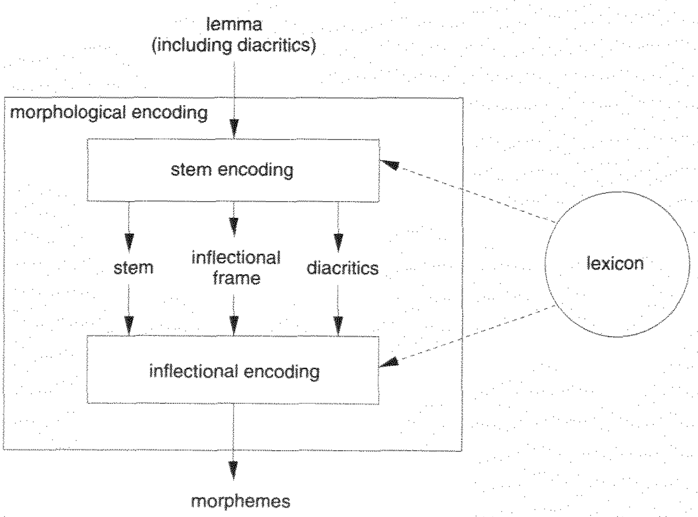
Het experiment 6 in het vorige hoofdstuk onderzocht het gedrag van sterk werkwoorden als *krijg–kreeg*. In dat experiment werden de verledentijdsvormen gebruikt, in **Experiment 10** wordt dit experiment herhaald in de tegenwoordige tijd. Hoewel er op zich niets onregelmatigs is aan de tegenwoordige tijd van *krijgen*, vinden we toch dat de tegenwoordige tijd zich net zo gedraagt als de onregelmatige verleden tijd: Beide vormen vereisen een voor werkwoorden onregelmatig inflectioneel frame, met slechts twee slots.

Er is tot nu aangenomen dat alle zelfstandige naamwoorden een zelfde inflectioneel frame hadden, ongeacht of het om enkelvoudige naamwoorden ging *(winkel)* of om samenstellingen *(windvlaag)*. In **Experiment 11** wordt deze assumptie getoetst door setjes van enkelvoudige naamwoorden te vergelijken met setjes van enkelvoudige naamwoorden en een samenstelling. Er blijkt geen verschil tussen deze twee condities te zijn en dit bevestigt de aanname dat enkelvoudige naamwoorden en samenstellingen geheel gelijk zijn voor wat betreft hun inflectionele structuur.

Het laatste experiment **(Experiment 12)** betreft de definitie van het aantal slots in het inflectionele frame: Is het inderdaad zo dat het aantal mogelijke inflectionele affixen het frame bepaalt, of is het aantal feitelijk aanwezige affixen van belang? Om dit te testen werden enkelvoudige verledentijdsvormen *(heet+te)* vergeleken met zelfstandige naamwoorden in het meervoud *(hemel+s)*. Beide typen woorden hebben één feitelijk affix en alleen de werkwoorden kunnen nog een tweede affix dragen (het meervoudsaffix, zoals in *heet+te+n*). Wederom vonden we een sterke reductie van het preparatie-effect in de conditie met het werkwoord, hetgeen erop duidt dat er inderdaad een verschil in inflectionele frames is tussen de gebruikte werkwoorden en zelfstandige naamwoorden. Gezien het gelijke aantal werkelijke suffixen, moet het aantal mogelijke suffixen een cruciale rol spelen.

Afsluitend deel

De belangrijkste conclusies uit de experimenten kunnen het beste worden samengevat aan de hand van Figuur 1. Allereerst wordt het morfologisch encoderingsproces opgesplitst in een deel dat *stem encoding* wordt genoemd en dat derivaties en samenstellingen voor zijn rekening neemt, en in een deel dat *inflectional encoding* wordt genoemd en dat de inflecties toevoegt. Dit is in overeenstemming met de split morphology hypothesis: Stam encoderen behelst namelijk voornamelijk het opzoeken van vormen in het lexicon en inflectioneel encoderen is het toevoegen van affixen aan de opgezochte vorm.



Figuur 1: Onderverdeling van het morfologisch encoderingsproces

De inflectionele frames zijn alleen gevoelig voor het aantal affixen dat in die laatste stap aan de stam wordt toegevoegd en verschaffen ons een methode om dit proces verder te onderzoeken. Uit een van de vervolgexperimenten bleek dat sterke werkwoorden (bv. *krijgen*) een afwijkend inflectioneel frame hebben, omdat voor deze werkwoorden het tijdsaffix niet door inflectie tot stand komt, maar er een speciale vorm uit het lexicon gehaald wordt *(kreeg)*.

In **Hoofdstuk 6** worden de mogelijke consequenties van de gevonden resultaten bekeken. Allereerst vereisen de voorgestelde inflectionele frames dat affixen die niet gerealiseerd worden toch een soort resultaat achterlaten. Concreet betekent dit dat een enkelvoud als *anker* geproduceerd wordt als *anker+ø* , waarbij de '*ø*' aanduidt dat er op die plek een enkelvoudsaffix staat, dat onzichtbaar is in het Nederlands. De problemen met dit soort 'lege' morfemen worden besproken en ik stel beperkingen op het aannemen van dit soort morfemen voor.

Verder lijkt het erop dat er in versprekingen niet alleen echte morfemen optreden, maar ook *semi-morphs*. Dit zijn woordonderdelen die op morfemen lijken, maar volgens de traditionele analyse geen morfemen zijn omdat ze in geen enkele context een vaste betekenis hebben. Betoogd wordt dat de morfologie betekenis niet als leidraad moet nemen, maar zich moet concentreren op wederkerende vormelementen. Verder moeten derivatie en inflectie kunnen putten uit hetzelfde morfeemrepertoire, zodat verklaard kan worden hoe een vorm soms twee verschillende toepassingen kan vinden, waarbij de een derivationeel en de ander inflectioneel is.

Een belangrijk stuk van Hoofdstuk 6 is gewijd aan de relatie tussen taal en denken en de rol van inflecties hierin. In het Levelt/Roelofs-model wordt aangenomen dat lexicale concepten een op een overeenkomen met stammen in de taal en er voor het woord *zeil* dus een concept ZEIL is. Ik stel voor dat de precieze afbakening van concepten weliswaar verschilt tussen culturen (en dus ook tussen talen), maar dat de belangrijkste verschillen tussen talen liggen in welke verdere informatie er nog bij de concepten verschaft moet worden. Voor het Nederlands is dat bijvoorbeeld tijdsinformatie, dat op het werkwoord uitgedrukt moet worden. Maar voor andere talen, zoals het Chinees, is zulke tijdsinformatie van geen belang.

Drie laatste korte punten in dit hoofdstuk betreffen de status van de derivationele regels in het voorgestelde psycholinguïstisch model, hoe het inflectioneel frame toegepast kan worden op talen die een veelvoud aan inflecties hebben (zoals het Turks), en hoe het probleem van het definiëren van inflectie en derivatie misschien ooit opgelost kan worden door experimentatie.

Met dit laatste punt is de cirkel rond: Hoewel inflectie en derivatie niet duidelijk gedefiniëerd kunnen worden, bleken experimenten naar deze verschijnselen bijzonder vruchtbaar te zijn. Het inflectionele frame speelt een belangrijke rol in de productie van geïnflecteerde woorden en diverse eigenschappen van dit productieproces zijn verduidelijkt, waaronder de behandeling van onregelmatige werkwoorden. Uiteindelijk kunnen deze en soortgelijke experimentele gegevens leiden tot een beter begrip van de taalkundige aspecten van inflectie en derivatie, waaronder de zo problematische definitie van het verschil tussen beide processen.

Curriculum Vitae

I was born in Tilburg, The Netherlands, on Sunday 19 October 1969. Still young, we moved to a little village, Riel, where I learned judo, horseback riding, and hockey: The necessities of life. Despite my earlier preference for math and computers (my secondary education resulted in a *gymnasium* β diploma), I went to Nijmegen to study Linguistics. After some short excursions to computer science and philosophy, my remaining interests found an outlet in a second study, Cognitive Science. Both studies were completed in 1994.

During and after my studies, I was lucky enough to work with William Marslen-Wilson (then in London) and with Anne Cutler and James McQueen (MPI), who taught me the tricks of the trade. In January 1996, I became a PhD student with Ardi Roelofs and Pim Levelt. The work reported in this thesis (and some more) was carried out during the three years of my stipend. Thanks to the MPI, I was able to visit many conferences and courses, most notably the ESCOP summerschool in Bressanone in 1997.

Next to my studies and research, I have always loved to get involved in writing and editing (*Ivoren Toren* and *De MaxKrant*), computer programming, organising things (the MPI series in psycholinguistics), and a bit of teaching. I hope to continue all of the above in the year 2000, which will mostly be spent in Tucson, Arizona and Urbana-Champaign, Illinois on a post-doc grant from the Niels Stensen Stichting.