On the production of morphologically complex words with special attention to effects of frequency

ISBN: 978-90-76203-30-0

Printed and bound by Ponsen & Looijen bv, Wageningen

On the production of morphologically complex words with special attention to effects of frequency

Een wetenschappelijke proeve

op het gebied van de Sociale Wetenschappen

Proefschrift

ter verkrijging van de graad van doctor

aan de Radboud Universiteit Nijmegen

op gezag van de rector magnificus prof. Mr. S.C.J.J. Kortmann,

volgens besluit van het College van Decanen

in het openbaar te verdedigen op dinsdag 27 november 2007

om 15.30 precies

door

Heidrun Bien

geboren op 8 juni 1976

te Münster, Duitsland

Promotor: Prof. dr. W.J.M. Levelt

Copromotor: Prof. dr. R. H. Baayen

Manuscriptcommissie:
Prof. dr. H. Schriefers (NICI, RU Nijmegen)
Dr. D. Davidson (FCDC, RU Nijmegen)
Prof. dr. G. Libben (University of Alberta, Canada)

CONTENTS

ACKNOWLEDGEMENTS

Finally, the work is done and I can express my gratitude to the people who stood beside me throughout my Ph.D. project.

First of all, I would like to thank my promoter Pim Levelt for taking me on as his last Ph.D. student in the Utterance Encoding Group. Pim, you were a great model of precision, persistence and focus and I am still astonished how you managed to give speedy feedbacks from all over the world. Your determination was essential to the success of this project and you taught me what it takes to become a good researcher.

Next, I want to thank Harald Baayen for being my co-promoter. Harald, you introduced me to the world of UNIX and R. Most of all, I am grateful for your contagious enthusiasm and for more than one unschedulded six-hour-meeting.

Here, I also have the chance to thank Pienie Zwitserlood for having employed me as a student assistant in 1999, for having supervised my master thesis and for having encouraged me start a Ph.D. You opened the world of psycholinguistics to me and remained a stable factor over the years.

Thanks to numerous people for their participation in one of my experiments (and even more so, for showing up at the second session, after having gone through the first).

Working at the institute would not be possible without the support of the technical group, the administration, the secretaries, the library, and the reception. Thanks to all of you for taking care.

A major advantage of being a Ph.D student at this institute is to have great colleagues around at all times, even at the most unusual ones such 3 a.m. At the more usual times, our lively Pim-broodje lunches were nutritious to body and soul, increasingly so as the Utterance Encoding Group gradually dissolved and Evelyn and I were in danger of developing signs of social phobia. Thanks to Petra, Kerstin,

Keren, and Martijn, to Mirijam, Frank, Anita, Claudia, and Lara, to Rebecca, Annelie, and Sarah, Carmel, Laura, and Bettina and to all the others who have been there over the years. You have made this time special.

As a healthy mind comes in a healthy body, one contributor to this thesis was the sport in and outside the MPI. Ad's volleyball group has been a constant companion and a lot of fun. I am specialy grateful to John and Kees for having me in their 'bedrijvencompetitie' team. To me, it was the perfect combination of being challenged, seeing myself improve, and winning. During a Ph.D. project, such feelings are worth more than gold. And I can not thank Annelie enough for having walked the Vierdaagse with me. It was a painful but wonderful experience. See, it is true: if you can finish the 4daagse, you can finish your thesis.

Mette, 'Coos & neighbours', you have made me feel home in Bottendaal. To you and to my friends at the KeizerKarel volleyball I want to say thanks for comfort, patience, distraction and the feeling of belonging.

Berna, thanks for bearing with me in these unbearable times.

I am honoured to have Annelie Tuinman and Mette Pyckevet as my paranimfs.

INTRODUCTION

CHAPTER 1

1.1. Productivity in human speech

Language is often referred to as the one truly unique human capacity. Different species have evolved a variety of communication systems, which serve a variety of communicative needs, such as warning, attracting the attention of potential mates, claiming one's territory. When compared to human languages, animal communicative systems differ in several crucial respects. One apparent difference is that they contain substantially fewer units, most of which are either iconic (directly representative of what they express) or indexical (drawing attention directly to what they express), but rarely symbolic. A symbolic unit represents its referent and can be used to give information about that referent, regardless of whether the latter is physically present or remote in space or time. Human language units are almost always symbolic, providing unlimited potential for communication. Another difference is that units in animal communicative systems do not typically combine with other units to form new meaningful entities, while human languages are very productive. A finite number of basic elements (words or morphemes) can be combined to form an unlimited number of expressions to communicate ideas.

Remarkably, all that people are usually aware of when speaking their mother tongue is the thought that they wish to express and the sentence they hear themselves articulate almost immediately afterwards. What happens in between remains largely subconscious. When we want to learn more about the phenomenon of human language production, we need to take a closer look at the structures and processes involved. In this doctoral dissertation, I focus on speech production at the level of words.

Words are the central units of language. It is the word that carries meaning. Words consist of phonemes (which do not themselves contain any meaning), and can be combined to form sentences (which turn out to be more or less meaningful). Adult speakers in our society know an estimated 50,000 to 100,000 words in their mother tongue. Considering words such as *rainbow*, *rainy*, *rained*, and *snowball*, *snowy*, *snowed*, it is evident that some words seem to be related to one-another by rules. Indeed, this often involves structural units smaller than the word: the morpheme.

Some morphemes can occur as independent words (free morphemes), whereas others cannot (bound morphemes). Words that contain only one morpheme (e.g., *dog*) are referred to as monomorphemic or morphologically simplex words. Words that contain two or more morphemes are referred to as morphologically complex words. Any given morphologically complex word belongs to one of the following three categories: a) compounds, b) derived words, or c) inflected words. Looking again at our examples (*rainbow, rainy, rained, snowball, snowy, snowed*), there are representatives of all three categories. Simply expressed, compounds are a combination of two (or more) independent words (e.g., *rain+bow*; *snow+ball*). In derivation, a free morpheme (stem) is combined with (at least) one bound morpheme (affix) to construct a new word (e.g., *rain+y*, *snow+y*). This derived word is related to its stem, but may belong to a different word class. In inflections, a free morpheme (stem) is combined with (at least) one bound morpheme (affix) to construct a new form of the same word (e.g., *rain+ed*, *snow+ed*), without changing word class.

Languages differ in their morphological complexity. Dutch is of higher morphological complexity than English, but still relatively low on the continuum of morphological complexity. Agglutinating languages (such as Turkish, Finnish, and Swahili) tend to have a high rate of morphemes per word 'glued together'. Some agglutinating languages, such as Turkish, are highly regular. Generally, morphological regularity refers to the fact that, when a given affix is glued to stems, all stems will undergo the same transformation in meaning. In other words, the morphologically complex words formed by the same affixation will all have the same relationship to their stems. Therefore, it is even possible to attach an affix to newly created words or words that are adopted from other languages, and to efficiently communicate the intended change in meaning. Any proficient listener of Dutch will understand the word *downloadbaar* as *can be downloaded* (even, if the meaning of *to download* is unknown). Likewise, speakers of Dutch can inflect a newly created verb or a verb that is adopted from another language and use it appropriately in the syntactical context of Dutch (e.g., '*ge-sms-t*'). Thus, morphological rules define possible new complex words in a language (Aronoff, 1994) and also condition how a word fits into a sentence. While the source of some of the morphological variation is the grammar of word combination, the focus of my study is on the production of isolated, morphologically complex words.

In the case of compounding, there are rules about the combination of base morphemes to define the meaning of the new word. In Dutch (as in English), regularly, the rightmost morpheme is the head of the compound. A head determines the semantic class of a compound, its word class, case, and gender, etc. A modifier specifies that semantic class, while the type of specification varies between compounds ('is made of', 'is used as', 'is shaped as', 'is found in', etc.). Morphological rules enable speakers to form and understand new morphologically complex words, which increases the productivity of their language.

1.2. Morphologically complex words

Based on the fact that morphologically complex words have internal structure, the following question arises: Are morphologically complex words stored in the mental lexicon as complex words? In some cases this is likely to be the case at some level. For example, not all compounds are transparent in a sense that their meaning can be directly inferred from the meaning of their constituents (opaque compound such as *honeymoon*). In fully transparent compounds or at levels beneath conceptualization, storage of morphemes could be sufficient as complex words can be assembled from their constituting morphemes. The extra processing load of building complex words during production might be made up for by the advantage of a smaller store of memorized forms to search through.

Full-storage models of speech production (e.g., Butterworth, 1983) assume that all words (morphologically simplex or complex) are stored in the same way, as separate nodes at the word form level. In contrast, no separate stored form representations for complex words are assumed in the production model of Levelt, Roelofs and Meyer (1999). For all words, simplex or complex, transparent or opaque, activation spreads from concepts to lexical representations (lemmas). A lemma contains all syntactic information about the word but no phonological information. This process, following conceptualization and resulting in the selection of one lemma, is part of grammatical encoding. After the selection of the lemma, phonological encoding starts with access to the word form level. According to Levelt et al. (1999), the word form level contains morphemes, and words that share a morpheme (*snow*, *snowflake*, *snowy,* etc.) all involve access to the same form node for the shared morpheme (*snow*).

The idea of a lexical level, containing individual representations (lemmas), and a form level, containing shared, morphemic nodes, becomes most clear in the case of homophones. Homophones are words that differ in meaning but sound the same (e.g., *bank*). Levelt et al. (1999) assume homophones to have their own lemmas but to access the same word form (Jescheniak & Levelt, 1994; Jescheniak, Meyer, & Levelt, 2003; but see Caramazza & Miozzo, 1998; Carmazza, Costa, Miozzo, & Bi, 2001; Shatzman & Schiller, 2004). Similarly, opaque compounds and derived words that are not fully decompositional are assumed to have their own representations at higher levels, but to access nodes at the word form level in no other way than transparent complex words do. Indeed, there is converging evidence from different paradigms that the constituents of complex words are involved in speech production irrespective of whether the complex word is semantically transparent or not (Zwitserlood, Bölte, & Dohmes, 2000; Roelofs & Baayen, 2002; Melinger, 2003; Dohmes, Zwitserlood, & Bölte, 2004; Gumnior, Bölte, & Zwitserlood, 2006).

Roelofs (1996) addressed the question whether the form lexicon underlying speech production contains morphologically decomposed entries (decomposition hypothesis vs. full listing hypothesis). If compounds such as *schuimbad* (foam bath) and *schoolbel* (school bell) are stored in a decomposed fashion, individual nodes for *schuim*, *bad*, *school*, and *bel* will be accessed. To test this hypothesis, Roelofs made use of a reliable effect in speech production: the frequency effect. In a seminal study, Oldfield and Wingfield (1965) had found that pictures with low-frequency names (e.g., *syringe*) took longer to name than pictures with high-frequency names (e.g., *basket*). Many studies thereafter have found effects of word frequency on speech onset latencies to be replicable and robust (Jescheniak et al., 1994; Levelt, Praamstra, Meyer, Helenius, & Salmelin, 1998). Jescheniak et al. (1994) located the word frequency effect at the word form level. Using the lemma based task of gender decision a frequency effect was found, but one that diminished over repetitions. In contrast, a frequency effect in word form access was shown to be robust over repetitions. Assuming, first, that the frequency effect is located at the word form level, and, second, that compounds are stored in a decomposed fashion, individual nodes for the constituents should be accessed with their individual frequencies. Roelofs (1996) used the implicit priming paradigm (Meyer, 1990, 1991), in which subjects produce words from learned paired-associates. The compounds were presented in either homogeneous or heterogeneous response sets, with homogeneity referring to an overlap of the first morpheme (e.g., *schuimbad*, *schuimkop*, *schuimspaan*). Stimulus presentation in the homogenous response sets resulted in shorter onset naming latencies than presentation in heterogeneous response sets, due to a preparation effect for the modifier. Crucially, this preparation effect was larger for low-frequency modifiers (e.g., *schuim*) than for high-frequency modifiers (e.g., *school*) suggesting that it is the constituents rather than the entire compound word that is accessed at the word form level.

In this dissertation, I study the production of morphologically complex words in Dutch. Based on the assumption that frequency effects arise at the lexical level, the retrieval of wordforms is informative with respect to the storage of wordforms. I focus on the production of one subtype of each of the three types of morphologically complex words: compounds (noun-noun compounds such as *handtas* (handbag)) in Chapter 2, derived words (deverbal adjectives such as *drinkbaar* (drinkable)) in Chapter 3, and inflected words (inflected, regular verbs such as *draaiend* (spinning)) in Chapter 4.

1.3. Predictors

A chronometric paradigm measures the speech onset latencies of the participants. These reaction times are the dependent variables for the statistical analyses. Among the independent variables, I distinguish between two kinds of predictors: control variables and critical variables.

I measure the speech onset latencies to learn more about the effects of variables that are of particular interest (the critical variables). When measuring speech onset latencies, however, one needs to be aware of the fact that a number of other variables, one is not particularly interested in, might also influence the latencies. I will refer to these variables as control variables, as it is my aim to bring these unwanted sources of variation under experimental control. One such variable is REPETITION. In my experiments, I measure ten latencies per word per participant in sequence. The number of repetitions might influence the latency as some participants could become faster (due to practice) or slower (due to fatigue) within the sequence. Another control variable relates to acoustic characteristics of a word's initial phoneme. As the onset latencies are recorded via microphone, some prevoicing might not be detected by the voice key (Kessler, Treiman, & Mullennix, 2002). This potentially leads to longer onset latencies for words which start with a plosive. Therefore, a first action of control is taken at the stage of material selection. I try to reach a fair distribution of initial phonemes and characteristics over conditions and over the distribution of other variables. In addition, I include the variables PLOSIVE (plosive, fricative, or other initial phoneme) and VOICED (voiced or unvoiced initial phoneme) in the statistical analyses.

Of theoretical interest are the following three groups of variables: a) Frequency Variables (1.3.1), b) Morphological Variables (1.3.2.), and c) Phonological Variables (1.3.3.). The first set of predictors brings together the variables SURFACE FREQUENCY OF THE COMPLEX WORD, CUMULATIVE STEM FREQUENCY, LEMMA FREQUENCIES, and POSITIONAL FREQUENCY. The morphological group includes the variables POSITIONAL ENTROPY, DERIVATIONAL ENTROPY, and INFLECTIONAL ENTROPY. The phonological group contains the variables PHONOLOGICAL WORD LENGTH, NEIGHBORHOOD DENSITY, POSITION-SPECIFIC NEIGHBORHOOD DENSITIES, and COHORT ENTROPIES. In what follows, we will take a closer look at these variables, the motivation to study them and the expectations concerning their effects on the production onset latencies of morphologically complex words.

1.3.1. Frequency Variables

Surface Frequency of the Complex Word

A first frequency measure to consider in the analysis of the speech onset latencies of morphologically complex words is the frequency of occurrence of the complex word as it is actually produced. I will refer to this frequency as the SURFACE FREQUENCY OF THE COMPLEX WORD. According to fully non-decompositional models of speech production (e.g., Butterworth, 1983), morphologically complex words are stored in the same way as morphologically simplex words, with individual nodes at the word form level. If the production of a complex word involves the activation of a separate word form node for the complex word, the production latency of a complex word should relate to its own frequency of occurrence.

Lemma Frequency

Both for morphologically complex words as well as for compound constituents, I further include the LEMMA FREQUENCY measure (e.g., Jescheniak et al., 1994), defined as the summed frequencies of a word's inflectional variants (i.e. the summed frequencies of the inflectional paradigm). In the case of noun-noun compounds, their constituents, and deverbal adjectives, it is the frequency of the singular form plus the frequency of the plural form (e.g., *handbag*, *handbags*). In the case of inflected verbs, it is the sum of the frequencies of all inflections (e.g., *dream*, *dreaming*, *dreamed*).

Cumulative Stem Frequency

In contrast to fully non-decompositional models, fully decompositional models of speech production assume that all morphologically complex words are assembled from their constituting morphemes at the form level. Decompositional models, therefore, predict frequency effects for the constituents rather than the full word. Based on the assumption that the production of any word containing a given morpheme (be it compounded, derived, inflected or morphologically simplex) involves access to the very same morpheme node, the relevant frequency for fully-decompositional models is the sum of all occurrences of the given morpheme. I will refer to this sum of frequencies of all contextual variants of a morpheme as the CUMULATIVE STEM FREQUENCY (e.g., Laudanna & Burani, 1985; Burani & Caramazza, 1987; Schreuder & Baayen, 1997).

While the prediction of a frequency effect for the initial morpheme follows straightforwardly for decompositional models of speech production, the prediction of a frequency effect for later morphemes depends on the notion of incrementality. Levelt et al. (1999) assume that the morphemes of a morphologically complex word are retrieved one after the other from beginning to end. The consecutive process of phonological encoding can start as soon as the initial morpheme of the compound has been retrieved. Assuming both decompositionality and incrementality, Levelt et al. (1999), therefore, predict frequency effects for initial morphemes only.

Positional Frequencies

An intermediate position between fully decompositional models and fully non-decompositional models of speech production is the assumption of some form of structured storage, in which morphemes might be stored as separate entities with information about their composability. In such a model, the frequency in which a morpheme (e.g., *hand*) occurs as a modifier in a compound (e.g., *handbag, handcuff, etc.*) might be considered a better predictor of the production latency of one of these compounds (e.g., *handbag*) than the frequency of occurrence of the morpheme (e.g., *hand*) as an independent word or as a constituent in any morphologically complex word.

Next to the cumulative stem frequency and the constituent lemma frequency (both of which are position-independent), I, therefore, compute the POSITIONAL FREQUENCY, defined as the sum of the frequencies of all members in the constituent family (Krott, Baayen, & Schreuder, 2001). For *hand* in *handbag*, it is the sum of the frequencies of all compounds that contain *hand* as modifier.

While the cumulative stem frequency is an unconditional frequency, both the lemma frequency and the positional frequency are conditional frequencies: the former is syntactically conditioned with respect to the word, the latter is structurally conditioned with respect to position.

1.3.2. Morphological Variables

As described above, words are also members of paradigms and the frequencies within the paradigms may influence the production latency of a word. Next to the sums of frequencies defined in the previous section, I include three measures applying Shannon's entropy (Shannon, 1948; Shannon & Weaver, 1949). The crucial difference between the sums of frequencies and the corresponding entropies over the distributions of these frequencies is that the entropies capture aspects of both type and token frequencies. As token-weighted counts of types (Moscoso del Prado Martín, Kostic, & Baayen, 2004; Baayen, Schreuder, & Feldman, 2006), the entropy is affected both by the size of the paradigm (i.e., the number of tokens) and their relative frequencies. All tokens being evenly frequent, the entropy is high. From the comprehender's point of view, the concept of entropy can be understood as a measure of uncertainty. From the speaker's point of view, a high entropy characterizes a paradigm full of equally likely forms.

Positional Entropies

Constituent families are the basis for calculating both the previously presented variable positional frequency and the variable POSITIONAL ENTROPY. Rather than summing up the frequencies of the constituent family members (as done to compute the positional frequency), we now compute Shannon's entropy (Shannon, 1948; Shannon & Weaver, 1949). Generally, high positional entropies indicate constituent families with many members, or constituent families with members that are of similar frequency, while low positional entropies translate to constituent families with only a few members or constituent families with a large variation in frequency. Therefore, the POSITIONAL ENTROPY of *hand* in *handbag* reflects the frequency distribution of all compounds that share *hand* as a modifier.

Derivational Entropy

Next to a position-specific measure of entropy, I look at the predictive value of an entropy measure that is position-independent. The DERIVATIONAL ENTROPY of *hand* in *handbag* reflects the frequency distribution of all morphologically complex words that share *hand* as a constituent, independent from its position within the word.

Inflectional Entropy

Words differ not only in the summed frequencies of their inflectional variants (the LEMMA FREQUENCY) but also in the distribution of these frequencies. While Dutch verbs have seven inflectional variants, compounds and derived words have only two (the singular and plural form). For verbs, I include the predictor INFLECTIONAL ENTROPY, defined as Shannon's entropy estimated by the relative frequencies of a verb's inflectional variants. A high inflectional entropy indicates that a given verb stem (e.g., *dromen* (to dream)) is actually used in many or all of its infected forms (*droom, droomt, dromen, dromend, droomde, droomden, gedroomd*), and that these inflections are of similar frequencies. Under these circumstances, the production of a specific inflected form might be relatively harder than when the forms in the inflectional paradigm are few and of different frequencies.

1.3.3. Phonological Variables

Phonological Word Length

Effects of word length have been shown to exist in production research (e.g., Meyer, Roelofs, Levelt, 2003). As speech unfolds over time, the time required for the mere articulation as well as perception of a word increases with the number of phonemes. Although the perception of a word is a prerequisite for its recognition, comprehension research has shown that (depending on the location of the uniqueness point (e.g., Marslen-Wilson, 1990) a word may be recognized even before its last phoneme has been perceived. Similar to the difference between perception and recognition, there might be a difference between articulation and the planning of articulation with respect to the number of phonemes. If a word is fully planned before the onset of articulation and if the time required for planning increases with the number of phonemes, PHONOLOGICAL WORD LENGTH might affect not only the mere articulation time, but also the speech onset latency. If PHONOLOGICAL WORD LENGTH explains a significant portion of the variance in the speech onset latencies, I, therefore, expect it to be inhibitory.

Neighborhood Density

Phonological Neighbors are words that can be transformed into one another by substituting a single phoneme (Greenberg & Jenkins, 1964; Coltheart, Davelaar, Jonasson, & Besner, 1977). Effects of the number of phonological neighbors of a word (a word's NEIGHBORHOOD DENSITY) have been encountered in both comprehension and production studies, however, typically inhibitory in the former and facilitatory in the latter (e.g., Luce & Pisoni, 1998; Vitevitch, 2002; but see Vitevitch & Stamer, 2006, for contrasting results in Spanish).

In studies on auditory word recognition, dense neighborhoods were found to slow processing and reduce identification accuracy. According to the Neighborhood Activation Model (Luce et al., 1998), effects of word frequency are directly tied to the number and nature of phonologically similar words activated by a stimulus input. The relative ease of word recognition depends on the word's neighborhood density and its frequency relative to the frequencies of its neighbors. A high-frequency word with many or high-frequency neighbors might be relatively harder to recognize than a low-frequency word with a few or low-frequency neighbors. Recently, facilitative effects of neighborhood density were also found in studies on speech production. Vitevitch (2002) found words with dense neighborhoods to be produced more quickly than words with sparse neighborhoods and explains his finding by a co-activation of phonologically similar words that increases the activation of the target word. Scarborough (2004) showed that vowel-to-vowel co-articulation is more likely in words with sparse phonological neighborhoods.

Neighborhood Density effects in speech production are crucial with respect to the notion of joint activation at the word form level. In the speech production model of Levelt et al. (1999), several concepts can be activated at the conceptual level, but only one lemma will be eventually selected and activate its word form. Therefore, beneath the level of lemma selection, the model predicts no interference by irrelevant, non-selected word forms. Other models of speech production (e.g., Dell, 1986) assume spreading activation and competition also at the word form level. Findings of neighborhood effects in the production of words may help to distinguish between these theories.

Given that long or morphologically complex words tend to have few or no neighbors, I only include the neighborhood densities for the constituent morphemes.

Position-Specific Neighborhood Densities

The idea to take into consideration the phonological neighbors of a word originally comes from studies that involved reading, presenting whole words at once. Relatively short complex words can be read based on a single fixation. However, as the speech signal unfolds over time, both in speaking and listening, words are processed from beginning to end. I, therefore, look at the influence of phonological neighbors in an additional, more specific and hopefully suitable way: I count the number of neighbors exchanging the first, second, third, etc. phonemes separately (Sevald & Dell, 1994, for initial neighbors). These POSITION-SPECIFIC NEIGHBORHOODS (N1, N2, N3, etc.) add up to the total number of phonological neighbors of a word.

As with the neighborhood densities presented above, I compute the position-specific neighborhood densities for the constituent morphemes but not for the morphologically complex words, because the latter tend to be zero. Likewise, I include the specific neighborhoods of only the first three positions (N1-N3), because N4 is often zero (some morphemes consist of only three phonemes).

Cohort Entropies

A second group of variables based on the notion of incremental speech processing are the COHORT ENTROPIES (H1, H2, H3, etc), defined as the likelihood of a word given all other words beginning with the same first (H1), first two (H2), first three (H3), and so on phonemes. Recent speech corpora studies on fine phonetic detail (Van Son & Pols, 2003; Van Son & Van Santen, 2005; Kuperman, Pluymakers, Ernestus, & Baayen, subm) found greater reduction of segments with small information loads in their production cohort compared to segments with great information loads, strongly suggesting influences of other word forms on the production of the intended word form. Cohort entropies and position-specific neighborhoods tend to be negatively correlated.

1.4. Statistical analyses

In this doctoral dissertation, I present analyses of variance (Chapter 2), as well as two different types of multiple regression analyses (Chapter 2-5). As presented in Chapter 2, the structure of noun-noun compounds allowed for a matching of frequencies and the application of analysis of variance designs. The additional application of a stepwise multilevel analysis of covariance with participant as main grouping factor (Pinheiro & Bates, 2000; Baayen, Tweedie, & Schreuder, 2002; Quené & Van den Bergh, 2004) on the same data provided valuable insight with respect to the predictive values of other variables of interest. In morphologically complex words that contain bound morphemes (such as the studied deverbal adjectives and inflected verbs), a similar matching of frequencies for analysis of variance designs was unfeasible (see Cutler, 1981 for a discussion of difficulties in matching of item sets in psycholinguistics). In Chapters 3, 4 (and 5), I, therefore, exclusively used multiple regression analyses.

1.4.1. Statistical Analysis of multiple regression

Very recently, a new technique for the analyses of data with repeated measurements has become available: the multilevel analysis of covariance with subject and word as crossed random effects (Bates, 2005; Bates & Sarkar, 2005; Baayen, in press). The superiority of mixed-effects models has convinced me to reanalyze all datasets in this fashion. In the following section, I elaborate why I chose for this newly available mixed-effects modeling with Subject and Word as crossed random effects by comparing different statistical analyses of multiple regression.

By-Item regression

Multiple regression with Subjects and Words can be done in several different ways. One possibility is to average over Subjects to obtain by-Word means, and to use these by-Word means as the dependent variable in a standard ordinary least squares regression with by-Word predictors such as FREQUENCY and LENGTH. I refer to this procedure as a by-Item regression (our Items are Words).

By-Subject regression

A second possibility is to fit separate regression models to the response latencies of the individual subjects. For an experiment with 20 subjects, this technique results in 20 regression models, with 20 estimates for each coefficient (FREQUENCY, LENGTH, etc.). Intercepts and slopes for the population are estimated by averaging over the 20 estimates. Significance is assessed by means of one-sample t-tests applied to the 20 estimates for the intercept and to the 20 estimates for each slope in the model. These t-tests evaluate whether a given coefficient is significantly different from zero, i.e., whether the associated predictor has a non-zero slope. This regression technique, to which I will refer as by-subject regression, was advocated by Lorch and Myers (1990). This regression technique has become the gold standard in psycholinguistic research.

Mixed-effects models

With the advent of mixed-effects modeling (Pinheiro et al., 2000; Baayen et al., 2002; Bates, 2005; Bates & Sarkar, 2005; Baayen, in press), new possibilities for regression analysis have become available. Mixed-effects models are models which incorporate two kinds of predictors, fixed effects and random effects. The distinction between fixed effects and random effects pertains to factorial predictors. Factors with a fixed, usually small number of levels, such as Sex (male versus female) or Animate (animate versus inanimate), are fixed effects. When the levels of a factor represent a sample from a larger population, we are dealing with a random effect. Examples of random effects are Subject (the subjects in an experiment usually do not exhaust the population of speakers) and Word (the words in an experiment usually represent an (ideally random) sample from the words in use in a given language community). Technically, random effects are modeled as random variables that follow a normal distribution with a mean equal to zero and some unknown standard deviation that is to be estimated from the data.

Let us look at a model with Subject as random effect

$$\mathbf{y}_i = \mathbf{X}_i\beta + \mathbf{Z}_i\mathbf{b}_i + \epsilon_i \quad (1)$$

For an experiment with four words, and with FREQUENCY and LENGTH as predictors, $\mathbf{X_i}$ represents the data matrix for subject i.

$$\mathbf{X}_i = \begin{pmatrix} 1 & F_1 & L_1 \\ 1 & F_2 & L_2 \\ 1 & F_3 & L_3 \\ 1 & F_4 & L_4 \end{pmatrix} \quad (2)$$

The transposed vector of population coefficients $\boldsymbol{\beta}^T = (\beta 0, \beta 1, \beta 2)$ specifies the intercept and the slopes for FREQUENCY and LENGTH. The product of $\mathbf{X}_i\boldsymbol{\beta}$ represents the expected values for an unseen subject.

$$\mathbf{X_i}\beta = \begin{pmatrix} \beta_0 + F_1\beta_1 + L_1\beta_2 \\ \beta_0 + F_2\beta_1 + L_2\beta_2 \\ \beta_0 + F_3\beta_1 + L_3\beta_2 \\ \beta_0 + F_4\beta_1 + L_4\beta_2 \end{pmatrix} \quad (3)$$

In a mixed-effects model, the intercept and slopes are calibrated to provide more accurate fits for the subjects in the experiment. Calibration for the intercept implies that the average speed with which a subject responds is allowed to vary from subject to subject. Calibration with respect to the slope of a predictor relaxes the assumption that the effect of the predictor is identical for all subjects. In what follows, we assume that subjects differ substantially with respect to their average speed and that they are differently sensitive to the effect of FREQUENCY. We also assume that they all show the same effect of LENGTH. Therefore, in the present example, we assume that calibration is required for the intercept and for the slope of FREQUENCY[1]. The random effects for subject i are modeled formally by defining $\mathbf{Z}_i$ to contain the first two columns of $\mathbf{X}_i$,

$$\mathbf{Z}_i = \begin{pmatrix} 1 & F_1 \\ 1 & F_2 \\ 1 & F_3 \\ 1 & F_4 \end{pmatrix} \quad (4)$$

and by introducing a (transposed) vector with the adjustments to intercept and slope for subject i, $\boldsymbol{b}_i^T = (\boldsymbol{b}_{0i}, \boldsymbol{b}_{1i})$. The product of $\mathbf{Z}_i$ and $\mathbf{b}_i$ is

---

[1] A mixed-effect model has minimally random intercepts for one random effect. Given a model with one random grouping factor (here, Subject), one must have random intercepts, and can have one or more random slopes. Addition of more random grouping factors (e.g., Items) will lead to more random intercepts, and possibly to more random slopes.

$$\mathbf{Z}_i\mathbf{b}_i = \begin{pmatrix} b_{0i} + F_1 b_{1i} \\ b_{0i} + F_2 b_{1i} \\ b_{0i} + F_3 b_{1i} \\ b_{0i} + F_4 b_{1i} \end{pmatrix} \quad (5)$$

with $\boldsymbol{b_{0i}}$ bringing individual differences with respect to the intercept into the model, and $\boldsymbol{F_1 b_{1i}}$ representing the fine-tuning of the slope for FREQUENCY. It follows straightforwardly from (1), (3), and (5) that

$$\begin{aligned} \mathbf{y}_i &= \mathbf{X}_i\beta + \mathbf{Z}_i\mathbf{b}_i + \epsilon_i \\ &= \begin{pmatrix} \beta_0 + F_1\beta_1 + L_1\beta_2 + b_{0i} + F_1 b_{1i} + \epsilon_{i1} \\ \beta_0 + F_2\beta_1 + L_2\beta_2 + b_{0i} + F_2 b_{1i} + \epsilon_{i2} \\ \beta_0 + F_3\beta_1 + L_3\beta_2 + b_{0i} + F_3 b_{1i} + \epsilon_{i3} \\ \beta_0 + F_4\beta_1 + L_4\beta_2 + b_{0i} + F_4 b_{1i} + \epsilon_{i4} \end{pmatrix} \quad (6) \\ &= \begin{pmatrix} (\beta_0 + b_{0i}) + (\beta_1 + b_{1i})F_1 + L_1\beta_2 + \epsilon_{i1} \\ (\beta_0 + b_{0i}) + (\beta_1 + b_{1i})F_2 + L_2\beta_2 + \epsilon_{i2} \\ (\beta_0 + b_{0i}) + (\beta_1 + b_{1i})F_3 + L_3\beta_2 + \epsilon_{i3} \\ (\beta_0 + b_{0i}) + (\beta_1 + b_{1i})F_4 + L_4\beta_2 + \epsilon_{i4} \end{pmatrix} \quad (7) \end{aligned}$$

In this model, we have in all three random effects: $\mathbf{b_{0i}} \sim \boldsymbol{N}\ \mathbf{(0, \sigma_{b0})}$ and $\mathbf{b_{1i}} \sim \boldsymbol{N}\ \mathbf{(0, \sigma_{b1})}$, and the residual error $\boldsymbol{\epsilon_k} \sim \boldsymbol{N}\ \mathbf{(0, \sigma^2)}$. It can be described as a mixed-effects model with random intercepts and with random slopes for FREQUENCY. As can be seen from (7), a mixed-effects regression model with random intercepts and random slopes for FREQUENCY is similar to a by-subject regression, with as key difference that the non-significant differences in the slopes for LENGTH have been made part of the error term. One of the advantages of mixed-effects modeling is that more precise estimates of the adjustments $\boldsymbol{b_{0i}}$ and $\boldsymbol{b_{1i}}$ are obtained thanks to shrinkage[2], which allows more precise prediction. Another advantage is that longitudinal effects due to, for instance, familiarization or fatigue, can be brought into the model.

The multi-level analysis reported in 2.4 was carried out using a model with the overall structure of (1). As in the by-subject regression advocated by Lorch et al

[2] The concept of shrinkage can be best explained considering an example. In random regression (Lorch and Meyer, 1990), subjects that are extremely fast or extremely slow in one experiment, cause a wide range of estimates for the intercept. However, it is known that in repetition estimates tend to regress towards the mean. In a repetition, subjects are likely to cause less extreme estimates. Mixed-effect regression anticipates this shrinkage towards the mean and brings it into the model. Random regression, therefore, provides a (too) tight fit to the data, while mixed-effect regression provides better predictions.

(1990), this model does not bring variability that is linked to the words into the model. At the time that the PNAS paper (sections 1 to 5 of Chapter 2) was written, the only way in which Word could be brought into the model equation was by nesting Word under Subject, as mixed-effects models were initially developed for nested hierarchical designs (e.g., for data with pupils nested under schools nested under towns). Such models are often referred to as hierarchical linear models (HLMs). Although it has been argued that nesting of Word under Subject is exactly the right thing to do (e.g., Quené et al., 2004), recent studies (e.g., Baayen, in press) suggest that ignoring by-item variance that is crossed with rather than nested under Subject may lead to Type I error rates that are as high as 25% to 30%. By nesting Word under Subject, the model fitting algorithm is not constrained to assigning a given word a constant effect, however small, across all subjects. Effectively, the model is free to consider the set of experimental words as completely different items for each subject. This freedom also characterizes the by-subject regression.

This can be demonstrated by considering two experimental designs. In one design, we expose all subjects to the same set of words, with FREQUENCY and LENGTH as predictors. In the other design, we expose each subject to a set of words that has no overlap with the sets of words shown to other subjects. Again, FREQUENCY and LENGTH are considered as predictors. The assessment of the significance of FREQUENCY and LENGTH in a random regression model as advocated by Lorch et al. (1990) proceeds in exactly the same way for both designs, even though the variance due to the words is much more constrained in the first design.

In general, random regression, mixed-effects regression with Subject as random effect, and mixed-effects regression with Word nested under Subject perform with nearly identical high Type I error rates when the data are characterized by an independent random effect for Word (Baayen, in press; Baayen, Davidson, & Bates, subm.).

Mixed-effects models with Subject and Word as crossed random effects

It has recently become possible to fit mixed-effects models which contain independent random effects for Subject and for Word. Such models are referred to as including crossed random effects. The structure of the new mixed-effects models that I have considered in this doctoral dissertation is described by the following equation (i denotes subjects, j denotes items).

$$\mathbf{y}_{ij} = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{S}_i\mathbf{s}_i + \mathbf{W}_j\mathbf{w}_j + \epsilon_{ij} \tag{8}$$

In the analyses, W_j was always a vector of ones and w_j, therefore, was always restricted to a by-item adjustment to the intercept. Hence, continuing with the above example, the expected onset latencies for this model are

$$\begin{aligned}
\mathbf{y}_{ij} &= \mathbf{X}_i\boldsymbol{\beta} + \mathbf{S}_i\mathbf{s}_i + \mathbf{W}_j\mathbf{w}_j + \epsilon_{ij} \\
&= \begin{pmatrix} \beta_0 + F_1\beta_1 + L_1\beta_2 + s_{0i} + F_1 s_{1i} + w_{01} + \epsilon_{i1} \\ \beta_0 + F_2\beta_1 + L_2\beta_2 + s_{0i} + F_2 s_{1i} + w_{02} + \epsilon_{i2} \\ \beta_0 + F_3\beta_1 + L_3\beta_2 + s_{0i} + F_3 s_{1i} + w_{03} + \epsilon_{i3} \\ \beta_0 + F_4\beta_1 + L_4\beta_2 + s_{0i} + F_4 s_{1i} + w_{04} + \epsilon_{i4} \end{pmatrix} && (9) \\
&= \begin{pmatrix} (\beta_0 + s_{0i} + w_{01}) + (\beta_1 + s_{1i})F_1 + L_1\beta_2 + \epsilon_{i1} \\ (\beta_0 + s_{0i} + w_{02}) + (\beta_1 + s_{1i})F_2 + L_2\beta_2 + \epsilon_{i2} \\ (\beta_0 + s_{0i} + w_{03}) + (\beta_1 + s_{1i})F_3 + L_3\beta_2 + \epsilon_{i3} \\ (\beta_0 + s_{0i} + w_{04}) + (\beta_1 + s_{1i})F_4 + L_4\beta_2 + \epsilon_{i4} \end{pmatrix} && (10)
\end{aligned}$$

Mixed-effects models with Subject and Item as crossed random effects are much more conservative than the corresponding by-subject regression models, and in fact tend to report significance levels that are similar to those obtained with by-item regression models.

To summarize, mixed-effects models, in general, offer the advantage of shrinkage estimates for the by-Subject and by-Item adjustments to slopes and intercept, as well as the possibility to bring longitudinal effects under control. The newly available mixed-effects models with crossed random effects for Subject and Item are superior to other mixed-effects models, because they are not prone to increased Type I error rates. Therefore, the statistical analyses reported in Chapters 3-5 all make use of the general model defined in (10). At the end of Chapter 2, I present a reanalysis of the data reported in Bien, Levelt and Baayen (2005) which had been analyzed there with only Subject as random effect. The reanalysis includes Word as random effect, and leads to a more parsimonious model.

References

Aronoff, M. (1994): Morphology by itself. MIT Press, Cambridge, Mass.

Baayen, R.H., Davidson, D.J., & Bates, D.M. (submitted). Mixed-effects modeling with crossed random effects for subjects and items.

Baayen, R.H. (2007) *Analyzing linguistic data. A practical introduction to statistics*. Cambridge University Press, in press.

Baayen, R. H., Feldman, L. F., & Schreuder, R. (2006). Morphological influences on the recognition of monosyllabic monomorphemic words. *Journal of Memory and Language*, *53*, 496-512.

Baayen, R.H., Tweedie, F.J., & Schreuder, R. (2002). The subjects as a simple random effect fallacy: Subject variability and morphological family effects in the mental lexicon, *Brain and Language*, *81*, 55–65.

Bates, D.M., & Sarkar, D. (2005). lme4: Linear mixed-effects models using S4 classes, R package version 0.9975-7.

Bates, D.M. (2005). Fitting linear mixed models in R, *R News*, *5*, 27-30.

Bien, H., Levelt, W.J.M., & Baayen, R.H. (2005). Frequency effects in compound production. *Proceedings of the National Academy of Sciences*, *102(49)*, 17876-1788.

Burani, C., & Caramazza, A. (1987). Representation and processing of derived words. *Language & Cognitive Processes*, *2*, 217-227.

Butterworth, B. (1983). Lexical representation. In B. Butterworth (Ed.), *Language production: Vol. 2. Development, writing and other language processes*. London: Academic Press.

Caramazza, A., & Miozzo, M. (1998). More is not always better. A response to Roelofs, Meyer, & Levelt. *Cognition*, *69*, 231-241.

Caramazza, A., Costa, A., Miozzo, M., & Bi, Y. (2001). The specific-word frequency effect: Implications for the representation of homophones in speech production. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *27*, 1430-1450.

Coltheart, M., Davelaar, E., Jonasson, J.T., & Besner, D (1977). Access to the internal lexicon. In Dornic, S. (Ed.), Attention and Performance VI. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Cutler, A. (1981). Making up materials is a confounded nuisance, or: Will we be able to run any psycholinguistic experiments at all in 1990? *Cognition*, *10*, 65-70.

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

Dohmes, P., Zwitserlood, P., & Bölte, J. (2004). The impact of semantic transparency of morphologically complex words on picture naming. *Brain and Language*, *90*, 203-212.

Gumnior, H., Bölte, J., & Zwitserlood, P. (2006). A chatterbox is a box: Morphology in German word production. *Language and Cognitive Processes*, *21*, 920-944.

Greenberg, J.H., & Jenkins, J.J. (1964). Studies in the psychological correlates of the sound system of American English. *Word*, *20*, 157-177.

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

Jescheniak, J.D., Meyer, A.S., & Levelt, W.J.M. (2003). Specific word frequency is not all that counts in speech production. Reply to Caramazza et al. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *29*, 432-438.

Kessler, B., Treiman, R., & Mullennix, J. (2002). Phonetic biases in voice key response time measurements. *Journal of Memory and Language*, *47*, 145-171.

Krott, A., Baayen, R. H., & Schreuder, R. (2001). Analogy in morphology: modelling the choice of linking morphemes in Dutch. *Linguistics*, *39(1)*, 51-93.

Kuperman, V., Pluymaekers, M., Ernestus, M., & Baayen, R. H. (subm). Morphological predictability and acoustic salience of interfixes in Dutch compounds. To appear in Journal of the Acoustical Society of America.

Laudanna, A., & Burani, C. (1985). Address mechanisms to decomposed lexical entries. *Linguistics, 23,* 775-792.

Levelt, W.J.M., Praamstra, P., Meyer, A.S., Helenius, P., & Salmelin, R. (1998). An MEG study of picture naming. *Journal of Cognitive Neuroscience*, *10*, 553-567.

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

Lorch, R.F., & Myers, J.L. (1990). Regression analyses of repeated measures data in cognitive research. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 149-157.

Luce, P.A., & Pisoni, D.B. (1998). Recognizing spoken words: The Neighborhood Activation Model. *Ear & Hearing*, *19*, 1-36.

Marslen-Wilson, W. (1990). Activation, competition, and frequency in lexical access. In G. Altmann (Ed.), *Cognitive Models of speech processing*, pp. 148-172. Cambridge: MIT Press.

Melinger, A. (2003). Morphological structure in the lexical representation of prefixed words: Evidence from speech errors. *Language and Cognitive Processes*, *18* (3), 335-362.

Meyer, A.S., Roelofs, A., & Levelt, W.J.M. (2003). Word length effects in object naming: The role of a response criterion. *Journal of Memory and Language*, *48*, 131-147.

Meyer, A.S. (1990). The time course of phonological encoding in language production: The encoding of successive syllables of a word. *Journal of Memory and Language, 29,* 524-545.

Meyer, A.S. (1991). The time course of phonological encoding in language production: The encoding inside a syllable. *Journal of Memory and Language, 30,* 69-89.

Moscoso del Prado Martín, F., Kostic, A., & Baayen, R. H. (2004). Putting the bits together: An information theoretical perspective on morphological processing. *Cognition*, *94*, 1-18.

Oldfield R.C., & Wingfield A. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology*, *17*, 273-281.

Pinheiro, J.C., & Bates, D.M. (2000). Mixed-effects models in S and S-PLUS, *Statistics and Computing Series*, Springer-Verlag, New York, NY.

Quené, H., & Van den Bergh, H. (2004). On multi-level modeling of data from repeated measures designs: A tutorial, *Speech Communication*, *43 (1-2)*, 103-121.

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

Roelofs, A., & Baayen, H. (2002), Morphology by itself in planning the production of spoken words. *Psychonomic Bulletin & Review, 9*, 132-138.

Scarborough, R. A. (2004). Degree of coarticulation and lexical confusability. *Proceedings of the 29th Meeting of the Berkeley Linguistics Society, February 14-17, 2003*.

Sevald, C.A., & Dell, G.S. (1994). The sequential cuing effect in speech production. Cognition, 53, 91-127.

Shannon, C. E., & Weaver, W. (1949). *The mathematical theory of communication*. University of Illinois Press, Urbana.

Shannon, C.E. (1948). A mathematical theory of communication, *The Bell System Technical Journal*, *27*, 379-423.

Shatzman, K.B., & Schiller, N.O. (2004). The word frequency effect in picture naming: Contrasting two hypotheses using homonym pictures. *Brain and Language*, *90*, 160-169.

Van Son, R.J.J.H., & Pols, L.C.W. (2003). How efficient is speech? *Proceedings of the Institute of Phonetic Sciences*, *25*, 171-184.

Van Son R.J.J.H., & Van Santen, J.P.H. (2005). Duration and spectral balance of intervocalic consonants: A case for efficient communication. *Speech Communication*, *47*, 100-123.

Vitevitch, M.S., & Stamer, M.K. (2006). The curious case of competition in Spanish speech production. *Language & Cognitive Processes, 21,* 760-770.

Vitevitch, M.S. (2002). The influence of phonological similarity neighborhoods on speech production. *Journal of Experimental Psychology: Learning, Memory and Cognition, 28,* 735-747.

Zwitserlood, P., Bölte, J., & Dohmes, P. (2000). Morphological effects on speech production: Evidence from picture naming. *Language and Cognitive Processes*, *15*, 563-591.

FREQUENCY EFFECTS IN COMPOUND PRODUCTION

CHAPTER 2

Heidrun Bien, Willem J. M. Levelt, & R. Harald Baayen [1]

Abstract

Four experiments investigated the role of frequency information in compound production by independently varying the frequencies of the first and second constituent as well as the frequency of the compound itself. Pairs of Dutch noun-noun compounds were selected, such that there was a maximal contrast for one frequency, while matching for the other two frequencies. In a position-response association task, participants first learned to associate a compound with a visually marked position on a computer screen. In the test phase, participants had to produce the associated compound in response to the appearance of the position mark and we measured speech onset latencies.

The compound production onset latencies varied significantly according to factorial contrasts in the frequencies of both constituting morphemes, but not according to a factorial contrast in compound frequency, providing further evidence for decompositional models of speech production. In a stepwise regression analysis of the joint data of all four Experiments, however, compound frequency was a significant non-linear predictor, with facilitation in the low-frequency range and a trend towards inhibition in the high-frequency range. Furthermore, a combination of structural measures of constituent frequencies and entropies explained significantly more variance than a strict decompositional model including cumulative stem

[1] This chapter (section 2.1-5) originally appeared in Proceedings of the National Academy of Science (2005), 102, 17876-17881. In section 2.6, I present an additional analysis of the compound naming latencies with crossed random effects for subject and item.

frequency as the only measure of constituent frequency, suggesting a role for paradigmatic relations in the mental lexicon.

2.1. Introduction

High-frequency words are produced more quickly than low-frequency words. Since the seminal study of Oldfield and Wingfield (1965), the effect of word-frequency has emerged to be replicable and robust. In their series of timed picture naming studies in seven languages, Bates, D'Amico, Jacobsen, Szekely, Andonova, Devescovi, Herron, ChingLu, Pechmann, Pl'eh et al. (2003) found large frequency effects in all the seven languages studied. Jescheniak and Levelt (1994) observed that the frequency effect for lemma retrieval diminished quickly over repetition, but that the frequency effect for a word's form (lexeme) remained stable across repetitions. The cumulative homophone effect reported in that study suggests that the effect of word frequency arises at the level of word form, rather than conceptualization or articulation. Word frequency has, therefore, been attributed (Levelt, Roelofs, & Meyer, 1999) to the access of a word's phonological code (but see Caramazza, Costa, Miozzo, & Bi, 2001; Jescheniak, Meyer, & Levelt, 2003). The general finding that a word's frequency is correlated with its production latency has become a powerful experimental tool.

In this study, we address frequency effects in the production of Dutch compounds in order to distinguish decompositional from non-decompositional models of production. Fully non-decompositional theories predict frequency effects for each individual form of occurrence. In such theories, only the specific frequency of morphologically complex words such as *handbag* is expected to be predictive of their production latency. We will refer to this form-specific frequency as the word form frequency. In the case of compounds such as *handbag*, fully non-decompositional theories distinguish the word form frequency of the singular *handbag* from the word form frequency of the plural *handbags*.

In a fully decompositional model such as (Levelt et al., 1999; Levelt, 2001) and its computer simulation WEAVER++ (Roelofs, 1997), all complex words are assembled from their constituent morphemes. The more often a morpheme has been used, the lower its activation threshold. Hence, the proper frequency measure for predicting production onset latencies in WEAVER++ for words such as *hand* is the summed frequency of all variants of that word, whether part of inflected (*hands*), derived (*handy*) or compound words (*handbag*). Each of those occurrences is assumed to leave a frequency trace on the stem *hand*. In what follows, we will refer to the summed frequencies of a word as its CUMULATIVE STEM[2] FREQUENCY (see Laudanna & Burani, 1985; Burani & Caramazza, 1987; Schreuder & Baayen, 1997) for CUMULATIVE STEM FREQUENCY effects reported for comprehension).

According to WEAVER++ (Roelofs, 1997), the CUMULATIVE STEM FREQUENCIES of the constituents *hand* and *bag* should be the relevant frequency measures for predicting the production latency of a compound such as *handbag*. Roelofs (1996), using the implicit priming paradigm (Meyer, 1990, 1991), addressed the question of whether the form lexicon underlying speech production contains morphologically decomposed entries. In this paradigm, subjects produce words from learned paired-associates. Homogenous response sets, in which all response words began with the same morpheme, resulted in shorter naming onset latencies than heterogeneous response sets, in which all initial morphemes of response words were different. Crucially, this preparation effect was larger for words with initial low-frequency morphemes than for ones with high-frequency morphemes and the effect was stable in repeated measurements. This is because low-frequency morphemes have more to gain from implicit priming than high-frequency morphemes. This finding supports decompositional theories in which the constituents of words like *handbag* are individually accessed during the production process. Currently, WEAVER++ implements the most parsimonious decompositional theory, by assuming that the form representation for *hand* in *handbag* is the same as the one for the word *hand* itself.

An intermediate position between non-decompositional and fully decompositional models is to assume structured storage. Instead of storing *handbag* at the form level as two independent monomorphemes, it might be stored with information about their combination. In a model with structured storage, the

[2] CUMULATIVE STEM FREQUENCY and CUMULATIVE ROOT FREQUENCY denote the same variable

frequencies of *hand* as the first constituent of any compound or *bag* as head of any compound can be more precise predictors than their frequencies as independent words. We will refer to a set of compounds sharing the same left (or right) constituent as the left (or right) constituent family, following (Krott, Baayen, & Schreuder, 2001). For each constituent family, we have a distribution of the compound frequencies of its members. One way to obtain a point estimator of such a distribution is to sum the frequencies of its members (henceforth, POSITIONAL FREQUENCY). Another point estimator of this distribution is to compute Shannon's entropy for the probability distribution estimated by the relative frequencies (the frequencies normalized with positional frequency, henceforth, POSITIONAL ENTROPY) (Shannon, 1948; Shannon, & Weaver, 1949).

Both the POSITIONAL FREQUENCY and the POSITIONAL ENTROPY are measures calculated for the range of alternative compounds sharing the same morpheme in the same position. While the POSITIONAL FREQUENCY adds up the frequencies of the constituent family members, the POSITIONAL ENTROPY takes into account the probability distribution within the family. Other constituent frequency measures that are of potential interest are the summed frequencies of all other complex words in which the constituent appears (henceforth, COMPLEMENT FREQUENCY) and the entropy of the constituents calculated over the full range of morphologically complex words in which the constituent appears (henceforth, DERIVATIONAL ENTROPY).

Finally, we define the summed frequencies of a word's inflectional variants as LEMMA FREQUENCY. In the case of a simple or complex noun, the inflectional variants are the singular and the plural forms. The CONSTITUENT LEMMA FREQUENCY of *hand* in *handbag*, therefore, refers to the sum of the frequencies of *hand* and *hands*, while the lemma frequency of the compound (henceforth, COMPOUND FREQUENCY) refers to the sum of the frequencies of *handbag* and *handbags*, respectively. In general, the LEFT (or RIGHT) CONSTITUENT LEMMA FREQUENCY, CUMULATIVE STEM FREQUENCY, POSITIONAL FREQUENCY, POSITIONAL ENTROPY, COMPLEMENT FREQUENCY and DERIVATIONAL ENTROPY are strongly correlated.

2.1.1 Key questions

In this paper we address three key questions. First, can separate frequency effects for the constituents of compounds be ascertained? Constituent frequency effects would provide further evidence against full unstructured storage models. Roelofs (1996) observed constituent frequency effects for the left constituent using implicit priming. We seek to replicate this finding, using immediate naming. In addition, we examine whether a similar frequency effect can be found for the right constituent. It is not self-evident that a frequency effect for the right constituent should exist, even within decompositional theories. Various studies (Roelofs, 1996; Cholin, Schiller, & Levelt, 2004) have shown that production proceeds incrementally, suggesting that the frequency of the second constituent might become relevant only after completion of the planning and initiation of the articulation of the first constituent. However, there are circumstances in which the length of the word co-determines object naming onset latencies (Meyer, Roelofs, & Levelt, 2003), indicating that speakers may plan the complete phonological word before speech onset. This evidence is in line with recent studies addressing the acoustic realization of complex words. Stems pronounced in isolation tend to have longer durations, and tend to be produced with a different intonation contour, than the same stems appearing as the initial constituents of complex words, both for inflection and derivation (Kemps, Ernestus, Schreuder, & Baayen, 2005; Kemps, Wurm, Ernestus, Schreuder, & Baayen, 2005; Koester, Gunter, Wagner, & Friederici, 2004). This suggests that the planning of the articulation of the first constituent is to some extent dependent on the presence of a second constituent.

Second, does the frequency of the compounds contribute to its response latency? In non-decompositional models, compound frequency should be the only relevant measure. But in strict decompositional models, it should be irrelevant. In a model with structured storage a compound frequency effect cannot be ruled out, but it might be strongly modulated by the role of the constituent families.

Third, are constituent frequency effects on response latency best predicted from the cumulative stem frequencies of the constituents, or do we rather see different effects of the more specific frequencies measures, including positional measures based on the constituent families? This comparison allows us to distinguish between full decomposition as in WEAVER++ and structured storage.

In our experiments, we systematically manipulated the frequency of the first morpheme, of the second morpheme, and of the compound itself, to examine their individual influences on compound production latency. For each experiment, we selected pairs of Dutch compounds as targets, such that there was a maximal contrast for one frequency factor, while the other two (exception: Exp 3) were matched. The contrasts were constructed in terms of lemma frequency but coincided with a range of additional contrasts such as cumulative stem frequency, positional frequency or derivational entropy for the constituents. Following the factorial analyses of Experiments 1-4, we present a regression analysis addressing the question, which of these frequency measures are the appropriate predictors.

All compounds used in this study are semantically transparent Dutch noun-noun compounds, with the first constituent being the modifier and the second constituent being the head noun.

2.2. Material and Method

From the CELEX lexical database (Baayen, Piepenbrock, & Guliker, 1995, CD-ROM), we selected Dutch noun-noun compounds on the basis of three frequency counts: the lemma frequency (the summed frequencies of the word's inflectional variants) of the compound[3], the lemma frequency of its left constituent, and the lemma frequency of its right constituent. (All CELEX frequencies reported here and below are counts based on a corpus of 42 million words.) For each experiment, we selected 32 compounds which consisted of 16 pairs.

2.2.1. Pairs in Experiment 1: The frequency of the head noun

In Experiment 1, the compounds in a pair shared the first morpheme (e.g., *<u>lucht</u>brug* - *<u>lucht</u>buks*, airlift - airgun), they were matched for compound frequency, and differed with respect to the frequency of the second constituent, which was either high (mean:

[3] For all compounds we additionally collected familiarity ratings of 16 participants and the frequencies with which the compounds appear in Google, to double check the CELEX frequencies.

13354, median: 5731, range: 437-50439) or low (mean: 145, median: 61, range: 1-867).

The sixteen compounds with low-frequency second constituents were arranged into eight pairs (e.g., *luchtbuks* - *broodkruim*) with the constraint that the pair mates had minimal phonological overlap, no obvious semantic relation, and had similar compound frequencies. The same was done for the sixteen compounds with high-frequency second constituents. Presentation was blocked by condition and the order of the blocks was counterbalanced. Half of the participants started out with the eight pairs with low-frequency second constituents, the other half began with the eight pairs with high-frequency heads. We blocked the conditions in order to minimize effects of criterion-setting that would lead to elongated responses for otherwise short reaction times and to speeded responses for otherwise long reaction times (Meyer et al., 2003; Lupker, Brown, & Colombo, 1997). The sixteen subsets of target compounds were complemented by three practice subsets, with compounds of similar structure and frequency.

2.2.2. Pairs in Experiment 2: The frequency of the modifier

In Experiment 2, the compounds in a pair shared the second morpheme, and were matched for compound frequency. The first morpheme carried a factorial contrast between high (mean: 8072, median: 7111, range: 1424-23062) and low (mean: 660, median: 356, range: 39-2645) frequency. As described for Experiment 1, we blocked the conditions by rearranging the items into eight subsets of compounds with low-frequency modifiers and eight subsets of compounds with high-frequency modifiers and added three practice subsets of similar structure and frequency.

2.2.3. Pairs in Experiment 3: Frequency contrasts for head and modifier

In Experiment 3, we selected thirty-two compounds, pairwise-matched for compound frequency. Within a pair, one compound had high-frequency constituents (mean: 10400, median: 7213, range: 409-48452) and the other low-frequency constituents (mean: 618, median: 291, range: 4-4416). A given constituent appeared in only one compound. We used the same blocking strategy as described for the Experiments 1 and 2 and complemented the resulting sixteen subsets of condition-intern repairings by three practice subsets of similar structure and frequency.

2.2.4. Pairs in Experiment 4: The frequency of the compound

In Experiment 4, thirty-two compounds were selected, pairwise-matched both according to the frequency of the first morpheme and according to the frequency of the second morpheme, while the frequency of the compound was either high (mean: 973, median: 897, range: 516-2369) or low (mean: 48, median: 39, range: 6-132). As described for the previous experiments, we blocked the conditions, creating eight subsets of low-frequency compounds and eight subsets of high-frequency compounds that were complemented by three similar practice subsets.

2.2.5. Participants

For each experiment, twenty-four native speakers of Dutch[4] were recruited from the subject pool of the Max Planck Institute for Psycholinguistics. None of them took part in more than one of the experiments. They received € 5 for their participation.

2.2.6. Position-Response Association Task

Participants were tested individually in a dimly lit sound-attenuated booth. They were comfortably seated in front of a CRT computer screen, a Sennheiser microphone and a cordless mouse and they were wearing headphones. On average, a session lasted 45 min.

We used a position-response association task (Cholin, Levelt, & Schiller, 2006), in which participants learned to associate the two compounds in a subset with visually marked positions on the left and right part of a computer screen. For each subset, the experimental procedure consisted of a learning phase, a practice phase and a test phase. Each phase was introduced by an attention signal presented on the screen for two seconds, and ended with a pause signal that remained on the screen until the following phase was initiated by the experimenter.

[4] The gender of the participants was as follows: Experiment 1: 20 female, 4 male; Experiment 2: 19 female, 5 male; Experiment 3: 17 female, 7 male; 19 female, 5 male.

In the learning phase, participants were presented with the two compounds over headphones. Simultaneously with hearing the first compound, they saw the icon of a loudspeaker appearing on the left side of the screen. Simultaneously with hearing the second compound, the same icon appeared on the right side of the screen. This procedure was repeated once.

In the practice phase, both icons (left and right) were visible with the cursor of the mouse in the center of the screen. The subject was then acoustically presented with one of the two compounds and had to click on the associated icon. Both compounds were presented twice and in random order. We provided the participants with feedback on their accuracy by displaying the number of errors (0-4) on the screen.

In the test phase, 20 trials of a distractor task alternated with 20 trials of the experimental task. A test phase always began with a distractor trial. In a distractor trial, one of five single-digit numbers (1, 2, 3, or 6) was presented in the center of the screen and had to be named as fast and correctly as possible. We included those distractor trials to avoid that participants would have to produce exactly the same word during consecutive trials. In other words, the insertion of distractor trials made it difficult for participants to use the break between two experimental trials to already prepare one of the target words. In an experimental trial, the icon of a loudspeaker was presented either on the left or right position of the screen, prompting the participant to say aloud the associated compound, again as fast and correctly as possible. Each position appeared a total of ten times. The two positions were presented in pseudo-random order with the restriction of a maximum of four consecutive repetitions of one position.

Simultaneous with the presentation of the icon, the voice key was activated for 1500 ms. Naming onset latencies longer than 1500 ms were counted as time-outs. The experimenter monitored the participant's responses through headphones and took notes of incorrect naming, hesitations and voice key errors.

2.3. Results

Only those compounds for which a correct response was obtained were included in the analysis. Time-out trials (>1500 ms) and extreme outliers (i.e., latencies outside a range of two standard deviations around the mean latency for each subject per condition (high or low) as well as for each item) were also removed from the analysis.

2.3.1. Results Experiment 1: The frequency of the head noun

Altogether, 421 trials were excluded (5%) in Experiment 1. Mean onset latencies, standard deviations, and error rates are summarized in Table 2.1.

We analyzed the onset latencies both by subjects and by items, with FREQUENCY as a within participants factor, and ORDER as a between participants factor. Compounds with a high right constituent frequency elicited shorter onset latencies (on average 14 ms) than compounds with a low-frequency head, both in the analysis by participants ($F1(1,22) = 5.8$, $p = 0.025$), and in the analysis by items ($F2(1,30) = 18.1$, $p <.001$). There was no effect of ORDER in the by-participant analysis ($F(1,22) = 2.1$, $p = 0.16$) and no interaction of ORDER by FREQUENCY ($F(1,22) = 1.8$, $p = 0.19$). In the by-item analysis, ORDER emerged as a significant main effect ($F(1,30) = 16.9$, $p < 0.001$) in interaction with FREQUENCY ($F(1,30) = 5.9$, $p = 0.021$). The interaction points to a significant difference between the 6 ms FREQUENCY effect for the high-low block order, and the 21 ms FREQUENCY effect for the low-high block order. An analysis of the error scores revealed no significant main effects, nor any interactions.

Apparently, the blocking strategy, which was chosen in order to avoid criterion setting, created an alternative problem, the interaction of frequency and practice. Due to practice, participants became faster, leading to shorter onset latencies in the second block compared to the first block. The speeding up was strong in the otherwise slow low-frequency set, while the already fast production of high-frequency items may benefit less from an additional effect of practice, underestimating the difference between low- and high-frequency items in the high-low order of presentation.

In short, the frequency of the head noun co-determines production onset latencies, even though it is not the initial constituent of a compound, providing evidence for decompositionality as well as evidence against strict incrementality. Apparently, articulation is not initiated before the phonological code of the head noun has been retrieved. Experiment 2 investigates the predictivity of the frequency of the initial constituent.

Table 2.1: Mean onset latencies in ms for Experiments 1-4 for the main effect of frequency (with standard deviations and error percentages) and for the Block Orders LH (first block low-frequency items, second block high-frequency items) and HL (first block high-frequency items, second block low-frequency items).

Exp	Freq	Mean	In LH	In HL
1	high	457 (111, 3)	437	476
	low	471 (116, 2)	458	482
2	high	443 (118, 5)	439	447
	low	468 (129, 5)	487	450
3	high	414 (105, 6)	405	424
	low	441 (115, 5)	445	437
4	high	442 (108, 4)	430	454
	low	434 (104, 4)	433	435

2.3.2. Results Experiment 2: The frequency of the modifier

Altogether, 814 trials (10%) were removed from the analysis following the criteria described above.

Compounds with a high LEFT CONSTITUENT FREQUENCY elicited shorter onset latencies (on average 25 ms) than compounds with a low-frequency modifier, both in the analysis by participants ($F1(1,22) = 19.4$, $p <.001$), and in the analysis by items ($F2(1,30) = 48.5$, $p <.001$). In the by-participant analysis, there was no main effect of ORDER ($F(1,22) = 0.2$, $p = 0.675$) but an interaction of ORDER by FREQUENCY ($F(1,22) = 13.5$, $p = 0.001$). In the by-item analysis, ORDER emerged as a significant main effect ($F(1,30) = 9.0$, $p = 0.005$) in interaction with FREQUENCY ($F(1,30) = 37.7$, $p < .001$). As in Experiment 1, the difference between the high and low frequency conditions was bigger in the low-high block order (48 ms) than in its reverse (3 ms). An error analysis revealed no significant main effects, nor interactions.

In short, this experiment shows that the frequency of the initial constituent affects compound production, as expected in a decompositional theory of compound production. We next investigated whether constituent frequency effects can be observed in the absence of compounds in the experiment that share head or modifier. In order to maximize constituent effects, only two conditions were tested: high versus low frequency for both constituents.

2.3.3. Results Experiment 3: Frequency contrasts for head and modifier

Time-out trials, voice key errors, extreme outliers, and incorrect naming responses were removed from the data set (863 trials, 11%).

Analyses of variance with FREQUENCY as a within factor and ORDER as a between factor revealed that compounds with high-frequency constituents elicited shorter onset latencies (on average 27 ms) than compounds with low-frequency constituents ($F1(1,22) = 20.7$, $p < .001$; $F2(1,30) = 42.7$, $p < .001$). There was no main effect of ORDER ($F1(1,22) = 0.2$, $p = 0.631$; $F2(1,30) = 1.4$, $p = 0.241$), but an interaction of ORDER by FREQUENCY ($F1(1,22) = 8.5$, $p = 0.008$; $F2(1,30) = 11.1$, $p = 0.002$). As before, the interaction suggested that the difference between the high and low frequency conditions was more prominent in the Low-High block order (40ms) than in the High-Low block order (13ms). Analysis of the error scores revealed no significant main effects, nor any interactions.

This experiment provides further support for the constituent frequency effects of Experiments 1 and 2, though not differentiating between the frequency effects of the first and of the second morpheme. It also rules out the possibility of a confound due to prior experience with a head or a modifier in the experiment.

In our final experiment we addressed the question of whether the production latency of a compound might be additionally affected by the compound's own frequency of occurrence.

2.3.4. Results Experiment 4: The frequency of the compound

A total of 652 trials (8%) was excluded from the analyses following the criteria defined above.

We analyzed the onset latencies with analyses of variance both by participants and by items. FREQUENCY was a within factor, and ORDER was a between factor. On average, the high-frequency compounds elicited onset latencies that were 8 ms longer than the low-frequency compounds. This difference in the direction of an anti-frequency effect did not reach full significance in the by-participants analysis ($F1(1,22) = 3.5$, $p = 0.074$), but reached significance in the by-item analysis ($F2(1,30) = 6.8$, $p = 0.014$).

In the analysis by participants, there was no main effect of ORDER ($F(1,22) = 0.4$, $p = 0.515$) but an interaction of ORDER by FREQUENCY ($F(1,22) = 5.6$, $p = 0.028$). In the analysis by items, ORDER emerged as a significant main effect ($F(1,30) = 9.9$, $p = 0.004$) in interaction with FREQUENCY ($F(1,30) = 11.7$, $p = 0.002$).

The interaction points to a significant difference between the 19 ms anti-frequency effect for the high-low block order, and the 3 ms frequency effect for the low-high order. We have argued so far that the interaction of block order and frequency reflects a practice effect, which leads to an underestimation of the frequency effect, when the slow block is presented last. In Experiment 1-3, the slower block clearly was the block with the items of the low-frequency condition. Here, the situation seems to be reversed, with the items of high compound-frequency displaying a practice effect. If so, the anti-frequency effect would be underestimated in the low-high block order. An analysis of the error scores revealed no effects.

In summary, high-frequency compounds were not produced any faster than low-frequency compounds. If anything, they elicited longer naming onset latencies.

2.4. Comparing Frequency Measures

Material selection for the Experiments 1-4 was based on LEMMA FREQUENCY. Table 2.2 shows for Experiments 1-3, that the high versus low frequency conditions for the constituents implemented contrasts in all the different measures of constituent frequency and entropy that we defined in the introduction.

Table 2.2: The frequency characteristics of the material in Experiments 1-3 (log-transformed) for the left/right constituent. CUMULATIVE STEM FREQUENCY (CumFreq), LEMMA FREQUENCY (LemFreq), COMPLEMENT FREQUENCY (ComplFreq), POSITIONAL FREQUENCY (PosFreq), POSITIONAL ENTROPY (PosEntr), DERIVATIONAL ENTROPY (DerEntr).

		CumFreq		LemFreq		ComplFreq		PosFreq		PosEntr		DerEntr	
		left	right	left	right	left	right	left	right	left	right	left	right
Exp 1	high	9.44	10.01	8.89	9.59	8.59	8.95	8.00	7.89	13.87	10.47	21.39	14.78
	low	9.44	6.88	8.89	5.76	8.59	6.49	8.00	5.91	13.87	2.46	21.39	9.36
Exp 2	high	9.61	9.59	9.00	9.01	8.73	8.59	7.90	7.99	12.30	12.97	22.98	16.86
	low	7.29	9.59	6.50	9.01	6.64	8.59	6.28	7.99	7.27	12.97	9.31	16.86
Exp 3	high	9.99	9.82	9.43	9.03	9.04	8.99	7.77	7.90	13.12	11.32	22.80	18.24
	low	7.25	7.86	6.42	6.44	6.65	6.95	5.68	6.64	6.00	6.17	12.79	12.26

To further examine their predictivity, we included these measures in a stepwise regression analysis of the joint data of Experiments 1-4, along with NEIGHBORHOOD DENSITY (Vitevitch, 2002), defined as the number of words that are similar to a target on the basis of the substitution of a single phoneme only (Coltheart, Davelaar, Jonasson, & Besner, 1977) and factors controlling for the sensitivity of the voicekey, addressing the nature of the onset phoneme. We started out with a variety of specifications such as voicing, fricative, nasal etc., but only the factor PLOSIVE vs. non-plosive turned out to be a significant predictor. Figure 2.1 visualizes the partial effects of the covariates.

A stepwise multilevel analysis of covariance (Pinheiro & Bates, 2000; Baayen, Tweedie, & Schreuder; 2002; Baayen, 2004; Quené, & Van den Bergh, 2004) with participant as main grouping factor revealed effects of the manner of articulation of the initial consonant ($\beta = 0.0240$, $t(28822) = 7.5595$, $p < 0.0001$), PLOSIVES elicited longer naming onset latencies (Panel 9), probably an artifact of the voicekey.

Panels 1 and 4 illustrate the facilitatory, linear effects of the left and right positional entropies adjusted for the effects of the other covariables (LEFT POSITIONAL ENTROPY: $\beta = -0.0081$, $t(28822) = -4.1796$, $p < 0.0001$; RIGHT POSITIONAL ENTROPY: $\beta = -0.0098$, $t(28822) = -4.1716$, $p < 0.0001$). Panel 2 and 5 picture the facilitatory, linear effect of the LEFT COMPLEMENT FREQUENCY ($\beta = -0.0077$, $t(28822) = -6.2234$, $p < 0.0001$), and the inhibitory, linear effect of the RIGHT COMPLEMENT FREQUENCY ($\beta =$

0.0055, t(28822) = 3.7080, p = 0.0002). Notice, that this effect of complement frequency is significant for both head and modifier, but in the opposite direction.

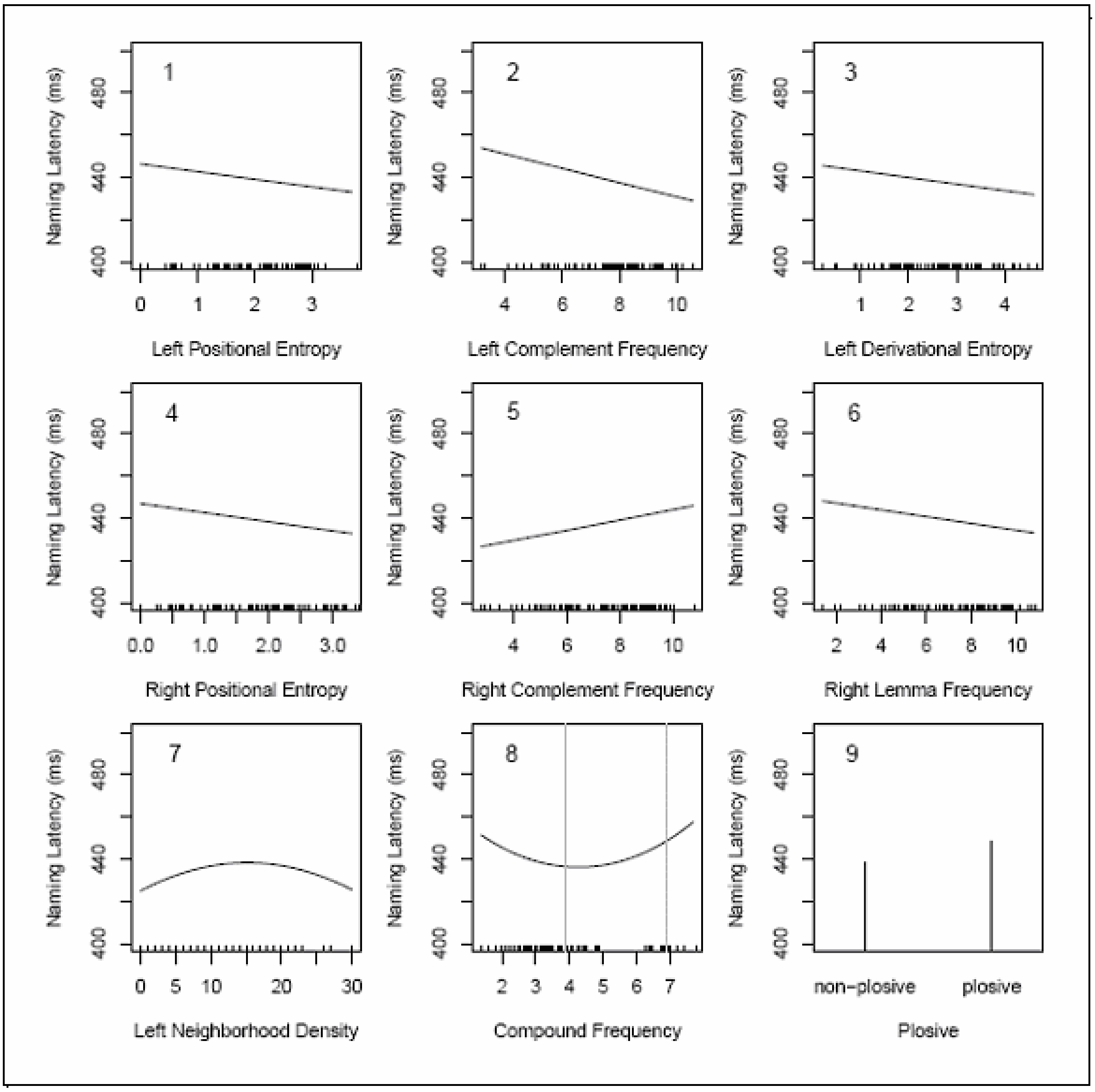

Figure 2.1: Partial effects of the predictors in the multilevel covariance analysis of the data of Experiments 1–4. The left vertical axis shows the effect in log units; the right axis shows the effect in milliseconds. Values pertain to words that do not begin with a plosive and are adjusted for the effects of the other covariates at their median value.

The more often the modifier appears as a constituent in other complex words, independent from its position within those words, the faster the compound is named.

In contrast, the more often the head constituent appears within other complex words, the slower the compound is named. For the head, and only for the head, however, we also observe an effect of lemma frequency, and this effect is facilitative (RIGHT LEMMA FREQUENCY, $\beta = -0.0036$, $t(28822) = -2.9686$, $p = 0.0030$).

The higher the frequency of the head as an independent word, the faster the compound is named (Panel 3). For the modifier, we further observe a linear, facilitatory effect of derivational entropy ($\beta = -0.0071$, $t(28822) = -3.8231$, $p = 0.0001$, LEFT DERIVATIONAL ENTROPY) as plotted in Panel 6.

Panel 7 shows the nonlinear curve for the neighborhood density of the initial constituent ($\beta = 0.0041$, $t(28822) = 5.2589$, $p < 0.0001$ for the linear component of the LEFT NEIGHBORHOOD Density, for its quadratic component, $\beta = -0.0001$, $t(28822) = -4.9338$, $p < 0.0001$), suggesting facilitation for left constituents with very sparse or very dense phonological similarity neighborhoods. There was no effect of the neighborhood density of the right constituent. Panel 8 illustrates the nonlinear curve for COMPOUND FREQUENCY ($\beta = -0.0351$, $t(28822) = -6.3141$, $p < 0.0001$ for its linear component, $\beta = 0.0041$, $t(28822) = 6.5850$, $p < 0.0001$ for its quadratic component).

While in the lower range of compound frequencies we see a facilitatory effect, this effect levels off and turns into inhibition in the higher range of compound frequencies. The two grey, vertical lines in this panel mark the averages of the low- and high-frequency conditions used in Experiment 3, illustrating why we did not observe a reliable effect of COMPOUND FREQUENCY in that experiment. The factorial contrast tended to balance low-frequency facilitation and high-frequency inhibition.

Finally, we compared the predictivity of our model with the predictivity of a strict decompositional model, a model in which the CUMULATIVE STEM FREQUENCIES of the left and right constituents were the only measures of constituent frequency entered into the regression equation. Both models included the non-frequency predictors NEIGHBORHOOD DENSITY and PLOSIVE. The more complex model explained significantly more variance than the strict decompositional model (Log-likelihood ratio test, $p < 0.001$) with a 61 percent increase in the variance explained by linguistic predictors.

Summing up, the factorial analyses of Experiments 1-4 showed that the naming latency of a compound was affected by the frequencies of its constituents. The regression analysis revealed that the naming latency of a compound is affected by a combination of different measures of frequency and entropy for both constituents. Interestingly, a qualitative difference emerged with respect to how the left and right constituents were affected. While for the modifier all significant effects of frequencies and entropies are facilitative, there is facilitation as well as inhibition for the head-constituent. The total outcome is facilitation in both cases, with greater facilitation for the modifier. In the regression analysis, we also observed that compound frequency was one of the factors with explanatory value. The nonlinear effect of COMPOUND FREQUENCY suggests facilitation within the lower frequency range, combined with inhibition in the higher frequency range. This inhibition might represent a floor effect, however, as it might be an artifact of modeling nonlinearity with a simple, quadratic polynomial.

2.5. Discussion

This study addressed three key questions concerning the role of frequency in compound production. First, are there separate frequency effects for the constituents of a compound? Second, does the frequency of the compound itself affect its naming latency? Third, if we find effects of constituent frequency, which measures of frequency and entropy are the best predictors for the compound production onset latencies?

Experiments 1-3 addressed the first question by means of factorial contrasts. For pairs of compounds matched for COMPOUND FREQUENCY and sharing one constituent, a frequency contrast on the other constituent affected the production onset latencies. Both for the head (Experiment 1) and for the modifier (Experiment 2), a higher CONSTITUENT FREQUENCY led to shorter naming onset latencies. This advantage of high frequency constituents persisted in Experiment 3, in which both constituents were of either high or of low frequency. These results allowed us to conclude that the frequencies of both constituents indeed codetermine the production latency of a compound.

Experiment 4 addressed the second question by means of a factorial contrast, matching for CONSTITUENT FREQUENCIES and contrasting COMPOUND FREQUENCIES. High frequency compounds were not produced any faster than their matched counterparts with low frequencies. In fact, there was an indication that a high COMPOUND FREQUENCY might be inhibitory, but this inhibitory effect was small and not fully significant.

In order to ascertain which frequency or entropy measures are the optimal predictors for the naming onset latencies, we analyzed the joint data of Experiments 1-4 by means of a multilevel regression analysis, which revealed that the production onset latencies were best predicted not by the constituent's CUMULATIVE STEM FREQUENCY but rather by a combination of different, partly position-specific frequency and entropy measures and compound frequency.

These results shed new light on the role of decompositionality and incrementality in production. If compounds were similar to monomorphemic words, as in full-listing models, their naming onset latencies should depend on COMPOUND FREQUENCY only. Our experiments show, however, that COMPOUND FREQUENCY plays a minor role only, leading to facilitation only for the lowest ranges of compound frequencies, and possibly to inhibition for the higher frequency ranges. The presence of an effect of COMPOUND FREQUENCY in the regression analysis shows that the position-response association task is, in fact, sensitive to word frequency, which has been demonstrated before (Cholin, 2004). Since the effects that we observed for constituent frequencies were both larger and more robust, we conclude that our data challenge models with only unstructured storage and no decomposition for complex words (see Ayala & Martin, 2002; Mondini, Luzzatti, Saletta, Allamano, & Semenza, 2005 for similar conclusions based on aphasic speakers).

The CONSTITUENT FREQUENCY effects observed for the left constituent replicate the frequency effect reported for initial constituents in (Roelofs, 1996). The frequency effects observed for the right constituent provide further support for the possibility that production onset latencies are determined not only by the first morpheme or syllable, but also by subsequent parts of the word as mentioned in (Meyer et al., 2003). The observation that frequency effects for the first constituent are more facilitatory than for the second constituent supports theories of incremental morphological processing in production. However, the effect of the second constituent argues against full incrementality. Speakers apparently plan the articulation of the first constituent with an eye on what is to be produced next. This look-ahead may also shape the details of the acoustic realization (Kemps et al., 2005a, 2005b; Koester et al., 2004).

The finding that the general frequency effects of the left and right constituents can be made more precise in terms of structural measures of constituent frequency and entropy offers new insights into the details of morphological processing in lexical access that invite further theoretical reflection. Our data suggest that the mental lexicon is highly sensitive to the specific morphophonological context in which a word has to be articulated.

The CUMULATIVE STEM FREQUENCY is a context-independent predictor of the speaker's familiarity with a given word form (e.g., *hand*), whereas position-specific measures are contextually conditioned predictors (e.g., *hand* in *handbag* or *handcuff*). This contextual sensitivity may well reflect the differences in the phonetic details of the production of *hand* by itself versus the production of *hand* as a head or a modifier. The POSITIONAL FREQUENCY effects are in line with the predictions of decompositional models with structured storage.

The POSITIONAL ENTROPY effects provide further evidence for the role of paradigmatic relations (the links between morphologically related words) in the mental lexicon (Ernestus & Baayen, 2004; Krott et al., 2001). Paradigmatic effects in lexical processing show that words are not isolated processing units, but rather structured units participating in networks of morphological relations. For instance, our POSITIONAL ENTROPY effects argue for structured storage, because the more often constituents appear in other compounds in the same structural position the faster are their production onset latencies in immediate naming.

The similarity in the magnitude of the POSITIONAL ENTROPY effects for the left and right constituents suggests that the paradigmatic effects do not differentiate between the constituent that has to be pronounced first, and the constituent that has to be pronounced last. However, from the perspective of incremental processing, simultaneous activation of the head with the modifier should be disadvantageous. In fact, there is evidence for some disadvantage associated with co-activation of the head: the inhibitory effect of the RIGHT COMPLEMENT FREQUENCY. For the initial constituent, the modifier, all measures of frequency and entropy are facilitatory, but for the final constituent, the head, the inhibitory effect of the RIGHT COMPLEMENT FREQUENCY modulates the facilitatory effects of the other measures. Apparently, selecting the target's first constituent is the harder, the more other morphologically complex non-compound words include the head constituent. Consequently, the overall frequency effect for the modifier emerged as stronger than the overall frequency effect for the head.

Considered jointly, our experiments support decompositional models of speech production, in which the paradigmatic relations entertained by the constituents of a compound with other morphologically complex words containing the constituents, as well as their structural position within those other words co-determine the details of the planning and articulation of the compound.

2.6. A reanalysis of the compound naming latencies with crossed random effects for subject and item[5]

A multilevel analysis of covariance with subject and item as crossed random effects (Bates, 2005; Bates & Sarkar, 2005; Baayen, in press) was fitted to the compound data, using a stepwise variable selection procedure.

2.6.1. Results

Inspection of the distribution of the residuals of the fitted model showed clear divergences from normality. Therefore, outliers with an absolute standardized residual exceeding 2.5 standard deviations from the mean were removed from the data (2.0 % of the correct namings within the time limit), after which the model was

[5] For a short introduction to this newly available technique, its advantages and our rational for using it, the reader is referred to section 4 of Chapter 1.

refitted. Model criticism by means of quantile-quantile plots[6] did not reveal problems with the goodness of fit of the trimmed model. In the previous regression analysis reported in 2.4, a relatively higher number of outliers was excluded (2.4 % of all correct namings within the time limit), applying the same threshold for outliers. Fewer outliers indicate a better fit between model and data.

Table 2.3: Fixed effects of the multilevel regression model with subject and stem as crossed random effects.

	Estimate	Std.Error	DF	t-value	p-value
Intercept	6.2757	0.0457	28935	137.457	0.0000
Repetition	0.0015	0.0005	28935	3.281	0.0010
Plosive	0.0239	0.0083	28935	2.885	0.0039
Left Derivational Entropy	-0.0104	0.0044	28935	-2.385	0.0171
Left CumStem Frequency	-0.0093	0.0029	28935	-3.173	0.0015
Right Positional Entropy	-0.0081	0.0044	28935	-1.819	0.0689
Compound Frequency	-0.0396	0.0146	28935	-2.711	0.0067
Compound Frequency^2	0.0046	0.0016	28935	2.811	0.0049

In what follows, I concentrate on the coefficients and associated statistics for this refitted model (see Table 2.3). Panels 1 to 6 of Figure 2.2 illustrate the partial effect of the predictors adjusted for the means of the other covariables.

Naming onset latencies increased with REPETITION (β = 0.0015, t(28935) = 3.281, p = 0.0010; see Panel 4 of Figure 2.2), and were also longer for words

[6] In a quantile-quantile plot, the quantiles of the data set are plotted against the theoretical quantiles of the standard normal distribution. If the points fall roughly along a straight line, the data set can be considered as a sample from a normal population. Systematic deviations from a line indicate length of tails and skewness of distribution, suggesting the removal of extreme outliers.

beginning with a PLOSIVE as opposed to words beginning with a non-plosive (see Panel 6).

COMPOUND FREQUENCY was a non-linear predictor with similar shape as observed in the previous analysis (2.4), with facilitation for lower and inhibition for higher frequencies (linear component: β = -0.0396, t(28935) = -2.711, p = 0.0067; nonlinear component: β = 0.0046, t(28935) = 2.811, p = 0.0049; see Panel 2).

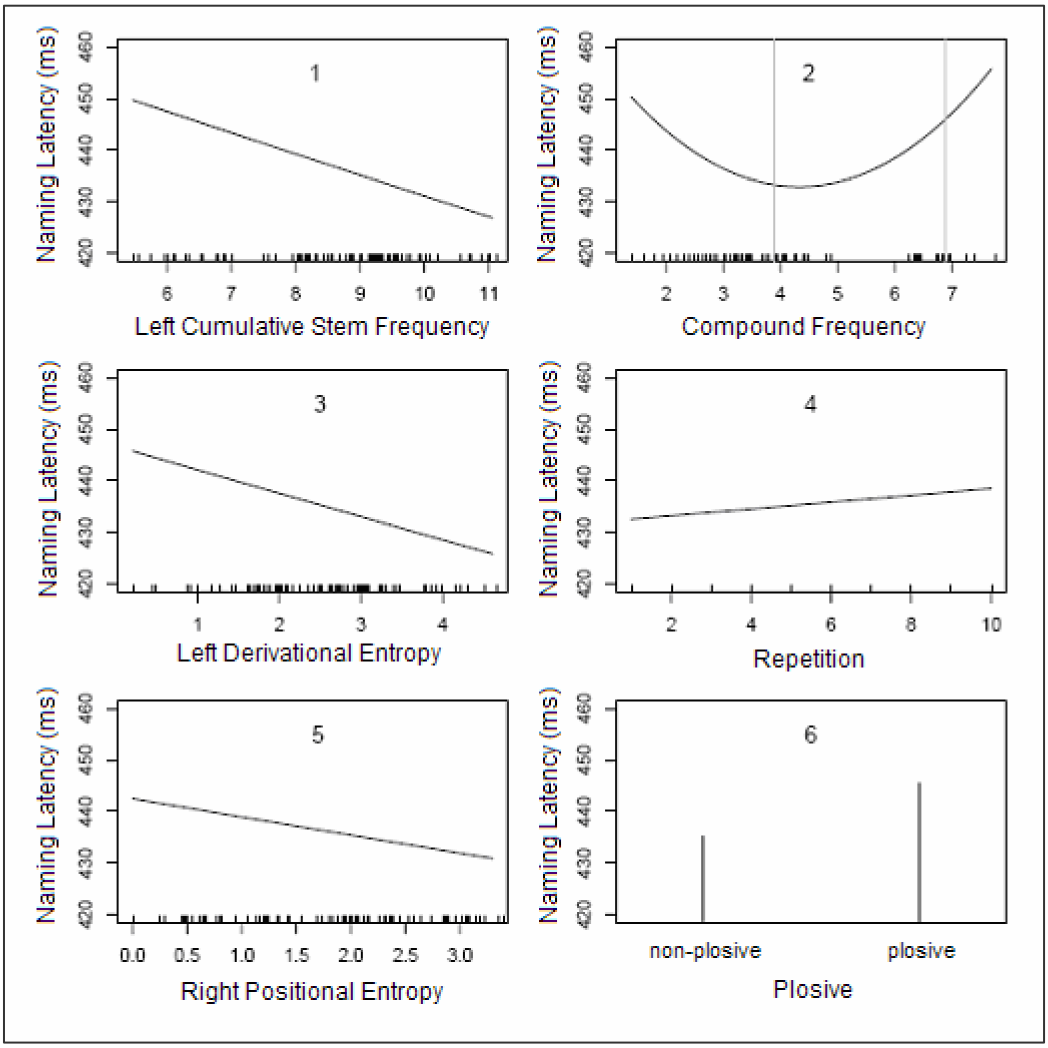

Figure 2.2: Partial effects of the predictors adjusted for the effects of the other covariables.

The cumulative stem frequency of the left constituent, summed over all its occurrences across the lexicon independent of position, revealed a facilitatory effect (β = -0.00934, t(28935) = -3.173, p = 0.0015); see Panel 1, LEFT CUMULATIVE STEM FREQUENCY). In addition, two entropy measures emerged as significant predictors:

Shannon's entropy calculated over the frequency distribution underlying the cumulative frequency count for the modifier (β = -0.0104, t(28935) = -2.385, p = 0.0171, Panel 3, LEFT DERIVATIONAL ENTROPY) and the positional entropy of the right constituent (β = -0.0081, t(28935) = -1.819, p = 0.0689, see Panel 5, RIGHT POSITIONAL ENTROPY), i.e., Shannon's entropy estimated for the frequency distribution of the compounds in which the right constituent occurs as right constituent. None of the other predictors of the earlier model were significant.

The model incorporated three random effects: random intercepts for word and for subject, and the residual error. The standard deviations of these random effects were 0.036 for word, 0.164 for subject, and 0.226 for the residual error. The inclusion of word as random effect in the model was strongly supported by a likelihood ratio test[7] ($Chi^2(1)$ = 446.15, p < 0.0001). The model with both subject and word as crossed random effects captures a substantial by-item variation that is not captured by the item-specific predictors in a model with subject as random effect only.

2.6.2. Discussion

When comparing this new model to our previous model, it is evident that the number of significant predictors is reduced from nine to six. The total number of parameters is reduced by two, because the new model contains three fixed effects less, but an extra random effect for word. Even though the more conservative model reports fewer significant predictors, it explains a higher proportion of variance (r-squared: 0.355 (new model) versus 0.341 (previous model). Following Occam's razor, the new, parsimonious model clearly is to be preferred.

Taking a closer look at the significant predictors of the two models, four variables turn out to be significant predictors in both models. Next to these four shared variables, the new model proposes significant effects for two new variables, replacing a total of five variables of the previous model.

[7] The likelihood ratio test compares a relatively more complex model to a simpler model to see if it fits a particular dataset significantly better. The more complex model must differ from the simple model only by the addition of parameters. Although adding parameters will always result in a higher likelihood score, there comes a point when it is no longer justified in terms of significant improvement.

As before, PLOSIVE-initial words evoke significantly longer naming onset latencies compared to non-plosive initial words. There still is a non-linear effect of COMPOUND FREQUENCY, with shortest naming onset latencies for words of medium frequencies. An effect of NEIGHBORHOOD DENSITY (nonlinear in the previous model) is no longer present.

Instead of three frequency measures for the left constituent (LEFT DERIVATIONAL ENTROPY, LEFT POSITIONAL ENTROPY, and LEFT COMPLEMENT FREQUENCY), the new model reports two (LEFT DERIVATIONAL ENTROPY and LEFT CUMULATIVE STEM FREQUENCY). Instead of three measures for the right constituent (RIGHT POSITIONAL ENTROPY, RIGHT COMPLEMENT FREQUENCY and RIGHT LEMMA FREQUENCY), the new model reports one (RIGHT POSITIONAL ENTROPY). All constituent-specific effects in the new model are facilitatory.

In the new analysis, I also checked for effects of position-specific neighborhoods and cohort entropies (see Chapter 1), two variables that we had not yet included in the previous analysis. The model reveals no significant effects of position-specific neighborhoods or cohort entropies.

While the new model draws a somewhat simpler picture of the variables which affect the production onset latencies of morphologically complex words, it nevertheless supports the main conclusions drawn from the previous analysis. The model argues against both full-storage and full-decomposition and supports the relevance of paradigmatic structure for speech production.

In a full-storage model, the frequency of the compound should be the only frequency that affects its naming latency. The data clearly show that this is not the case as we find constituent-specific effects of frequency and entropy.

In a full-decomposition model of speech production, the naming latency of morphologically complex words should be determined by constituent frequency, and perhaps constituent entropy, only. A model that assumes both decomposition and incremental processing (Levelt et al., 1999) predicts a frequency effect of the first constituent only. Fully decompositional models are challenged by the finding that the frequency of a compound codetermines its production latency.

In addition, full-storage and full-decomposition approaches are likewise challenged by the following findings, both of which suggest a role for paradigmatic structures in the mental lexicon and an influence of these paradigmatic structures on word production onset latencies. First, the positional entropy of the right constituent (RIGHT POSITIONAL ENTROPY) suggests context-sensitive processing of a compound's head. Second, the entropy effect of the modifier (LEFT DERIVATIONAL ENTROPY), which is significant beyond its frequency effect, demonstrates that all words are involved in the production of one word.

To summarize, the more conservative model with subject and word as crossed random effects explains more variance using fewer predictors. Drawing a somewhat simpler picture, the new analysis supports the main conclusions of the previous analysis, challenging models of full storage as well as models of full decomposition and suggesting that paradigmatic structure affected the production onset latencies of these transparent Dutch noun-noun compounds.

I want to conclude this chapter with a word of caution with respect to drawing conclusions about the significance and non-significance of specific paradigmatic variables. Some of our variables have not yet been used in prior studies of speech production and for other variables, our studies are among the first to include them. Replications are required, also for the following reason. Some of the paradigmatic variables show relatively high correlations. A non-significant predictor is not necessarily of zero predictivity and may even be a candidate to replace a significant predictor in a replication study. Therefore, I would like to see the effects of the present study replicated and focus at this stage on the more general conclusion that paradigmatic structure plays a role in the production of transparent noun-noun compounds in Dutch.

References

Ayala, J., & Martin, N. (2002). Decompositional effects in the production of compound words in aphasia. *Brain and Language, 83,* 81-83.

Baayen, R.H. (2004). Statistics in psycholinguistics: A critique of some current gold standards. *Mental Lexicon Working Papers*, 1-45.

Baayen, R.H. (2007). *Analyzing linguistic data. A practical introduction to statistics*. Cambridge University Press, in press.

Baayen, R.H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database* (CD-ROM), Linguistic data consortium. University of Philadelphia, Philadelphia, PA, Release 2.

Baayen, R.H., Tweedie, F.J., & Schreuder, R. (2002). The subjects as a simple random effect fallacy: Subject variability and morphological family effects in the mental lexicon, *Brain and Language*, *81*, 55–65.

Bates, D.M., & Sarkar, D. (2005). lme4: Linear mixed-effects models using S4 classes, R package version 0.9975-7.

Bates, D.M. (2005). Fitting linear mixed models in R, *R News*, *5*, 27-30.

Bates, E., D'Amico, S., Jacobsen, T., Szekely, A., Andonova, E., Devescovi, A., Herron, D., Lu, C.C., Pechmann, T., Pleh, C., Wicha, N., Federmeier, K., Gerdjikova, I., Gutierrez, G., Hung, D., Hsu, J., Iyer, G., Kohnert, K., Mehotcheva, T., Orozco-Figueroa, A., Tzeng, A., & Tzeng, O. (2003). Timed picture naming in seven languages. *Psychonomic Bulletin & Review*, *10(2)*, 344-380.

Burani, C., & Caramazza, A. (1987). Representation and processing of derived words. *Language & Cognitive Processes*, *2*, 217-227.

Caramazza, A., Costa, A., Miozzo, M., & Bi, Y. (2001). The specific-word frequency effect: Implications for the representation of homophones in speech production. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *27*, 1430-1450.

Cholin, J. (2004). The syllable in speech production. Effects of syllable preparation and syllable frequency. Ph.D. dissertation, Max Planck Institute series in psycholinguistics, 26, Radboud University Nijmegen, the Netherlands.

Cholin, J., Levelt, W.J.M., & Schiller, N.O. (2006). Effects of syllable frequency in speech production. *Cognition*, *99*, 205-235.

Cholin, J., Schiller, N.O., & Levelt, W.J.M. (2004). The preparation of syllables in speech production. *Journal of Memory and Language, 50*, 47-61.

Coltheart, M., Davelaar, E., Jonasson, J.T., & Besner, D (1977). Access to the internal lexicon. In Dornic, S. (Ed.), Attention and Performance VI. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Ernestus, M., & Baayen, R.H. (2004). Analogical effects in regular past tense production in Dutch. *Linguistics*, *42*, 873-903.

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

Jescheniak, J.D., Meyer, A.S., & Levelt, W.J.M. (2003). Specific word frequency is not all that counts in speech production. Reply to Caramazza et al. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *29*, 432-438.

Kemps, R., Ernestus, M., Schreuder, R., & Baayen, R. H. (2005). Prosodic cues for morphological complexity: The case of Dutch plural nouns. *Memory and Cognition*, *33*, 430-446.

Kemps, R., Wurm, L., Ernestus, M., Schreuder, R., & Baayen, R.H. (2005). Prosodic cues for morphological complexity in Dutch and English. *Language and Cognitive Processes, 20*, 43-73.

Köster, D., Gunter, T.C., Wagner, S., & Friederici, A.D. (2004). Morphosyntax, prosody, and linking elements: The auditory processing of German nominal compounds. *Journal of Cognitive Neuroscience, 16*, 1647-1668.

Krott, A., Baayen, R.H., & Schreuder, R. (2001). Analogy in morphology: modeling the choice of linking morphemes in Dutch. *Linguistics, 39(1)*, 51-93.

Laudanna, A., & Burani, C. (1985). Address mechanisms to decomposed lexical entries. *Linguistics, 23,* 775-792.

Levelt, W.J.M. (2001). Spoken word production: A theory of lexical access. *Proceedings of the National Academy of Sciences*, *98(23)*, 13464-513.

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

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

Meyer, A.S. (1990). The time course of phonological encoding in language production: The encoding of successive syllables of a word. *Journal of Memory and Language, 29,* 524-545.

Meyer, A.S. (1991). The time course of phonological encoding in language production: The encoding inside a syllable. *Journal of Memory and Language, 30,* 69-89.

Meyer, A. S., Roelofs, A., & Levelt, W.J.M. (2003). Word length effects in object naming: The role of a response criterion. *Journal of Memory and Language*, *48*, 131-147.

Mondini, S., Luzzatti, C. Saletta, P., Allamano, N., & Semenza, C. (2005). Mental representation of prepositional compounds: Evidence from Italian agrammatic patients. *Brain and Language*, *94*,178-187.

Oldfield R.C., & Wingfield A. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology*, *17*, 273-281.

Pinheiro, J.C., & Bates, D.M. (2000). Mixed-effects models in S and S-PLUS, *Statistics and Computing Series*, Springer-Verlag, New York, NY.

Quené, H., & Van den Bergh, H. (2004). On multi-level modeling of data from repeated measures designs: A tutorial, *Speech Communication*, *43 (1-2)*, 103-121.

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

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

Schreuder, R., & Baayen, R.H. (1997). How complex simplex words can be. *Journal of Memory and Language*, *37*, 118–139.

Shannon, C. E., & Weaver, W. (1949). *The mathematical theory of communication*. University of Illinois Press, Urbana.

Shannon, C.E. (1948). A mathematical theory of communication, *The Bell System Technical Journal*, *27*, 379-423.

Vitevitch, M.S. (2002). The influence of phonological similarity neighborhoods on speech production. *Journal of Experimental Psychology: Learning, Memory and Cognition, 28,* 735-747.

Appendix: Compounds

For each experiment, the 16 pairs of compounds are listed in 16 cells with the compound of the low-frequency condition on top of the compound of the high-frequency condition. The labels (low frequency, L1–L8; high frequency, H1–H8) assign each compound to one of the frequency–intern repairings (e.g. Exp. 1: L1, marktkraam–slagpin) as presented in the task.

Exp. 1	Exp. 2	Exp. 3	Exp. 4
(L1) marktkraam (H1) marktvrouw	(L1) zeepkist (H2) geldkist	(L1) schildklier (H1) deelstaat	(L1) deelstaat (H2) vraagstuk
(L1) slagpin (H8) slagkracht	(L1) strozak (H7) zandzak	(L1) dwangsom (H1) straatweg	(L1) straatweg (H7) zeeman
(L2) bedstee (H2) bedrust	(L2) rijstveld (H3) krachtveld	(L2) strohoed (H2) wijnglas	(L2) wijnglas (H4) voetstap
(L2) aardkluit (H5) aardschok	(L2) prachtstuk (H8) kunststuk	(L2) viltstift (H3) keelgat	(L2) bankstel (H5) zakdoek
(L3) hoofdtooi (H4) hoofdtaak	(L3) bronstijd (H1) steentijd	(L3) lintworm (H3) plaatsnaam	(L3) plaatsnaam (H3) dagboek
(L3) eindsprint (H6) eindstrijd	(L3) darmwand (H4) glaswand	(L3) mestvaalt (H2) bankstel	(L3) keelgat (H8) tandarts
(L4) regenscherm (H1) regentijd	(L4) halsband (H3) stemband	(L4) vloedlijn (H4) huisvriend	(L4) huisvriend (H2) hoofdstad
(L4) appelmoes (H3) appeltaart	(L4) prooidier (H8) lastdier	(L4) kruitdamp (H4) slagzin	(L4) slagzin (H3) zonlicht
(L5) postgiro (H3) postmeester	(L5) kleilaag (H1) luchtlaag	(L5) poolkap (H6) landmacht	(L5) kernbom (H6) vakbond
(L5) huissloof (H8) huisman	(L5) rumboon (H6) tuinboon	(L5) maiskolf (H5) ijsbaan	(L5) ijsbaan (H8) noodlot
(L6) kruisspin (H7) kruisvuur	(L6) kropsla (H7) veldsla	(L6) hooimijt (H7) jaarbeurs	(L6) landmacht (H7) grondwet
(L6) bloemtros (H4) bloemvorm	(L6) vilthoed (H4) punthoed	(L6) muilkorf (H8) geldprijs	(L6) koprol (H1) standpunt
(L7) voetveeg (H5) voetsteun	(L7) tolweg (H6) ringweg	(L7) roomsaus (H7) ringslang	(L7) ringslang (H5) maatstaf
(L7) grasspriet (H2) grasland	(L7) roomsaus (H2) wijnsaus	(L7) windhoos (H8) nachtploeg	(L7) jaarbeurs (H1) tijdschrift
(L8) broodkruim (H7) broodheer	(L8) loonsom (H5) hoofdsom	(L8) kropsla (H5) kernbom	(L8) nachtploeg (H4) rechtspraak
(L8) luchtbuks (H6) luchtbrug	(L8) grasmat (H5) deurmat	(L8) rumboon (H6) koprol	(L8) geldprijs (H6) weekblad

FREQUENCY EFFECTS IN THE PRODUCTION OF DUTCH DEVERBAL ADJECTIVES

3.1. Introduction

In this chapter, I study the role of frequency information in the production of a second type of morphologically complex words: derivations. In particular, I look at one subcategory of derivational morphology: deverbal adjectives. In derivational morphology, new words (lexemes) are formed through affixation. Affixation, the attachment of bound[1] morphemes to words, is the formal operation subserving both inflection and derivation, but it is only in derivation that the outcome is a new lexeme. In contrast, the outcome of inflection is a syntactically appropriate variant of the same lexeme. The derived word (the output word) can be of the same or of a different word class (noun, adjective, verb) than its base word (the input word). Booij (2002) presents examples for nine input-output word class relations in Dutch. As Figure 3.1 illustrates, input words of each word class can be derived to output words of each class. It follows straight-forwardly, that in six of these nine input-output combinations, the word class is changed.

All examples presented in Figure 3.1 involve the attaching of a suffix (an affix attaching to the end of a word stem). Dutch derivational affixation is, however, not restricted to suffixation. Five of the nine input-output combinations can be also formed through the attaching of a prefix (an affix attaching to the beginning of a word stem), such as in *ver-slaaf* (to addict to). In what follows, we will take a closer look at the case of deverbal adjectives (e.g., *lees-baar,* readable).

[1] Bound (as opposed to free) morphemes do not exist as independent words.

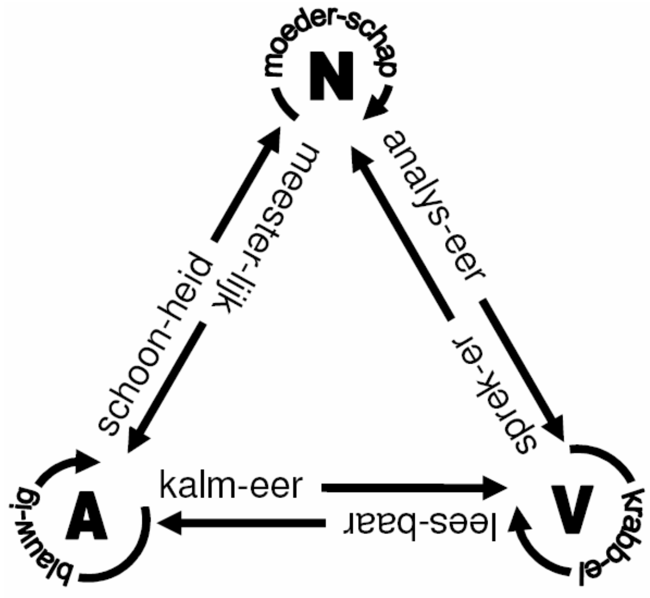

Figure 3.1: In Dutch derivational morphology, nine input-output word class combinations are found (N for noun, A for adjective, V for verb). The examples are taken from Booij (2002).

3.1.1. Deverbal adjectives

A Dutch deverbal adjective is formed by the attachment of an adjectival suffix (not prefix) to a verbal stem (see Table 3.1). As a result of the attachment, the word class is changed from verb to adjective, while the particular change in meaning depends on the particular suffix.

Table 3.1: Examples for the formation of deverbal adjectives in Dutch: A case of word class change during derivational affixation.

verb:	grijpen *(to grab)*	werken *(to work)*	schrapen *(to scratch)*
verb stem:	grijp	werk	schraap
adjectival suffix:	-baar	-zaam	-erig
deverbal adjective:	grijpbaar *(tangible)*	werkzaam *(effective)*	schraperig *(avaricious)*

In Dutch, the following nine adjectival suffixes form deverbal adjectives: *-achtig*, *-baar*, *-erig*, *-elijk*, *-ig*, *-lijk*, *-loos*, *-s*, and *-zaam*. It is not clear whether the suffixes -

elijk and *-lijk* should be considered as two different suffixes or not. They might very well be two phonological versions of the same underlying suffix, thereby reducing the number of deverbal adjectival suffixes to eight. However, *-elijk* and *-lijk* differ in the number of syllables. I, therefore, list *-elijk* and *-lijk* separately and treat them as different suffixes in the analyses.

As described above, the attaching of an adjectival suffix to a verbal stem always results in an adjective. However, these adjectival suffixes are not interchangeable. There are restrictions as to which adjectival suffix can be attached to which type of verb (for example transitive vs. intransitive verbs). Furthermore, each adjectival suffix brings along its own meaning as Table 3.2 illustrates.

Table 3.2: The attachment of a specific adjectival suffix determines the meaning of the resulting adjective.

verb:	*werken* (to work)			
verb stem:	*werk*			
adjectival suffix:	*-zaam*	*-loos*	*-baar*	*-elijk*
deverbal adjective:	*werkzaam* (effective)	*werkloos* (unemployed)	*werkbaar* (useful)	*werkelijk* (truly)

The attachment of a particular adjectival suffix to different verb stems, changes their meaning in the same way: *Werkzaam* (laborious) relates to *werken* (to work) as *voedzaam* (nutritious) relates to *voeden* (to feed), *grijpbaar* (touchable) relates to *grijpen* (to grab) as *drinkbaar* (drinkable) relates to *drinken* (to drink), and so on. Therefore, it is even possible to attach adjectival suffixes to new verbs or verbs of other languages, and to efficiently communicate the intended change in meaning. Any Dutch proficient listener will interpret an unknown word that ends in *–baar* (e.g., *downloadbaar*) as an adjective expressing ‘something is able to be V-ed’ (here: ‘something is able to be *download*-ed’, even if the listener does not know the exact meaning of the verb *to download*).

3.1.2. Research question

This doctoral dissertation studies the role of frequency information in the production of morphologically complex words in Dutch. In the previous chapter, I have looked at compounds, in particular at noun-noun compounds, which are morphologically complex words that consist of two free morphemes, each of which can in principal be head or modifier of a compound. It is only in the compound, that one (the initial) constituent becomes the modifier and the other (the final) becomes the head. In principal, any noun can take any position in a compound. There are even pairs of compounds, in which modifier and head are switched (e.g., *vraagprijs* - *prijsvraag*, asking price – price contest). Crucially, in contrast to noun-noun compounds, the constituting morphemes of derived words can not switch place. While the verbal stem is a free morpheme, the adjectival suffix is a bound morpheme that attaches to the end of the verbal stem, never to its beginning.

Factorial analyses of the compound naming onset latencies revealed frequency effects of both constituents but not of the compound itself, suggesting composition during production. Stepwise regression analyses of the joint data of the compound experiments revealed a superior predictivity of structural frequency and entropy measures, challenging full decomposition and suggesting a role for paradigmatic structure in speech production. I am curious to see, to what extent these findings replicate using a different type of morphologically complex words, deverbal adjectives.

With the deverbal adjectives, I do not run any factorial experiments, but immediately analyze the naming onset latencies of a wide range of deverbal adjectives with stepwise regression modeling. As with the compounds (Chapter 2), the main question is whether the naming latency of a morphologically complex word (here a deverbal adjective, e.g., *grijpbaar*) can be predicted by its own frequency of occurrence or by the frequency of its first or / and second constituent (e.g., *grijp*, *baar*). Other than with the compounds, however, it is impossible to find pairs of deverbal adjectives that are matched in two of these frequencies while having a contrast in the third frequency. Thus, no individual factorial experiments are run, which address the effect of one frequency variable at a time. The constituent and whole-word frequency variables are rather included along with other predictors in a stepwise multilevel analysis of covariance.

3.1.3. Predictors

As introduced in Chapter 1, variables from four groups are included in the analyses: control variables, frequency variables, morphological variables, and phonological variables. All variables were introduced and defined in Chapter 1, where I also described in more detail the motivation to include each variable.

In stepwise covariance modeling, a number of variables can be used simultaneously to predict the response latencies in the experiment. Variables that add no or very little predictive value, can be taken out step by step, thereby reducing the number of variables until a small set of good predictors is left. The final model will be the best trade-off between the two goals of trying to explain as much variance as possible while using as few predictors as possible. Sometimes a variable is taken out even though it has predictive value. This is the case whenever another variable is present that is strongly correlated with and a better predictor than the first. In the presence of the stronger, correlated predictor, the first predictor simply does not add sufficient information to secure its place in the model. Therefore, only the stronger predictor will appear in the final model of covariance. In such cases, discussing the remaining variables with an eye on the excluded ones can be very informative and enhance the general understanding of the model.

As control variables, I include REPETITION (see task description in 3.2.3.) and PLOSIVE, both of which helped to explain a significant proportion of variance in the previous analyses. Furthermore, the control variable VOICED is included.

Next to the above mentioned variables FREQUENCY OF THE DEVERBAL ADJECTIVE (e.g., *grijpbaar*) CUMULATIVE STEM FREQUENCY (e.g., *grijp*), and FREQUENCY OF THE SUFFIX (e.g., *baar*), I take along the POSITIONAL FREQUENCY OF THE STEM (e.g., *grijp* as initial constituent of a morphologically complex word). There was a positional frequency effects with the noun-noun compounds (e.g., *grijp* as initial constituent).

Furthermore, the morphological variable DERIVATIONAL ENTROPY is included, which refers to the token-weighted count of the numbers of types in which the stem (e.g., *grijp*) occurs as a constituent. The DERIVATIONAL ENTROPY is a measure of the amount of information carried by the stem's derivational paradigm.

As phonologically related predictors, I take along the PHONOLOGICAL WORD LENGTH, which is the number of phonemes of the deverbal adjective. Longer words tend to elicit longer naming onset latencies (e.g., Meyer, Roelofs, & Levelt, 2003).

Another phonological variable is the number of phonological neighbors, the NEIGHBORHOOD DENSITY (e.g., Vitevitch, 2002, 2006). Phonological neighbors are words that can be transformed into one another by exchanging only one phoneme (Greenberg & Jenkins, 1964; Coltheart, Davelaar, Jonasson, & Besner, 1977). NEIGHBORHOOD DENSITY effects in production are a curious finding, as they suggest a simultaneous activation of several word forms as discussed in the first Chapter. I have myself found a NEIGHBORHOOD DENSITY effect in the compound production onset latencies in the analysis presented in section 4 of Chapter 2, but not in the more conservative analysis presented in section 6 later in that chapter. Therefore, I am curious to see whether an effect of NEIGHBORHOOD DENSITY shows up in the production of deverbal adjectives.

Next to the total number of phonological neighbors, I take into account the number of neighbors exchanging the first, second, third, etc. phonemes separately (e.g., Sevald & Dell, 1994, for initial neighbors), because words are produced over time from the initial to the final phoneme. As a consequence, neighborhoods at different positions might enter the production process at different points in time. These POSITION-SPECIFIC NEIGHBORHOODS (N1, N2, N3, etc.) add up to (and are therefore correlated with) the total number of phonological neighbors of a word. The position-specific neighborhoods quickly decrease from initial to final phoneme, and many of the constituting morphemes in the material have zero neighbors at the fourth phoneme. Hence, only the POSITION-SPECIFIC NEIGHBORHOOD DENSITIES N1, N2 and N3 are included.

Based on a similar rationale, I include the COHORT ENTROPIES, which represent the specific amount of information that is carried by each additional phoneme (Van Son & Pols, 2003). COHORT ENTROPIES are the entropies estimated for the probability distributions of all words beginning with the same first (H1), first two (H2), first three (H3) phonemes.

3.2. Material and Method

3.2.1. Material selection

From the CELEX lexical database (Baayen, Piepenbrock, & Guliker, 1995, CD-rom), 124 Dutch deverbal adjectives (e.g., *drinkbaar* (drinkable)) were selected. The selection procedure was based on five criteria. First, all the nine suffixes forming deverbal adjectives in Dutch (*-achtig*, *-baar*, *-elijk*, *-erig*, *-ig*, *-lijk*, *-loos*, *-s*, *-zaam*) should be represented in the material. Second, each verbal stem should occur only once in the experiment. Third, the CELEX frequencies of both the deverbal adjectives and their verbal stems were required to be greater than zero. All CELEX frequencies reported here and below are counts based on a corpus of 42 million words. Fourth, no item should show unreasonable divergence in its CELEX frequency, its Google frequency and an average familiarity rating[2] based on 27 Dutch participants. Finally, selection was done in such a way that both the deverbal adjectives and their verbal stems were fairly distributed over a wide range of frequencies.

Table 3.4 lists the absolute and relative CELEX-frequencies of each deverbal suffix as well as their number and percentage of items used in the experiment. Figure 3.2 presents the frequency distributions of both the lemma frequencies of the deverbal adjectives (a) and the cumulative verbal stem frequencies (b) for the selected set of items.

Equal-sized groups of items per suffix were unfeasible, given both the selection criteria and the huge variation in frequency of occurrence of the deverbal suffixes. Compared to their CELEX token-frequencies, lower-frequency suffixes are overrepresented and higher-frequency suffixes are underrepresented in the

[2] On a 7 point scale, the mean rating for the selected material was 4.3, with a standard deviation of 2.0.

material. All selected items are listed in the Appendices A (grouped by suffixes) and B (assigned to item sets).

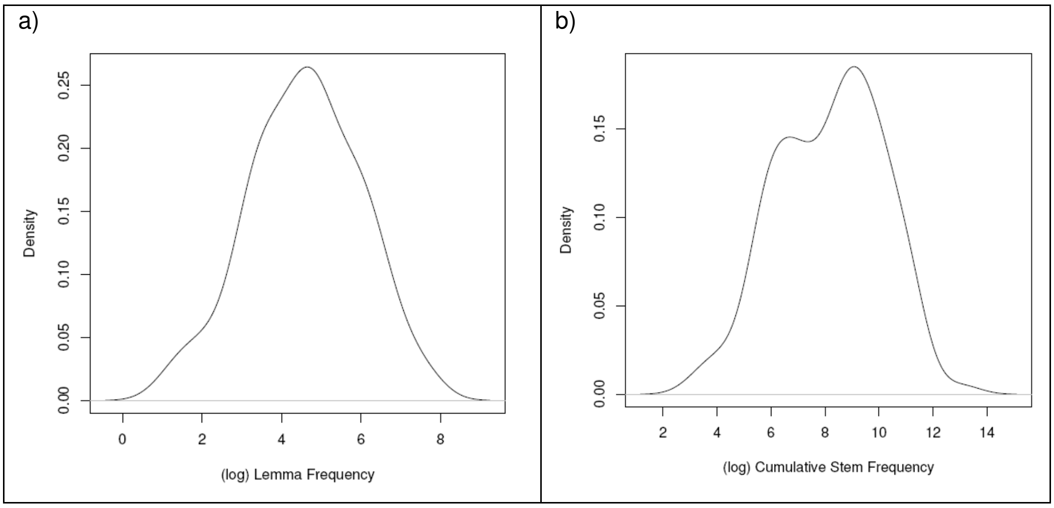

Figure 3.2: The selected deverbal adjectives were fairly distributed over a wide range of both the lemma frequencies of the deverbal adjectives (a) and cumulative stem frequencies (b).

Table 3.4: The nine suffixes forming deverbal adjectives in Dutch, their occurrence in the CELEX lexical database and in the experiment.

Suffix		***CELEX***				***Experiment***	
		Frequency		**%**		**Number**	**%**
		Token	**Type**	**Token**	**Type**	**of items**	**of items**
(1)	**-achtig**	9316	251	2	15	10	8
(2)	**-baar**	19869	185	4	11	33	27
(3)	**-elijk**	148443	228	31	13	13	11
(4)	**-erig**	5054	130	1	8	32	25
(5)	**-ig**	134216	529	28	31	11	9
(6)	**-lijk**	112976	136	24	8	4	3
(7)	**-loos**	11532	108	2	6	2	2
(8)	**-s**	23593	118	5	7	4	3
(9)	**-zaam**	7711	32	2	2	15	12
Total		**472710**	**1717**	**100%**	**100%**	**124**	**100%**

3.2.2. Experimental lists

The selected adjectives were arranged into 61 sets of two (e.g., *drinkbaar - buigzaam*, (drinkable – pliable)) with the constraint that the two adjectives within a set had minimal phonological overlap and no obvious semantic relation. Based on this first basic list of the 61 sets, I created three additional basic lists, balancing the order[3] of adjectives within the sets. The second basic list was created by exchanging the order within all 61 sets (e.g., *buigzaam - drinkbaar*, (pliable – drinkable)). To form the third (fourth) basic list, I exchanged the order of the adjectives within the first (last) 31 sets only, leaving the last (first) 30 sets as they were.

For each of the four basic lists, I then constructed 6 randomizations of the order of sets, creating a total of 24 different experimental lists. As the experiment had to be divided over two sessions, each experimental list was finally cut into two parts, one containing 31 sets, the other containing 30 sets. The sets of target adjectives within each part were preceded by three practice sets, which contained adjectives of similar structure and frequency.

3.2.3. Position-Response Association Task

Participants were tested individually in a dimly lit sound-attenuated booth. They were comfortably seated in front of a CRT computer screen, a Sennheiser microphone and button box and they were wearing headphones. On average, a single session lasted 70 min. Participants, who wanted to complete both parts on one day had to take a minimal break of 90 min in between the two sessions.

I used a position-response association task (Cholin, Levelt, & Schiller, 2006), in which participants learned to associate the two adjectives in a set with a position mark on the left or right side of the computer screen. For each set, the

---

[3] The order of the two adjectives within a set was important for the task (see 3.x.). The first target was presented *first* and had to be associated with an icon on the *left* side of the screen, while the second target was presented *second* and associated with an icon on the *right* side of the screen. This difference in presentation might have an effect on task performance. By balancing out the order of adjectives between lists (that is between subjects), such effects should be cancelled out in the overall analyses of naming latencies.

experimental procedure consisted of a learning phase and a test phase. Both phases were introduced by an attention signal presented on the screen for 2 seconds, and ended with a pause signal that remained on the screen until the participant initiated the following phase.

In the learning phase, participants were presented with the two adjectives over headphones. Simultaneously with hearing the first adjective, they saw the icon of a loudspeaker appearing on the left side of the screen. Simultaneously with hearing the second adjective, the same icon appeared on the right side of the screen. This procedure was repeated once. As a result, the participants established associations between the icon on the left (right) side of the screen and the first (second) adjective.

In the immediately following test phase, participants were repeatedly presented with the left or right icon as a prompt to name the associated adjective. Prompting was pseudo-randomized with maximally 4 consecutive repetitions of the same target. Each adjective was prompted ten times. I included distractor trials to make it difficult for participants to prepare one of the target words. In a distractor trial, participants named a single-digit number (1, 2, 3, or 6) which was presented in the center of the screen. In total, 20 distractor trials alternated with 20 experimental trials. The participants were instructed to name each target as quickly and correctly as possible and I measured the naming onset latencies as following. Simultaneously with the presentation of a prompt the voice key was activated for 1500 ms. Naming onset latencies longer than 1500 ms were counted as time-outs. I monitored the participant's responses through headphones and took notes of incorrect naming, hesitations and voice key errors.

3.2.4. Participants

From the subject pool of the Max Planck Institute for Psycholinguistics 24 native speakers of Dutch (20 females, 4 males) were recruited. A total of 17 participants completed both sessions within the same day, 7 participants took part on two different days. Each participant received a total of € 15 for completing both sessions.

3.3. Results

Due to computer problems, the onset latencies of one participant were not recorded correctly and excluded from the analyses. Furthermore, the list of selected items, mistakenly, contained two phonological neighbors (*zweterig* and *zweverig*). On top of that, the wav-file for *zweverig* was, accidentally, presented also in the learning phase of *zweterig*, collecting twice as many naming onset latencies for the former item and no latencies for the latter item. We, therefore, excluded both items from the analysis. Of the remaining 28060 experimental trials (23 participants producing 122 items, each ten times), a total of 959 (3%) time out trials (onset latencies >1500 ms), hesitations, wrong namings and voice key errors was removed prior to analyses.

The naming onset latencies were analyzed via a mixed-effects regression analysis with subject and item as crossed random effects (Pinheiro & Bates, 2000; Baayen, Tweedie, & Schreuder, 2002; Bates, 2005; Bates & Sarkar, 2005; Baayen, in press). Following a stepwise variable selection procedure, model criticism led to the removal of 2% data points with absolute standardized residuals exceeding 2.5 standard deviations from the mean.

Table 3.5: Fixed effects of the multilevel regression model with subject and stem as crossed random effects.

		Estimate	Std.Error	DF	t.value	p.value
INTERCEPT		6.07770	0.03288	26500	184.870	0.00000
REPETITION		0.00515	0.00049	26500	10.459	0.00000
VOICED	**unvoiced**	-0.01875	0.00869	26500	-2.159	0.03086
PLOSIVE	**other**	-0.00662	0.01156	26500	-0.572	0.56733
	plosive	0.02241	0.00837	26500	2.677	0.00743
N1		0.00251	0.00112	26500	2.246	0.02471
H2		-0.00550	0.00238	26500	-2.314	0.02068
CUM. STEM FREQUENCY		-0.00399	0.00191	26500	-2.089	0.03672

The final model incorporated three random effects: random intercepts for stem (STD = 0.034) and for subject (STD = 0.12), and the residual error (STD = 0.22). Table 3.5 summarizes the fixed-effect structure of the final model, including beta weights, standard errors, t-values and significance levels. Panels 1 to 6 of Figure 3.3 illustrate the partial effects of each predictor, adjusted for the effects of the other covariates at their medians.

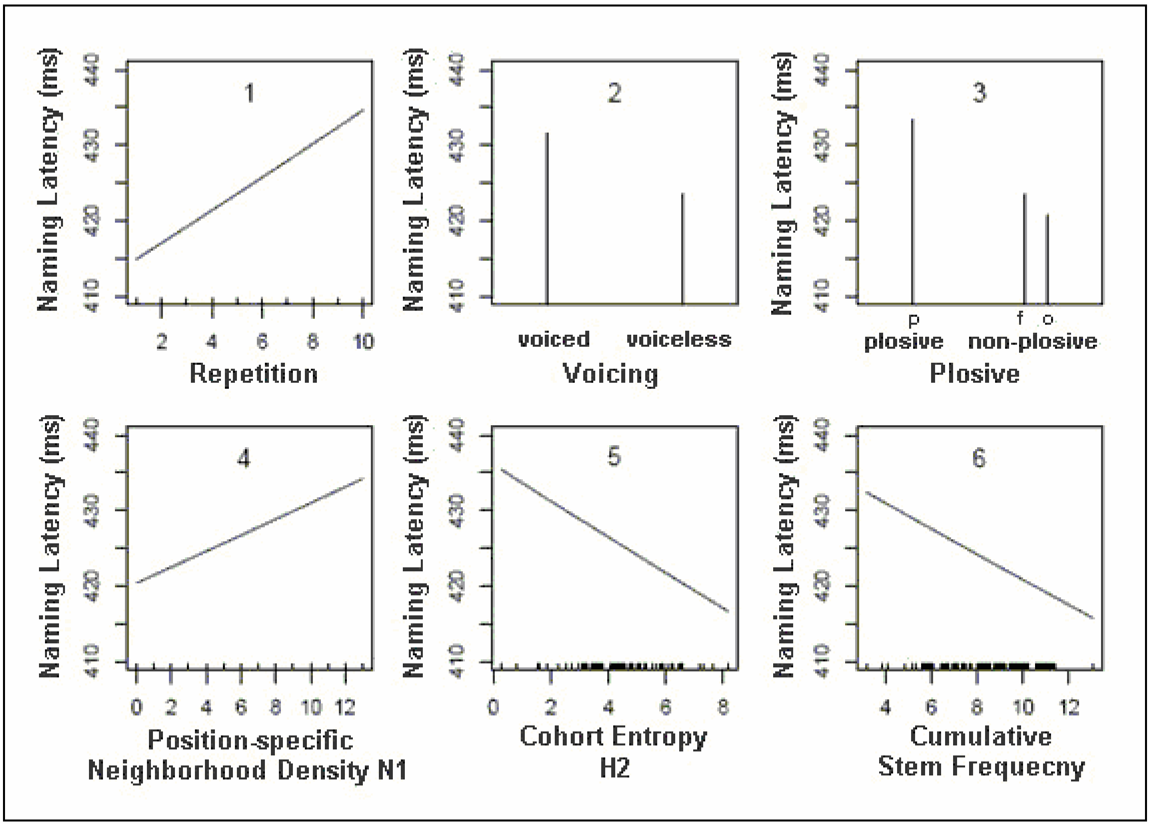

Figure 3.3: Deverbal adjectives: Partial effects of the predictors adjusted for the effects of the other covariables (panel 3: p=plosives, f=fricatives, o=other initial phonemes).

As in all experiments reported in the previous chapters, the participants started relatively fast within the test phases and slowed down towards their ends. The inhibitory effect of the control variable REPETITION ($\beta = 0.00515$, $t(26500) = 10.459$, $p < 0.0001$) is shown in panel 1 of Figure 3.3. Words with unvoiced initial segments were named faster than words with voiced initial segments ($\beta = -0.01875$, $t(26500) = -2.159$, $p = 0.03086$, panel 2, for VOICE) and PLOSIVE-initial words elicited longer onset latencies than words beginning with non-plosives (F(2, 26500) = 5.7625, p = 0.00315, for plosive, illustrated in panel 3, where label ‘f’ denotes fricatives and label ‘o’ denotes other initial phonemes).

Among the phonological variables, there was an inhibitory effect for N1, the POSITION-SPECIFIC NEIGHBORHOOD DENSITY of the initial phoneme (β = 0.00251, t(26500) = 2.246, p = 0.02471, panel 4). No other POSITION-SPECIFIC NEIGHBORHOOD DENSITY was significant, nor was there an effect of the overall NEIGHBORHOOD DENSITY. Panel 5 shows the facilitatory effect of the COHORT ENTROPY for the second phoneme (β = -0.00550, t(26500) = -2.314, p = 0.02068). No other COHORT ENTROPY was significant.

The model further reveals one predictor among the frequency variables: a facilitatory, linear effect of the CUMULATIVE STEM FREQUENCY (β = -0.00399, t(26500) = -2.089, p = 0.03672). The more often the verbal stem occurs anywhere in the lexicon, individually or as part of any morphologically complex word, the faster the deverbal adjective is named (Panel 6).

3.4. Discussion

The experiment collected a total of 29280 naming onset latencies of 124 Dutch deverbal adjectives. The statistical model that predicts these naming latencies best includes three linguistic variables (the POSITION-SPECIFIC NEIGHBORHOOD DENSITY N1, the COHORT ENTROPY H2 and the CUMULATIVE FREQUENCY OF THE STEM) next to three control variables (REPETITION, PLOSIVE (plosive versus non-plosive initial phoneme) and VOICE (voiced versus unvoiced initial phoneme)).

Starting relatively fast within each item block, the participants slowed down towards the end of each block. The same inhibitory effect of REPETITION occurred in the compound experiment[4], applying basically the same task (identical in the learning and test phase, but with an additional practice phase in the compound experiments). It might be argued that, if anything, repeated namings of the same word should lead to shorter, rather than longer onset latencies. However, the response-association task seems to be rather boring and cause a decrease in motivation or alertness from the first naming to the last repetition.

[4] in the analysis of covariance with subject and word as crossed random effects

Our analysis of covariance reveals a disadvantage of words starting with a PLOSIVE as opposed to non-plosives (fricatives and other initial phonemes). This disadvantage for plosive initial words emerged in both regression analyses of Chapter 2. It is most likely an artifact of the voice key, which is not activated by the prevoicing of plosives. In the present study, words with unvoiced initial phonemes are found to be produced faster than words with voiced initial phonemes. Studies that use voice key measurements generally tend to find the opposite pattern, i.e. shorter onset latencies for voiced initial segments than for unvoiced segments. As discussed by Kessler, Treiman and Mullennix (2002), this tendency actually reflects differences in loudness and is caused by voice keys which have a low sensitivity or which are set to a high triggering threshold. Furthermore, Kessler et al report that this tendency does not hold for all individual phonemes and some unvoiced phonemes are detected faster than voiced phonemes.

In Chapter 2, I had included the variable NEIGHBORHOOD DENSITY in the regression analysis of the compound naming onset latencies, to see whether recent findings of neighborhood effects in production could be replicated. As the speech signal unfolds over time, from initial to last phoneme, a whole-word count of NEIGHBORHOOD DENSITY seemed, however, not perfectly suitable for speech studies that involve acoustic stimuli. For the later analysis, I, therefore, additionally computed the POSITION-SPECIFIC NEIGHBORHOOD DENSITIES (N1, N2, N3, etc.), representing the number of neighbors a word has, when particularly the first, second, third, etc. phoneme is exchanged. Notice, that the overall NEIGHBORHOOD DENSITY is the sum of, and, therefore, correlated with, the POSITION-SPECIFIC NEIGHBORHOODS. As described in the introduction, a dense (absolute) neighborhood is generally assumed to be detrimental for comprehension but has been found facilitative for speech production (e.g., Vitevitch, 2002, 2006). My findings are not in line with a general facilitation of neighborhoods in production. The overall effect of neighborhood density in the subject as random effect analysis of Chapter 2 was non-linear. The breaking down of the overall density count into position-specific counts yielded results which suggest different effects for different phoneme positions. There was an inhibitory effect of the POSITION-SPECIFIC NEIGHBORHOOD DENSITY for the initial phoneme (N1). The non-significant densities

N2 and N3 showed clear trends of facilitation. In the presence of the position-specific neighborhood counts, the overall NEIGHBORHOOD DENSITY dropped out of the model. As mentioned previously, predictors falling out during stepwise analysis of covariance are not necessarily of zero predictivity. In the presence of the stronger, correlated predictor, they may simply not add sufficient information to secure their place in the model, where only the stronger predictors remain. After all, stepwise analysis of covariance aims to explain as much variance as possible using as few predictors as possible.

The results underline the importance of NEIGHBORHOOD DENSITIES for speech production onset latencies. The inclusion of neighborhoods computed for specific positions within the word was fruitful. The position-specific density N1 is a stronger predictor than the overall density count. A high neighborhood density of the initial phoneme slows the naming latency of the deverbal adjective. The more rhyme neighbors a word has (sharing all but the first phoneme), the harder it is to produce the target word. Sevald and Dell (1994) report that it is easier to produce a sequence of rhyme words (such as *pick, tick*) than a sequence of cohorts (such as *pick, pin*). While overlapping segments generally help when sequences of words are produced, there is an inhibitory component overlapping initial phonemes. In their sequential cuing model, Sevald et al propose that shared segments miscue the production of later sounds, explaining why miscuing can happen in sequences such as *pin, pick*, but not in sequences such as *pick, tick*. The results of the present study suggest that, when N1-neighbors are produced in sequence (as in pick, tick), they not only benefit from their segmental overlap, they also co-activate each other as N1-neighbors. Under these circumstances, the disadvantage of a big N1-neighborhood (representing a miscuing in a much broader sense) turns into an advantage. When rhyme neighbors are not produced in sequence, however, bigger N1-neighborhoods mean more coactivation, making it harder to select the to-be-produced initial phoneme.

While the existence of many rhyme neighbors makes the production of a deverbal adjective more difficult, the existence of many words starting with the same two phonemes tends to facilitate its production. Naming is the fastest (i.e. the COHORT ENTROPYH2 is the highest) when the number of words in the cohort sharing the initial two phonemes is high and when these words have little variation in frequency. Converging evidence for an influence of the probability distribution in a production cohort on speech production stems from speech corpora studies analyzing the relative length of segments within words (Van Son & Pols, 2003, Van Son & Van Santen, 2005; Kuperman, Pluymakers, Ernestus, & Baayen, subm). The information load in a production cohort was found to be negatively correlated with the amount of reduction of segments. The results of the present study suggest that the frequency distribution in a production cohort does not only affect the length with which segments of the word are produced, but also the time it takes to plan the word. The higher the entropy over all words sharing the first and second phoneme with the deverbal adjective, the faster the adjective is named.

Among the frequency variables, it is the CUMULATIVE STEM FREQUENCY that remains in the final model. The FREQUENCY OF THE ADJECTIVAL SUFFIX is not a significant predictor for the naming latency of the deverbal adjective. Notice that the naming onset latencies of noun-noun compounds were affected by the positional entropy of the right constituent. The absence of a suffix-frequency effect may have several reasons. First, the frequency of the suffix might indeed have no effect on the naming latency. The encoding of a deverbal suffix might be easy enough to be done on the fly so that production can start after the encoding of the verbal stem. As a second explanation, the number of different suffixes used (nine) might be too small to show an effect. Third, each adjectival suffixes was contained in several items (see Table 3.4), while constituents in the compound experiment occurred only once. The repeated usage of the deverbal suffixes might have masked actual frequency differences.

The frequency of occurrence of the deverbal adjective itself dropped out off the model early on, while the frequency of the verbal stem stays a significant predictor. To underline that no variable is disregarded simply because it can not explain a significant proportion of variance in the presence of better predictors I present how non-significant these variables are even under the most inviting circumstances. Being included as the only predictor next to the significant control variables of the final model (REPETITION, VOICING, and PLOSIVE), none of the following variables yielded a significant fixed effect: SURFACE FREQUENCY (β = -

0.00187, t(26504) = -0.80, p = 0.42372); LEMMA FREQUENCY (β = -0.00236, t(26504) = -0.96, p = 0.33706); NEIGHBORHOOD DENSITY (β = 0.00080, t(26504) = 1.73, p = 0.08364). The only competitor for the CUMULATIVE STEM FREQUENCY was the POSITIONAL FREQUENCY (β = -0.00619, t(26504) = -2.69, p = 0.00715). However, in the presence of the POSITION-SPECIFIC NEIGHBORHOOD N1 and the COHORT ENTROPY H2, the CUMULATIVE STEM FREQUENCY was a better predictor and helped to explain more of the variance in the naming onset latencies of the deverbal adjectives.

To summarize, it is not the frequency of occurrence of the adjective itself, but the frequency of the verbal stem that predicts the latency with which a deverbal adjective is named. The more often *grijp* occurs in any form (independently or as part of a compounded, derived of inflected word), the shorter the naming latency of *grijpbaar*. The data are in line with the assumption of decomposition in the production of Dutch deverbal adjectives and challenge the assumption that morphologically complex words are fully listed at the word form level.

Chapter 5 presents a joint analysis of the naming onset latencies of the deverbal adjectives and the naming onset latencies of regular Dutch inflected verbs (Chapter 4). A merged analysis of these naming latencies (there is a high overlap in the verb stems in the item pools) provides the chance to directly compare how particular variables influence the production of these two types of morphologically complex words.

References

Baayen, R.H. (2007) *Analyzing linguistic data. A practical introduction to statistics*. Cambridge University Press, in press.

Baayen, R.H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database* (CD-ROM), Linguistic data consortium. University of Philadelphia, Philadelphia, PA, Release 2.

Baayen, R.H., Tweedie, F.J., & Schreuder, R. (2002). The subjects as a simple random effect fallacy: Subject variability and morphological family effects in the mental lexicon, *Brain and Language*, *81*, 55–65.

Bates, D.M. (2005). Fitting linear mixed models in R, *R News*, *5*, 27-30.

Bates, D.M., & Sarkar, D. (2005). lme4: Linear mixed-effects models using S4 classes, R package version 0.9975-7.

Booij, G.E. (2002). *The morphology of Dutch*. Oxford: Oxford University Press.

Cholin, J., Levelt, W.J.M., & Schiller, N.O. (2006). Effects of syllable frequency in speech production. *Cognition*, *99*, 205-235.

Coltheart, M., Davelaar, E., Jonasson, J.T., & Besner, D (1977). Access to the internal lexicon. In Dornic, S. (Ed.), Attention and Performance VI. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Greenberg, J.H., & Jenkins, J.J. (1964). Studies in the psychological correlates of the sound system of American English. *Word*, *20*, 157-177.

Kessler, B., Treiman, R., & Mullennix, J. (2002). Phonetic biases in voice key response time measurements. *Journal of Memory and Language*, *47*, 145-171.

Kuperman, V., Pluymaekers, M., Ernestus, M., & Baayen, R. H. (subm). Morphological predictability and acoustic salience of interfixes in Dutch compounds. To appear in Journal of the Acoustical Society of America.

Meyer, A.S., Roelofs, A., & Levelt, W.J.M. (2003). Word length effects in object naming: The role of a response criterion. *Journal of Memory and Language*, *48*, 131-147.

Pinheiro, J.C., & Bates, D.M. (2000). Mixed-effects models in S and S-PLUS, *Statistics and Computing Series*, Springer-Verlag, New York, NY.

Sevald, C.A., & Dell, G.S. (1994). The sequential cuing effect in speech production. *Cognition*, *53*, 91-127.

Van Son, R.J.J.H., & Pols, L.C.W. (2003). How efficient is speech? *Proceedings of the Institute of Phonetic Sciences*, *25*, 171-184.

Van Son R.J.J.H., & Van Santen, J.P.H. (2005). Duration and spectral balance of intervocalic consonants: A case for efficient communication. *Speech Communication*, *47*, 100-123.

Vitevitch, M.S. (2002). The influence of phonological similarity neighborhoods on speech production. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *28*, 735-747.

Vitevitch, M.S., & Stamer, M.K. (2006). The curious case of competition in Spanish speech production. *Language & Cognitive Processes, 21,* 760-770.

Appendix: Deverbal adjectives

a) The 124 deverbal adjectives grouped by suffix

-achtig (10)

beuzelachtig
huichelachtig
kittelachtig
twijfelachtig
tekenachtig
toverachtig
schilderachtig
regenachtig
weigerachtig
weifelachtig

-baar (33)

brandbaar
breekbaar
drinkbaar
draaibaar
deelbaar
huwbaar
laakbaar
leesbaar
haalbaar
kneedbaar
rekbaar
scheidbaar
smeerbaar
houdbaar
leefbaar
meetbaar
misbaar
hoorbaar
strijdbaar
strafbaar
telbaar
grijpbaar
vloeibaar
voelbaar
vindbaar
tastbaar
toonbaar
vatbaar
plooibaar
vangbaar
kwetsbaar
splijtbaar
wendbaar

-elijk (13)

hopelijk
gruwelijk
hatelijk
denkelijk
schromelijk
merkelijk
sterfelijk
vreselijk
plaatselijk
schadelijk
schrikkelijk
smakelijk
wenselijk

-erig (31)

beverig
broeierig
brommerig
dromerig
druilerig
dweperig
hebberig
hijgerig
jeukerig
kloterig
kruiperig
jankerig
kleverig
krakerig
lacherig
hangerig
tobberig
plagerig
pruilerig
huilerig
pronkerig
slaperig
rillerig
pestering
plakkerig
snauwerig
soezerig
piekerig
springerig
schraperig
zweterig[5]
zweverig[4]

-ig (11)

roezig
flossig
knorrig
happig
vluchtig
stellig
morsig
willig
zwierig
warrig
woelig

-lijk (4)

deerlijk
draaglijk
klaaglijk
heuglijk

-loos (2)

weerloos
reddeloos

-s (4)

Broeds
Speels
Waaks
Sleets

-zaam (15)

buigzaam
handzaam
minzaam
raadzaam
lijdzaam
duldzaam
duurzaam
leerzaam
volgzaam
zorgzaam
spaarzaam
voegzaam
voedzaam
werkzaam
zwijgzaam

[5] The phonological neighbors *zweverig* and *zweterig* were excluded from the analysis.

b) The 62 items sets

The 124 deverbal adjectives assigned to 62 item sets for presentation in the position-response association task. The nine Dutch deverbal suffixations are labeled in parenthesis (*-achtig* (1), *-baar* (2), *-elijk* (3), *-erig* (4); *-ig* (5), *-lijk* (6), *-loos* (7), *-s* (8), *-zaam* (9)).

sets 1-21		sets 22-42		sets 43-63	
1	sleets (8) willig (5)	22	twijfelachtig (1) meetbaar (2)	43	plagerig (4) vatbaar (2)
2	weerloos (7) minzaam (9)	23	duurzaam (9) lacherig (4)	44	huichelachtig (1) kleverig (4)
3	morsig (5) denkelijk (3)	24	toonbaar (2) zwijgzaam (9)	45	voegzaam (9) snauwerig (4)
4	woelig (5) krakerig (4)	25	leefbaar (2) huilerig (4)	46	jeukerig (4) Voelbaar (2)
5	dweperig (4) broeds (8)	26	piekerig (4) rillerig (4)	47	hoorbaar (2) flossig (5)
6	lijdzaam (9) kruiperig (4)	27	happig (5) rekbaar (2)	48	wenselijk (3) houdbaar (2)
7	splijtbaar (2) pronkerig (4)	28	sterfelijk (3) duldzaam (9)	49	waaks (8) Leesbaar (2)
8	roezig (5) kloterig (4)	29	gruwelijk (3) wendbaar (2)	50	scheidbaar (2) Hangerig (4)
9	klaaglijk (6) huwbaar (2)	30	smakelijk (3) vindbaar (2)	51	Misbaar (2) zorgzaam (9)
10	zwierig (5) plaatselijk (3)	31	schadelijk (3) telbaar (2)	52	Hebberig (4) weifelachtig (1)
11	vluchtig (5) draaglijk (6)	32	deerlijk (6) schraperig (4)	53	warrig (5) Grijpbaar (2)
12	broeierig (4) speels (8)	33	werkzaam (9) beuzelachtig (1)	54	vloeibaar (2) Zweterig[6] (4)
13	pruilerig (4) spaarzaam (9)	34	soezerig (4) laakbaar (2)	55	Haalbaar (2) Plakkerig (4)
14	raadzaam (9) pesterig (4)	35	schrikkelijk (3) weigerachtig (1)	56	heuglijk (6) kwetsbaar (2)
15	dromerig (4) plooibaar (2)	36	stellig (5) brommerig (4)	57	hatelijk (3) brandbaar (2)
16	slaperig (4) vreselijk (3)	37	kittelachtig (1) springerig (4)	58	breekbaar (2) druilerig (4)
17	tastbaar (2) volgzaam (9)	38	schilderachtig (1) merkelijk (3)	59	buigzaam (9) drinkbaar (2)
18	knorrig (5) voedzaam (9)	39	schromelijk (3) handzaam (9)	60	regenachtig (1) draaibaar (2)
19	jankerig (4) vangbaar (2)	40	toverachtig (1) beverig (4)	61	Zweverig[6](4) kneedbaar (2)
20	leerzaam (9) deelbaar (2)	41	tobberig (4) reddeloos (7)	62	hopelijk (3) smeerbaar (2)
21	hijgerig (4) strafbaar (2)	42	tekenachtig (1) strijdbaar (2)		

FREQUENCY EFFECTS IN THE PRODUCTION OF INFLECTED VERBS

CHAPTER 4

4.1. Introduction

After having looked at frequency effects in the production of compounds (Chapter 2) and derivations (Chapter 3), we now turn to the third type of morphologically complex words: inflections. In particular, we look at one subcategory of inflections: the production of inflected verbs.

All three types of morphologically complex words (compounds, derivations, and inflections) contain multiple morphemes. Among the morphologically complex words, compounds are unique as they generally[1] consist of free morphemes, while both derivation and inflection involve the attaching of bound morphemes (such as the Dutch *-laar*, *-lijk*, *-en*, *-t*, etc). Bound morphemes, in contrast to free morphemes, cannot exist as independent words. The attaching of a bound morpheme to a word is called affixation. Derivations and inflections differ with respect to the outcome of the affixation. In derivational morphology, affixation always forms a new word. As described in the previous chapter, the derived word can be of a different word class (noun, verb, adjective) than its base. In contrast, in inflectional morphology, affixation does not form a new word, but a syntactically appropriate variant of the same word.

[1]There are a few exceptions to this rule. Some compounds (such as *aardbei*) contain a constituent that does not exist as a free morphemes (*aard* is a free morpheme, *bei* is not).

4.1.1. Inflected verbs

In Dutch, the attaching of a particular inflectional affix to different verb stems forms variants that are appropriate for the same syntactical context (similar to the same change in meaning described for the attaching of particular deverbal suffixes in Chapter 3). Syntactically, *gedroomd* (has/have dreamt) relates to *dromen* (to dream) as *gemist* (has/have missed) relates to *missen* (to miss), *dromend* (dreaming) relates to *drome*n (to dream) as *missend* (missing) relates to *missen* (to miss), and so on. Speakers are, therefore, also able to inflect unknown words or words that are adopted from other languages and use them appropriately in any syntactical context (e.g., *ge-sms-t*). Table 4.1 lists the seven different verbal inflections that are used in Dutch.

Table 4.1: Verbal inflections in Dutch

<table>
<tr><th colspan="3" rowspan="2">syntactically appropriate variant</th><th rowspan="2">affixation</th><th colspan="2">examples</th></tr>
<tr><th>missen (to miss)</th><th>dromen (to dream)</th></tr>
<tr><td rowspan="6">Present Tense:</td><td rowspan="3">sg</td><td>1st p</td><td>-</td><td>mis</td><td>droom</td></tr>
<tr><td>2nd p</td><td rowspan="2">-t / -d</td><td rowspan="2">mist</td><td rowspan="2">droomt</td></tr>
<tr><td>3rd p</td></tr>
<tr><td rowspan="3">pl</td><td>1st p</td><td rowspan="4">-en</td><td rowspan="4">missen</td><td rowspan="4">dromen</td></tr>
<tr><td>2nd p</td></tr>
<tr><td>3rd p</td></tr>
<tr><td colspan="3">Infinitive</td></tr>
<tr><td colspan="3">Present Participle</td><td>-end</td><td>missend</td><td>dromend</td></tr>
<tr><td rowspan="6">Past Tense:</td><td rowspan="3">sg</td><td>1st p</td><td rowspan="3">-de / -te</td><td rowspan="3">miste</td><td rowspan="3">droomde</td></tr>
<tr><td>2nd p</td></tr>
<tr><td>3rd p</td></tr>
<tr><td rowspan="3">pl</td><td>1st p</td><td rowspan="3">-den / -ten</td><td rowspan="3">misten</td><td rowspan="3">droomden</td></tr>
<tr><td>2nd p</td></tr>
<tr><td>3rd p</td></tr>
<tr><td colspan="3">Past Participle</td><td>ge-…- t / -d</td><td>gemist</td><td>gedroomd</td></tr>
</table>

4.1.2. Research question

Ever since the seminal study of Oldfield and Wingfield (1965), it is known that words of higher frequency of occurrence can be named faster than words of lower frequency of occurrence. This correlation of word frequency and naming latency is generally referred to as the word frequency effect. In addition to their own frequency of occurrence, morphologically complex words have constituent morphemes each of which has its own frequency. Models of speech production differ with respect to the predictions they make about frequency effects in morphologically complex words. If morphologically complex words are stored in, and retrieved from, the mental lexicon in the same way as morphologically simplex words are, their naming onset latencies should be related to their frequency of occurrence as complex words. If, however, morphologically complex words are assembled from their constituents during production, their naming onset latencies should instead be related to the frequencies of these constituents. After having looked at noun-noun compounds (where there were both compositional and non-compositional effects, Chapter 2) and deverbal adjectives (which behaved compositionally, Chapter 3), I study the naming onset latencies of Dutch regularly inflected verbs.

According to Pinker's (1999) words and rules approach, regular inflections are generated by a rule. Other than irregular words, they don't supply a past-tense form from memory. The regular rule, therefore, applies by default. If regular inflections are computed on the fly, one expects to find a frequency effect for the verbal stem. Similarly, Levelt, Roelofs, and Meyer (1999) assume that a regular inflected form (e.g., *escorting*) is generated from the lemma (e.g., *escort*), which is marked for the relevant regular inflection (e.g., *+progressive*), activating the form nodes <escort> and <-ing>. The selection of a lemma of an irregular verb (e.g., *go*) +past is assumed to activate the form node <went>. Both approaches assume that the lexical representation of regularly inflected verbs is a representation in which the stem and the inflectional affix are composed at the word form level. Assuming that frequency effects are located at the word form level, both approaches, therefore, predict stem frequency effects in regular inflections. In contrast, full-listing models (e.g., Butterworth, 1983) predict that the naming latency with which an inflected verb is produced correlates with its frequency of occurrence as a complex word. It is further possible that high-frequency regular variants are stored as complex words, while low-frequency regular inflections are always composed from their constituents (e.g., Stemberger & MacWhinney, 1986). Finally, one can assume that inflected verbs, such as complex words in general, are stored with structural information, i.e., with links between constituting morphemes that reflect the probabilities with which these morphemes tend to be combined.

Compared to compounds and derivations, inflections are morphologically complex words that differ least from their stems. Compounds are combinations of two free morphemes, each of which can in principle occupy the slot of head or modifier. The result of compounding is a new word. Derivations are new words, formed through affixation of bound morphemes. Still, some derivational affixes (such as '*zaam*', '*baar*', '*erig*') are phonological words. Affixation in inflection, in contrast, does not form a new word, but a syntactically appropriate variant of the same word. Dutch inflectional affixes are not phonological words and are very short (they consist of maximally three phonemes). Therefore, on a continuum of constituent similarity, compounds would score the highest, inflections would score the lowest.

The dissimilarity of the constituent morphemes in inflected verbs and the very limited number of verbal inflectional affixes make it unfeasible to construct item pairs that carry a factorial contrast in one frequency while other frequencies are matched (as done with the noun-noun compounds in Chapter 2). I, therefore, collected naming onset latencies of inflected verbs that are distributed over a wide range of frequencies to analyze in stepwise analyses of covariance.

4.1.3. Predictors

As introduced in the first chapter, several predictors can be analyzed simultaneously using stepwise analysis of covariance. All variables used in this experiment, are defined and motivated in the first chapter of this thesis. Based on the reasons presented there and based on the experience with the previous experiments, I am interested in four kinds of predictors: control variables, frequency variables, morphological variables, and phonological variables.

More specifically, the control variables REPETITION (namings one to ten in the test phase, see 4.2.3.), PLOSIVE (plosive versus non-plosive initial phoneme) and VOICED (voiced versus unvoiced initial phoneme) are included.

In the frequency domain, I study the LEMMA FREQUENCY of the verb (the summed frequency of all inflectional variants of that word), its CUMULATIVE STEM

FREQUENCY (the summed frequency of all words containing the verb stem), and its POSITIONAL STEM FREQUENCY (the summed frequency of all morphologically complex words containing the verb stem as initial constituent).

From the group of morphological variables, I include the INFLECTIONAL ENTROPY, defined as Shannon's entropy estimated by the relative frequencies of a verb's inflectional variants. It is the token-weighted count of the numbers of types of inflectional variants. A high INFLECTIONAL ENTROPY indicates that a particular verb stem is actually used in many or all of its inflectional variants, and that these inflections occur with similar frequency. When the entropy in an inflectional paradigm is high, the production of a specific inflected form might be relatively harder than when the INFLECTIONAL ENTROPY is low.

As a first phonological variable, the PHONOLOGICAL WORD LENGTH is included in the stepwise analyses of covariance. In research of speech production, longer words have been observed to elicit longer naming onset latencies (e.g., Meyer, Roelofs, & Levelt, 2003). However, PHONOLOGICAL WORD LENGTH was not a significant predictor in the previous analyses presented in Chapters 2 and 3.

We further take along the two-level factor PREFIX. One of the seven types of Dutch verbal inflections, the past participle, contains a prefix (*ge-*). I am curious to see, whether the presence or absence of a prefix has an effect on the production onset latencies, independent of PHONOLOGICAL WORD LENGTH. The factor PREFIX contains two levels: prefixed (all past participle forms) versus unprefixed (all other inflections).

As with the noun-noun compounds and deverbal adjectives, I take along the number of phonological neighbors of the inflected verbs. Both the whole-word NEIGHBORHOOD DENSITY and the POSITION-SPECIFIC NEIGHBORHOODS (N1, N2, and N3) are computed. Phonological neighbors are words that can be transformed into one another by exchanging only one phoneme (Greenberg & Jenkins, 1964; Coltheart, Davelaar, Jonasson, & Besner, 1977). A word’s neighborhood density refers to the total number of neighbors the word has. Position-specific neighbors refer to all words that differ from the to-be-pronounced word in exactly the first, second, third, etc. phoneme.

We, finally, include the COHORT ENTROPIES H1, H2, and H3, which are entropy measures based on the production cohort. The COHORT ENTROPIES represents the likelihood of the target word given all other words beginning with the same first (H1), first two (H2), first three (H3), etc. phonemes. When a cohort has many members and when these members are of similar frequencies, COHORT ENTROPIES are the highest. I observed significant effects for a POSITION-SPECIFIC NEIGHBORHOOD (N1) and a COHORT ENTROPY (H2) with the deverbal adjectives (Chapter 3) but not with the noun-noun compounds (Chapter 2).

4.2. Materials and Method

4.2.1. Material selection

We selected 126 Dutch verb stems from the CELEX lexical database (Baayen, Piepenbrock, & Guliker, 1995). I tried to use as many verb stems as possible that had also been used in the study on deverbal adjectives (Chapter 3).On the one hand, I want to study particular effects in the production of inflected verbs. On the other hand, I want to compare the particular effects in the production of the different types of morphologically complex words (see comparative analyses in Chapter 5). The higher the overlap in verb stems between the study on deverbal adjectives and the study on inflected verbs, the higher their comparability. I did, however, not reuse all verbal stems as I had a mayor restriction on material selection for the study on the production of inflected verbs: I wanted to use regular verbs only. Irregular inflections differ from regular inflections in several aspects. Some irregular inflections undergo vowel changes. Furthermore, inflectional variants of irregular verbs often differ in the number of syllables from inflectional variants of regular verbs. We, therefore, decided to not mix regular with irregular verbs in this study, but rather focus on the role of frequency information in the inflection of regular Dutch verbs.

Starting from the item pool of the derivation study, all stems of irregular verbs were excluded and replaced by about the same number of regular verbal stems that had not been used in the derivational study. The final overlap of the two items sets was 76%. For the benefit of statistical analyses the selection procedures for the previous experiments aimed to compile sets of items that are fairly distributed over a wide range of frequencies. The same intent guided the selection of the new items. Adding 32 new regular verb stems, a total number of 126 items was reached, which is comparable to the total number of items used in the previous experiments (128 in the compound study, 124 in the derivation study). For all selected verbs, CELEX[2] stem frequencies were greater than zero. I additionally collected the Google frequencies of all items as well as familiarity ratings[3] from 46 participants, both of which were well in line with the CELEX frequencies.

Table 4.2: 126 regular verbal stems were pseudo-randomly assigned to the seven types of Dutch verbal inflections (1-7), 18 items per type of inflection.

(1) present 1st singular	(2) present 2nd, 3rd singular	(3) present 1st, 2nd, 3rd plural / infinitive	(4) present participle	(5) past 1st, 2nd, 3rd singular	(6) past 1st, 2nd, 3rd plural	(7) past participle
bouw	bakt	broeden	brandend	beefde	beuzelden	geacht
broei	droomt	brommen	draaiend	deerde	daalden	gedeeld
duld	gunt	gruwen[3]	durend	hapte	dweepten	gehaald
huil	huichelt	haten	hopend	klaagde	hijgden	gehoord
kneed	kleeft	knorren	jankend	lachte	jeukten	gehuwd
leer	left	linen	krakend	pestte	kookten	gekwetst
maak	merkt	minnen	pakkend	piekte	morsten	gemist
plak	plaatst	plagen	plooiend	rekte	pleegden	gepronkt[4]
ren	regent	remmen	redden	rustte	pruilden	geraakt
scheur	schaatst	scheiden	rillend	schudde	schaadden	geruimd
snauw	smeert	smaken	schrapend	spoelde	schroomden	geschilderd
sticht	stelt	staken	soezend	stuitte	speelden	gespaard
tel	tekent	tasten	stoppend	trachtte	straften	gestoord
veeg	twijfelt	toveren	tobbend	voedde[3]	toonden	getild
volg	voelt	voegen	vloekend	weerde	vluchtten	gevloeid
weiger	weifelt	walgen	wakend	woelde	vulden	gevreesd
zaai	zaagt	wenden	werkend	zoemde	wensten	gewist
zeur	zoent	zweven	zwetend	zwierde	zorgden	gezet

The 126 selected verbal stems were evenly distributed over the seven types of Dutch verbal inflections introduced in Table 4.1 The assignment to a particular type

[2] All CELEX frequencies reported are based on a corpus of 42 million words.

[3] Familiarity was rated on a 7 point scale, yielding a mean of 4.4 and a standard deviation of 1.8 for the selected item pool.

[4] The items *gepronkt*, *gruwen*, and *voedde* were not included in the analyses (see 4.3)

of inflection was pseudo-randomized with two restrictions: First, the frequency of the inflected form had to be greater than zero. Second, initial phonemes should be fairly distributed over the seven groups, especially with respect to features such as plosiveness and voice. Table 4.2 lists all inflected verbs used in the experiment. A complete list of items, as paired in the position-response association task (see 4.2.3), is provided in the Appendix. Figure 4.1 shows, in logarithmic scale, that the selected set of items was fairly distributed over a wide range of both lemma frequencies and cumulative stem frequencies.

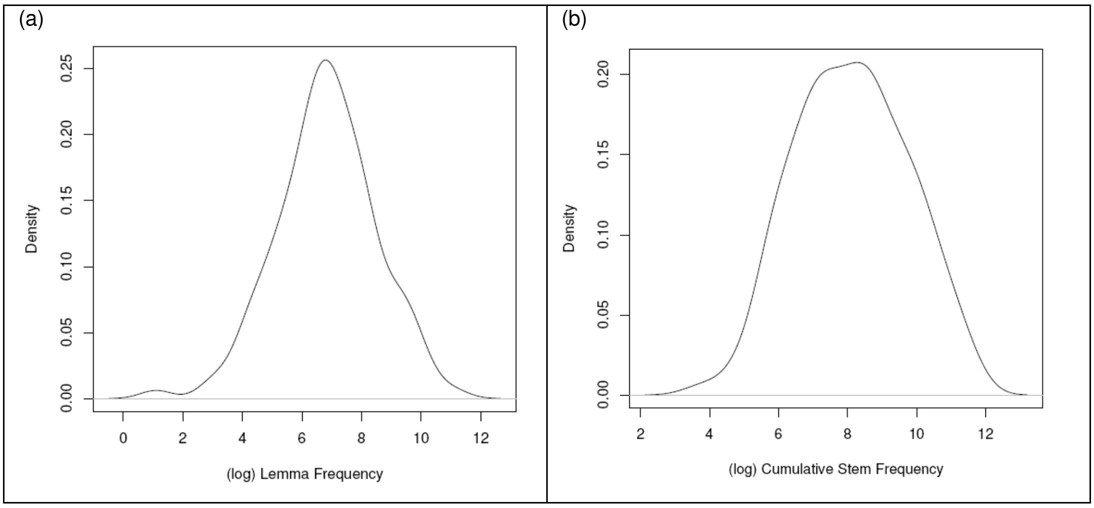

Figure 4.1: The distribution of LEMMA FREQUENCIES (a) and CUMULATIVE STEM FREQUENCIES (b) in the material.

4.2.2. Experimental lists

A total of 24 experimental lists was constructed, one list for each participant. In a first step, the 126 inflected verbs were assigned to 63 sets of two items each (item A and B), applying the following restrictions: The verbs within a set had to be of a different inflection, they had to have minimal phonological overlap and not be semantically related.

In a second step, I constructed four basic lists, each of which contained the 63 item sets in the same order (see table 4.3). The basic lists differed with respect to the order of the two items within the sets. This order plays an important role in the position-response association task used which is described in more detail in 4.2.3. In this task, the first item is presented *first* and associated with an icon on the *left* side of the screen, while the second item is presented *second* and associated with an icon on the *right* side of the screen. As this difference in presentation might affect the production onset latencies, I balanced the order of items within the sets, so that every item is presented left to one half of the participants and right to the other half of the participants.

Table 4.3: The four basic lists counterbalanced the left / right presentation of items in the position-response association task.

	Item sets 1 - 32		Item sets 33-63	
	item A	item B	item A	item B
Basic List 1	left	right	left	right
Basic List 2	right	left	right	left
Basic List 3	left	right	right	left
Basic List 4	right	left	left	right

As shown in Table 4.3, basic list 1 assigned all A-items to be presented on the left side, all B-items to be presented on the right side. In basic list 2, I reversed the order in all 63 sets, assigning all A-items to the right side and all B-items to the left side. For basic list 3, I reversed the order of items within the sets 1-32 only. For the final basic list (4), I reversed the order within the sets 33-63 only. In a third step, six randomizations of the order of sets within each basic list were constructed. To summarize, I constructed a total of 24 experimental lists with randomized set order and balanced order of items within each set.

As with the derivational experiment reported in the previous Chapter, a complete experimental list containing all 63 sets would be too long to be presented in one session. After randomization of set order, I therefore split each experimental list into two semi-lists, the first one containing 32 sets, the second one containing 31 sets. Finally, three practice sets with inflected verbs of similar structure were added to precede the experimental sets on each semi-list.

4.2.3. Position-Response Association Task

In order to measure the production onset latencies of the inflected verbs, I used a position-response association task (Cholin, Levelt, & Schiller, 2006), in which the two inflected verbs of a set are associated with a position mark on the left or right side of the computer screen. Each participant was tested individually in a sound-attenuated booth. The participant wore headphones and was comfortably seated in front of a Sennheiser microphone, a button box, and a CRT computer screen.

The position-response association task used in this experiment was identical to the one used with the deverbal adjectives (Chapter 3). The position-response association task that I used with the compounds (Chapter 2) included an additional practice phase between the learning and the test phase. In this experiment, I decided against the inclusion of a practice phase for three reasons. First, the participants had shown a close to zero error rate during the practice phases of the compound experiments, suggesting that they would have done well during the test phase without going through the practice phase. Second, leaving out the practice phase in the experiment on deverbal adjectives (Chapter 3) did not increase the percentage of hesitation, wrong naming, time outs, etc. during the test phase. Third, leaving out the practice phase freed valuable minutes that could be rather used to test additional item sets.

The experimental procedure consisted of learning phases followed by test phases for each set individually. The learning and test phase of one set form a block. Both the learning and the test phase started with the word 'ATTENTION' (*attention*) being presented in the center of the screen for two seconds. Both the learning and the test phase ended with the word 'PAUZE' (*pause*) being presented in the center of the screen. The pause signal remained on the screen until the participant initiated the following phase by pressing a button.

In the learning phase, participants established an association between the two inflected verbs of a given set and an icons appearing on the left versus right side of the computer screen. This association was established in the following way: The *first* item was presented over headphones. Simultaneously, an icon of a loudspeaker was presented on the *left* side of the screen. Then, the *second* item was presented over headphones with a simultaneous presentation of the loudspeaker item on the *right* side of the screen. This procedure was repeated once.

Having learned the association between the two inflected verbs and the icons on the two sides of the screen, participants could then be prompted by an icon to produce the associated inflected verb in the test phase. For each set, the test phase immediately followed the learning phase. In the test phase, each item was prompted ten times (20 experimental trials in total within each test phase). The order of prompting of the two items was pseudo-randomized with as restriction that no more than 4 consecutive repetitions of the same item were allowed.

A total of 20 distractor trials were added, each of which preceded an experimental trial. I included distractor trials to make it difficult for the participants to prepare one of the inflected words before a prompt was presented and to avoid immediate consecutive naming of the same item. In a distractor trial, participants simply named a single-digit number (1, 2, 3, or 6) that was presented in the center of the screen.

Participants were instructed to name each item and each number as quickly and correctly as possible. Simultaneously with the presentation of a prompt or a number the voice key was activated for 1500 ms. A triggering of the voice key within 1500 ms (due to the participant producing the prompted item or due to any other, loud enough sound) was automatically registered. The experimenter monitored the participant's responses through headphones and took notes of incorrect naming, hesitations and voice key errors for later exclusion of these trials. Naming onset latencies longer than 1500 ms were registered as time-outs.

Each participant took part in two sessions with semi-list A being presented in the first, and semi-list B being presented in the second session. A single session lasted 70 min on average. Participants had to take a minimal break of 90 min in between their two sessions. 16 of the 24 participants took part on two different days.

4.2.4. Participants

24 native speakers of Dutch (21 female, 3 male) were recruited from the subject pool of the Max-Planck Institute for Psycholinguistics, Nijmegen. eight participants completed both sessions on the same day, sixteen completed them on two different days. For completing both sessions of the experiment, each participant received € 15.

4.3. Results

Some naming latencies had to be excluded prior to the analyses. One participant did not complete the experiment. I included none of his latencies in the analyses, reducing the actual number of participants to 23. I further excluded all latencies of the experimental item *gruwen*. Many participants seemed to have problems recognizing this item via headphones, causing an unusual high number of hesitations, time outs trials and false naming responses. These latencies could, therefore, not be considered as proper production onset latencies for the intended item.

Furthermore, all onset latencies of the experimental item *voedde* were excluded, because this item accidentally appeared in two sets, once replacing the item *gepronkt*. Producing a single item twice as often as all other items, or coming across the item in two sets, while all other items appear in one set only, might affect the naming onset latencies measured for this item. I, therefore, decided to take no risk and exclude all naming latencies for this item from the analysis.

Of the remaining 28290 experimental trials, we included only those trials in the analyses that were named both correctly and within a latency of 1500 ms. A total of 1307 (4.6%) time out trials (onset latencies >1500 ms), hesitations, wrong naming and voice key errors was removed prior to the analysis. During the analysis, an additional 284 (1 %) extreme outliers (data points with absolute standardized residuals exceeding 2.5 standard deviations from the mean) were identified and excluded.

We analyzed the data using a stepwise multilevel analysis of covariance with subject and word as crossed random effects (Pinheiro & Bates, 2000; Baayen, Tweedie, & Schreuder, 2002; Bates, 2005; Bates & Sarkar, 2005; Baayen, in press) with participant as main grouping factor. Given the number of items used, no more than eight item-based parameters should be included in the model to avoid the risk of overfitting[5].

[5] As a rule of thumb to avoid overfitting, the number of item-based predictors should not exceed the number of items divided by 15.

The analysis resulted in a model with four predictors (including two interacting control variables). Table 4.4 summarizes the fixed-effects statistics, including beta weights, standard errors, t-values and p-values. Figure 4.2 pictures the partial effects of the significant predictors, each adjusted for the effects of the other covariables. As random effects the model incorporated random intercepts for word stem (STD = 0.038) and for subject (STD = 0.122), and the residual error (STD = 0.219).

Table 4.4: Multilevel analysis of covariance with subject and word as random effects resulted in these fixed effects. For the Predictor INITIAL PHONEME, there are adjustments for plosives and other initial phonemes, the fricatives lie on the intercept.

		Estimate	Std.Error	DF	t.value	p.value
INTERCEPT		6.0171	0.0289	26699	208.465	0.0000
REPETITION		0.0055	0.0007	26699	8.236	0.0000
INITIAL PHONEME	**other**	0.0561	0.0124	26699	4.522	0.0000
	plosive	0.0645	0.0109	26699	5.929	0.0000
REPETITION *by* INITIAL PHONEME	**other**	-0.0046	0.0012	26699	-3.795	0.0002
	plosive	-0.0020	0.0011	26699	-1.835	0.0665
H2		-0.0080	0.0024	26699	-3.315	0.0009
PREFIXED		0.0941	0.0146	26699	6.458	0.0000

Like in the experiments reported in the previous chapters, the participants started relatively fast within the test phases and slowed down towards their end. The inhibitory effect of the control variable REPETITION (β = 0.0055, t(26699) = 8.236, p<0.0001) interacts with the control variable INITIAL PHONEME in such a way that words starting with a fricative (f) or a plosive (p) are produced slower over repetitions, while there is no slowing down for words containing other (o) initial phonemes (F(2,26697) = 7.2966, p = 0.0007, panel 1 of figure 4.2).

The model reveals a disadvantage of prefixed inflections as opposed to inflections that involve no prefixation (β = 0.0941, t(26699) = 6.458, p<0.0001). The only Dutch verbal inflection carrying a prefix is the past participle form (*ge-…-t*).

Picturing the effect of PREFIX, panel 2, therefore, simultaneously contrasts the slower onset latencies for past participle items (right) as opposed to the items of all other verbal inflections (left).

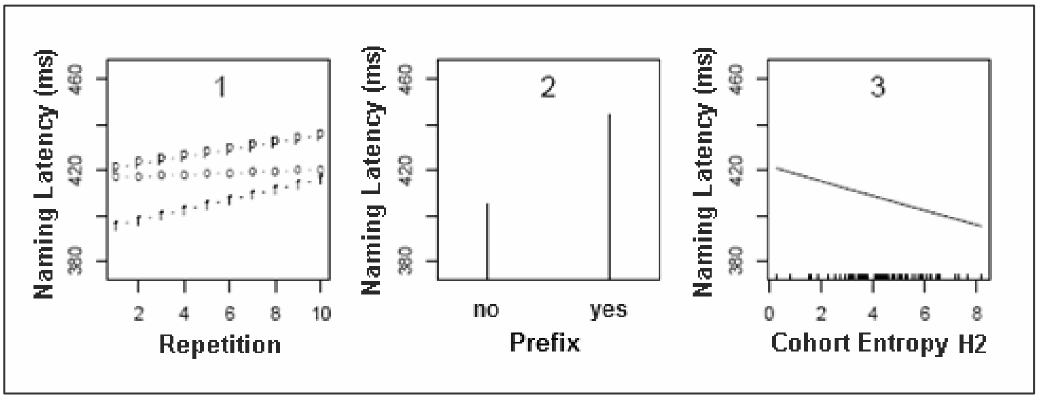

Figure 4.2: The partial effects of the significant predictors, adjusted for the effects of the other covariates. Panel 1, pictures the interaction of REPETITION and INITIAL PHONEME (plosive (-p-), fricative (-f-), other (-o-)).

We further observe a facilitative effect of the COHORT ENTROPY H2 (β = -0.0080, t(26699) = -3.315, p=0.0009). The more words exist that start with the same two phonemes as the target word, the faster the target word can be named (panel 3).

4.4. Discussion

Stimulated by previously associated visually presented symbols, 24 native speakers of Dutch repeatedly produced 126 Dutch inflected verb forms. The variance in the collected naming onset latencies is best modeled using the following information: how often has the item been produced previously in the experiment by the same participant (REPETITION), whether or not the item contains a prefix (PREFIX), and, finally, on the entropy in the production cohort H2.

Like in the previous experiments, the inhibitory effect of REPETITION seems to suggest that the test-phases of the position-association learning task failed too maintain the initial alertness and motivation over all 40 trials (20 namings of experimental items and 20 namings of random numbers). The interaction with INITIAL PHONEME can be interpreted as follows: repetitive articulations of fricatives and plosives are more tiring than repetitive articulations of other initial phonemes.

As an alternative explanation, the two experimental items within a test phase could become more activated and, therefore, stronger competitors with increasing number of repetition. After all, the task is very predictable with respect to the *moment* in which the trigger of an experimental item is presented (unpredictable is its *position*). Participants might, therefore, prepare both experimental items while waiting for the trigger.

Information on the PHONOLOGICAL WORD LENGTH or NEIGHBORHOOD DENSITIES (overall or position-specific) does not help to explain a significant proportion of variance in the naming onset latencies of the inflected verbs. Also not among the best predictors are the frequency of occurrence of the inflected verb form, the summed frequency of all inflectional variants of verb (LEMMA FREQUENCY), or Shannon's entropy estimated by the relative frequencies of a these inflectional variants (INFLECTIONAL ENTROPY). Neither do I find an effect of CUMULATIVE or POSITIONAL STEM FREQUENCY.

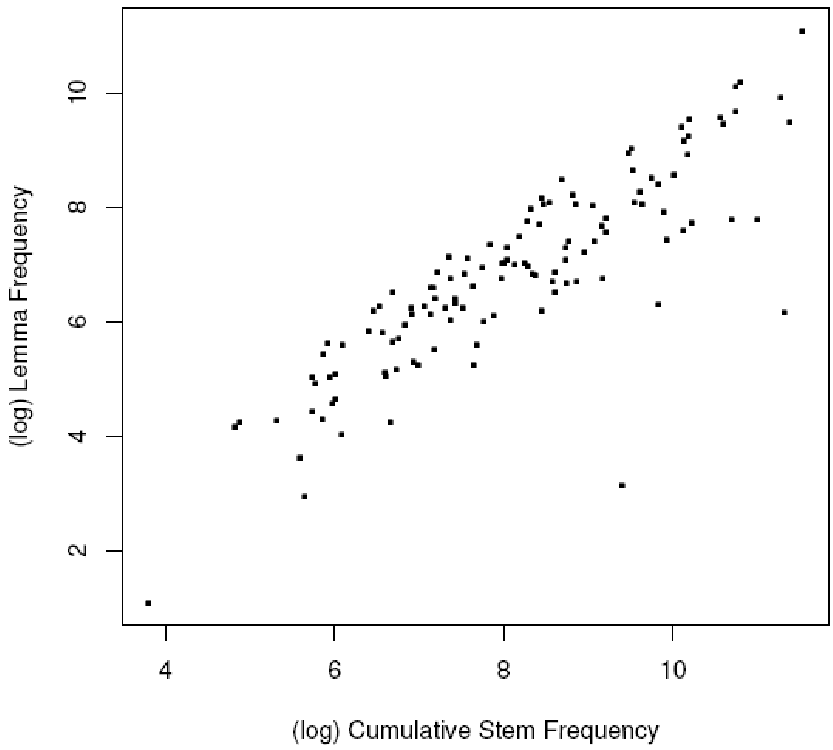

Figure 4.3: In Dutch verbal inflections, the LEMMA FREQUENCY and the CUMULATIVE STEM FREQUENCY are highly correlated (here, r = 0.85).

Generally, the LEMMA FREQUENCY is part of the CUMULATIVE STEM FREQUENCY. In the case of inflected verbs, the CUMULATIVE STEM FREQUENCY is little more than the LEMMA FREQUENCY. Therefore, these two measures of frequency are correlated (r = 0.85 in this item pool). The scatter plot shown in Figure 4.3 pictures the correlation between LEMMA FREQUENCY and CUMULATIVE STEM FREQUENCY in the present sample.

In multivariate regression analyses, highly correlated predictors cause a high colinearity. To test, whether the high correlation between the CUMULATIVE STEM FREQUENCY and the LEMMA FREQUENCY in the material is responsible for the absence of a frequency effects, I computed a new variable, to which I will refer as OTHER FREQUENCY. This variable represents the unique part of the CUMULATIVE STEM FREQUENCY that does not overlap with the LEMMA FREQUENCY. In other words, it is the summed frequency of all words containing the verbal stem, excluding its inflectional variants. There was no significant effect for OTHER FREQUENCY.

Given the general colinearity between the predictors, I further checked the predictive values of the most interesting nonsignificant predictors under the most inviting circumstances. Being included as the only predictor next to the control variables, the CUMULATIVE STEM FREQUENCY was still far from significance. The same is true for the POSITIONAL STEM FREQUENCY, the LEMMA FREQUENCY and SURFACE FREQUENCY. Significant in the absence of other predictors[6] was the NEIGHBORHOOD DENSITY (β = 0.00120, t(26701) = 2.36, p=0.0183), as well as the POSITIONAL NEIGHBORHOOD N1 (β = 0.00327, t(26701) = 3.29, p=0.0010).

Frequency counts play a role only in the facilitative effect of COHORT ENTROPY H2 (Van Son & Pols, 2003; Van Son & Van Santen, 2005). Note that the same effect explained a significant proportion of variance in the production onset latencies of deverbal adjectives (Chapter 3), but not in the latencies of noun-noun compounds (Chapter 2). The production of an inflected verb is affected by the size of, and the frequency distribution within, the cohort that shares the initial two phonemes. For verbs, this cohort contains all the inflected variants, leaving out the past participle form which carries a prefix. Given that there is no extreme variation in the size of the cohorts, it is mostly the distribution of frequencies within the

[6] with the exception of the control variables that had been significant in the final model

cohorts that determines their entropies. A cohort with little frequency variation speeds up the production latency of any member of the cohort. Taken together with corpora studies on fine phonetic detail (Van Son & Pols, 2003; Van Son & Van Santen, 2005; Kuperman, Pluymakers, Ernestus, & Baayen, subm) this finding suggests that the probability distribution in a production cohort affects both the length with which segments are produced within the word as well as the naming latency of the word.

Back to the question of how regularly inflected verbs are represented in the mental lexicon. The surface frequency of the inflected verb does not affect the frequency with which the verb is named, challenging model that assume a full listing of complex words (e.g., Butterworth, 1983). Neither does the frequency of the stem correlate with the naming onset latencies measured in the present study. Stem frequency effects are predicted in the decompositional model of Levelt et al (1999). They are likewise expected under the assumption that regular inflections are computed on the fly (Pinker, 1999). Even as the only predictor next to the control variables, no significant proportion of variance in the naming onset latencies was explained by either the frequency of the stem or by the frequency with which the surface form occurs.

Taken together, there is mixed evidence with respect to the decompositionality of morphologically complex words in speech production. While the naming onset latencies of transparent noun-noun suggest structured storage of constituting morphemes, the onset latencies of deverbal adjectives are predicted by the frequency of the verbal stem, arguing for decomposition. The onset latencies with which the regularly inflected verbs were named in the present study are neither related to the frequency of the stem not to the frequency of the complex word but reflect the frequency distribution within the cohort sharing the initial two phonemes. The following chapter is denoted to a joint analysis of naming onset latencies of the deverbal adjectives and the inflected verbs (which overlap in 76% in their verbal stems). The greater power of the joint analysis makes it possible to reveal more of the underlying structure that determines the onset latencies with which these two types of morphologically complex words are named.

References

Baayen, R.H. (2007) *Analyzing linguistic data. A practical introduction to statistics*. Cambridge University Press, in press.

Baayen, R.H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database* (CD-ROM), Linguistic data consortium. University of Philadelphia, Philadelphia, PA, Release 2.

Baayen, R.H., Tweedie, F.J., & Schreuder, R. (2002) The subjects as a simple random effect fallacy: Subject variability and morphological family effects in the mental lexicon, *Brain and Language*, *81*, 55–65.

Bates, D.M., & Sarkar, D. (2005). lme4: Linear mixed-effects models using S4 classes, R package version 0.9975-7.

Bates, D.M. (2005). Fitting linear mixed models in R, *R News*, *5*, 27-30.

Cholin, J., Levelt, W.J.M., & Schiller, N.O. (2006). Effects of syllable frequency in speech production. *Cognition*, *99*, 205-235.

Coltheart, M., Davelaar, E., Jonasson, J.T., & Besner, D (1977). Access to the internal lexicon. In Dornic, S. (Ed.), Attention and Performance VI. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Greenberg, J.H., & Jenkins, J.J. (1964). Studies in the psychological correlates of the sound system of American English. *Word*, *20*, 157-177.

Meyer, A.S., Roelofs, A., & Levelt, W.J.M. (2003). Word length effects in object naming: The role of a response criterion. *Journal of Memory and Language*, *48*, 131-147.

Oldfield R.C., & Wingfield A. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology*, *17*, 273-281.

Pinheiro, J.C., & Bates, D.M. (2000). Mixed-effects models in S and S-PLUS, *Statistics and Computing Series*, Springer-Verlag, New York, NY.

Appendix: Inflected verbs

Equally distributed over the seven types of Dutch verbal inflections ((1) 1st P. sg., present tense; (2) 2nd, 3rd P. sg., present tense; (3) 1st- 3rd P. pl., present tense / infinitive; (4) present participle; (5) 1st- 3rd P. sg., past tense; (6) 1st- 3rd P. pl., past tense; (7) past participle), the 126 items were assigned to 63 sets for presentation in the position-response association task.

sets 1-16		sets 17-32		sets 33-48		sets 49-63	
1	bouw (1) voelt (2)	17	trachtte (5) gezet (7)	33	schudde (5) draaiend (4)	49	remmen (3) kleeft (2)
2	kookten (6) staken (3)	18	haten (3) gewist (7)	34	gehaald (7) minnen (3)	50	gekwetst (7) speelden (6)
3	gruwen (3) pakkend (4)	19	weifelt (2) plagen (3)	35	gunt (2) morsten (6)	51	klaagde (5) zaai (1)
4	tekent (2) zorgden (6)	20	reddend (4) jeukten (6)	36	geraakt (7) jankend (4)	52	stoppend (4) zoent (2)
5	durend (4) merkt (2)	21	hapte (5) gemist (7)	37	toonden (6) smaken (3)	53	lenen (3) sticht (1)
6	ren (1) gehuwd (7)	22	toveren (3) snauw (1)	38	tasten (3) zwetend (4)	54	pruilden (6) lachte (5)
7	dweepten (6) maak (1)	23	tobbend (4) spoelde (5)	39	hijgden (6) zaagt (2)	55	hopend (4) wenden (3)
8	broeden (3) weerde (5)	24	gehoord (7) zoemde (5)	40	weiger (1) plaatst (2)	56	geacht (7) volg (1)
9	walgen (3) rekte (5)	25	scheur (1) brommen (3)	41	beefde (5) vluchtten (6)	57	pestte (5) stelt (2)
10	gestoord (7) zweven (3)	26	bakt (2) voegen (3)	42	scheiden (3) daalden (6)	58	twijfelt (2) geschilderd (7)
11	huil (1) wensten (6)	27	tel (1) zwierde (5)	43	regent (2) kneed (1)	59	getild (7) schroomden (6)
12	geruimd (7) droomt (2)	28	vulden (6) plooiend (4)	44	schaatst (2) deerde (5)	60	voedde (5) smeert (2)
13	leer (1) soezend (4)	29	schaadden (6) gedeeld (7)	45	rillend (4) broei (1)	61	veeg (1) schrapend (4)
14	brandend (4) gevloeid (7)	30	duld (1) piekte (5)	46	wakend (4) gepronkt (7)	62	beuzelden (6) vloekend (4)
15	huichelt (2) werkend (4)	31	rustte (5) norren (3)	47	woelde (5) pleegden (6)	63	krakend (4) stuitte (5)
16	straften (6) zeur (1)	32	leeft (2) gespaard (7)	48	gevreesd (7) plak (1)		

A JOINT ANALYSIS OF THE DEVERBAL ADJECTIVES AND INFLECTED VERBS

5.1 Revealing more of the underlying structure

The deverbal adjectives[1] studied in Chapter 3 and the inflected verbs studied in Chapter 4 have in common that they both contain verb stems. With respect to the specific verb stems, the two item pools overlap in 76%. I created this overlap on purpose to allow for additional comparative analyses of variables which influence the production of the two kinds of morphologically complex words. The overlap between the item pools is incomplete, because the pool of deverbal adjectives contains about one fourth of irregular verb stems, which were replaced by additional regular stems in the inflectional study (see 4.2.1 for the reasons to restrict Experiment 5 to regular verbs).

In parallel to the regression analysis of the joint data of four experiments reported in Chapter 2, I conducted a joint analysis of the data presented in Chapters 3 and 4. The two data sets were merged and a mixed-effects regression analysis with subject and item as crossed random effects was carried out (Pinheiro & Bates, 2000; Baayen, Tweedie, & Schreuder, 2002; Bates, 2005; Bates & Sarkar, 2005; Baayen, in press). While the joint analysis started fresh from the raw data, a number of latencies had to be excluded prior to the analysis. These latencies are identical to those excluded prior to the individual analyses in

[1] Dutch derivational morphology knows nine variants of input - output relations, because all word classes (noun, adjective, verb) can be derived from one another. Deverbal adjectives are the specific type of derivational morphology in which adjectives are derived from verbal stems through suffixation.

Chapters 3 and 4. Taking both data sets together, I excluded all latencies of two participants and five items (*zweterig; zweverig*; *gruwen voedde*, *gepronkt*)[2]. In addition, I excluded a total of 2266 (4%) specific trials, each of which qualified as one of the following: time out trial (latency >1500 ms), hesitation, wrong naming or voice key error. Of the original 57500 latencies measured, 54084 (94%) entered the analysis.

Following a stepwise variable selection procedure, model criticism led to the removal of an additional 2% of data points with standardized residuals exceeding 2.5. The model that was refitted to the data incorporated random intercepts for word stem (STD = 0.038) and for subject (STD = 0.122), and the residual error (STD = 0.225) and revealed the fixed-effects structure presented in Table 5.1.

As before, naming onset latencies slowed with REPETITION ($\beta = 0.0057$, $t(53181) = 11.17$, $p < 0.0001$), and PLOSIVE-initial words elicited longer onset latencies than words beginning with non-plosives. These main effects were modulated by an interaction. The repetition effect was strongest for fricatives (f) compared to plosives (p) and other initial segments (o). Furthermore, words with unvoiced initial segments were named faster than words with VOICED initial segments ($\beta = -0.0171$, $t(53181) = -2.36$, $p = 0.0183$). As before, we interpret these phonological effects as artifacts of the voice key (e.g., Kessler, Treiman, & Mullennix, 2002, see 3.4 for a discussion of the direction of the effect).

Word formation TYPE (inflection versus derivation) interacted with the INFLECTIONAL ENTROPY of the verbal stem. Inflected words elicited shorter onset latencies than derived words ($\beta = -0.1088$, $t(53181) = -2.58$, $p = 0.0099$), while only for the inflections, there was an inhibitory effect of INFLECTIONAL ENTROPY ($\beta = 0.0430$, $t(53181) = 3.63$, $p = 0.0003$ for the interaction, net effect of INFLECTIONAL ENTROPY: $\beta = -0.0085 + 0.0430 = 0.0345$).

There was a facilitatory effect of the Cohort Entropy H2 ($\beta = -0.0054$, $t(53181) = -2.80$, $p = 0.0051$) and an inhibitory effect of the positional neighborhood size for the initial phoneme ($\beta = 0.0018$, $t(53181) = 2.06$, $p = 0.0394$).

[2] For a detailed description on why these latencies were taken out prior to the analysis, the reader is referred to the subsections 3.3 and 4.3.

Table 5.1: Fixed effects of the multilevel regression model with subject and stem as crossed random effects (DF=53181). The final model contains five numerical predictors (REPETITION, INFLECTIONAL ENTROPY, COHORT ENTROPY H2, POSITIONAL ENTROPY N1, and LEMMA FREQUENCY), and five factors (levels in parentheses, with the reference level in italics[3]: PLOSIVE (*fricative*, other, plosive), VOICED (*voiced*, unvoiced), TYPE (*derivation*, inflection), PREFIXED (*unprefixed*, prefixed), ge-FREE SET (*non-ge-free set*, ge-free set)).

		Estimate	Std Error	t-value	p-value
INTERCEPT		6.0305	0.0301	200.51	0.0000
REPETITION		0.0057	0.0005	11.17	0.0000
PLOSIVE	other	0.0125	0.0106	1.18	0.2380
	plosive	0.0379	0.0083	4.59	0.0000
REPETITION *by* PLOSIVE	other	-0.0026	0.0009	-2.92	0.0035
	plosive	-0.0016	0.0008	-2.04	0.0414
VOICED	unvoiced	-0.0171	0.0072	-2.36	0.0183
TYPE	inflection	-0.1088	0.0422	-2.58	0.0099
INFLECTIONAL ENTROPY		-0.0085	0.0067	-1.26	0.2077
INFL. ENTROPY *by* TYPE	inflection	0.0430	0.0119	3.63	0.0003
COHORT ENTROPY H2		-0.0054	0.0019	-2.80	0.0051
POSITIONAL NEIGHBORHOOD N1		0.0018	0.0009	2.06	0.0394
PREFIXED	prefixed	0.0613	0.0113	5.44	0.0000
ge-FREE SET	ge-free set	0.0750	0.0284	2.65	0.0081
LEMMA FREQUENCY	linear	0.0035	0.0042	0.83	0.4065
	quadratic	0.0000	0.0004	0.04	0.9681
ge-FREE SET *by* LEMMA FREQUENCY	linear	-0.0222	0.0089	-2.51	0.0121
	quadratic	0.0015	0.0007	2.19	0.0285

[3] In the case of factorial variables, one level is generally modeled to lie on the intercept. The table, therefore, only lists the adjustment(s) for the other level(s).

Most items entered in this joint analysis carry their stem in initial position, followed by a suffix. Only one of the sixteen conditions, the past participle inflection, carries a prefix (*ge-*). The subset of PREFIXED words elicited longer response onset latencies (β = 0.0613, t(53181) = 5.44, p < 0.0001) than the complement set of unprefixed words. However, the effect of shorter onset latencies for unprefixed words only holds within those item sets that contain both a prefixed and an unprefixed form (non-ge-free sets), because there is a complementary inhibitory effect for items in ge-FREE SETS (β = 0.07501, t(53181) = 2.65, p = 0.0081). In other words, only in the presence of a prefixed item set partner, unprefixed forms are produced significantly faster.

Within the ge-FREE SETS, the LEMMA Frequency (specific to verbs, excluding homonymous nouns), emerged as a non-linear effect with shortest naming onset latencies for medium lemma frequencies (for the linear component: β = -0.0222, t(53181) = -2.51, p = 0.0121; for the quadratic component of the interaction: β = 0.0015, t(53181) = 2.19, p = 0.0285).

5.1.1 The partial effects for a subsets of items

Because of the complexity of this joint model, understanding is not enhanced by graphically displaying all partial effects and interactions as done in the previous chapters. It is, however, possible to graphically display the partial effects as they occur in a subset of the data.

Figure 5.1 presents the partial effects for those inflectional items, which had been presented in ge-FREE SETS[4]. In the following discussion of these partial effects displayed, I compare them both to the partial effects in other subsets of items as well as to effects found in analyses which were presented earlier in this thesis.

For the inflected verbs produced in ge-FREE SETS, Panel 1 of Figure 5.1 pictures the overall inhibitory effect of REPETITION (over all phonetic classes of the onset phonemes). The elongation of the production onset latencies within the individual test phases (from the first production of an item to its tenth production) is

[4] Excluded are all 36 inflectional items that either carry a prefix themselves or were tested with an item set partner that carries a prefix. Separate lists of all items produced in ge-FREE-SETS and items produced in non-ge-free sets are provided in the Appendices A and B.

a reliable, most likely task-specific, effect that occurred in all the position-response association learning task experiments conducted in the present study.

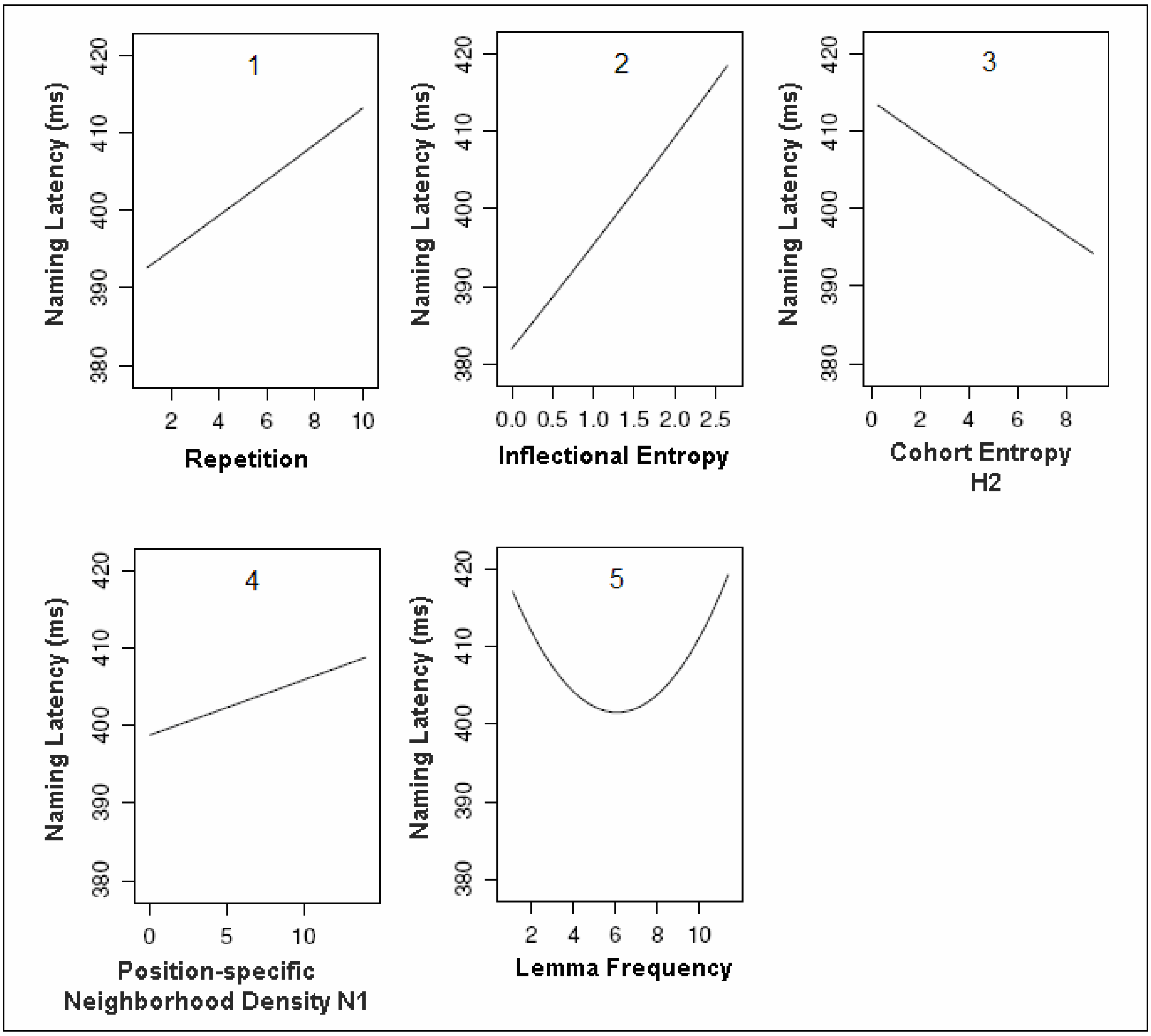

Figure 5.1: The partial effects for the subset of the 90 inflected verbs, which had been produced in ge-FREE SETS in the position-learning association task.

Panel 2 of Figure 5.1 shows the inhibitory effect of INFLECTIONAL ENTROPY as it occurs for the subset of inflected verbs produced in ge-FREE SETS. INFLECTIONAL ENTROPY had not emerged as a significant factor in the analysis of (all) the inflected verbs in Chapter 3. Analyzing the merged data set increases the statistical power and makes it possible to reveal more of the underlying structure that influences the production onset latencies of these morphologically complex words.

The facilitatory effect of the COHORT ENTROPY H2 (Panel 3) is familiar from both Chapter 3 ((all) inflected verbs) and Chapter 4 (the deverbal adjectives). With the transparent noun-noun compounds studied in Chapter 2, it was not the entropy of a positional cohort that facilitated production onset latencies, but rather both the derivational entropy of the initial constituent and the positional entropy of the head.

Panel 4 of Figure 5.1 pictures the inhibitory effect of the POSITIONAL NEIGHBORHOOD SIZE N1 for the subset of inflectional items produced in ge-FREE SETS. The same effect appeared in the analysis of the deverbal adjectives in Chapter 4 and is significant over all items in the merged model (see Table 5.1).

The analysis further revealed a non-linear effect of LEMMA FREQUENCY within the ge-FREE SETS (pictured in Panel 5). With shortest onset latencies for medium frequencies, this non-linear effect is very similar to the COMPOUND FREQUENCY effect presented in Chapter 2.

Note that the LEMMA FREQUENCY is computed over verbs only. It is the sum of frequencies of the inflectional variants of a verb. In Dutch, as in English, for some verbs the first person singular forms are homophonous to singular nouns (such as in *draai, teken, wens*). When only the frequencies of the inflectional variants are added up, the resulting sum of frequencies is a significant predictor for the latency in which an inflected verb is produced. When the frequency of the homophonous noun is added up along with the frequencies of the verb's inflectional variants, the resulting sum is not predictive. Note that besides the naming of single digit numbers as distracter trials, all items produced in the inflectional experiment were verb forms.

5.2 Differences and similarities in the production of deverbal adjectives and inflected verbs

The joint analysis reported in 5.2 is the most powerful of all mixed-effects modeling reported in this dissertation, because it is based on twice as many items (and data points) as the mixed-effects models presented in the previous chapters. A merging of the data sets of deverbal adjectives and inflected verbs is meaningful, because both types of morphologically complex words contain verbal stems. Tripling the power by merging also the compound data is not meaningful, because the chosen noun-noun compounds differ too much both in their structure and in their set of predictors. Thanks to its greater power, the joint analysis was able to not only reveal some effects that had not appeared in the separate analysis but also to compare how specific predictors affect the production onset latencies of these two types of morphologically complex words.

In the following, I discuss five aspects of the joint model in more detail. First, is the production of an inflected verb form easier than the production of a deverbal adjective? Second, what to think of the disadvantage for prefixed forms. Third, why is there a paradigmatic effect of inflectional entropy that is predictive for inflected verbs, but not for deverbal adjectives? Fourth, what do inhibitory effects of the positional neighborhood size N1 imply? And finally, how to look at the non-linear effect of the lemma frequency of the morphologically complex word.

5.2.1 Shorter naming onset latencies for inflected verbs than for deverbal adjectives

Inflected verbs were produced with shorter onset latencies than deverbal adjectives. Note, that these two types of morphologically complex words were tested in separate experiments with different participants. Under such circumstances, the faster production onset latencies for the inflected verbs could be attributed to faster responding participants in that experiment. To examine whether inflected verbs are produced with shorter onset latencies than deverbal adjectives, both types need to be tested within the same experiment using the same participants. The decision to test them in separate experiments was based on several reasons, among which a better comparison to the compound experiments, the total length of the experiment, and most prominent, the avoidance of repeated verb stems given their 76% overlap in the experiments. While the joint analysis of the two experiments is fruitful with respect to all other factors and numerical predictors, the confounding clearly restricts the interpretation of a main effect of TYPE (inflection, derivation).

Nevertheless, I want to discuss some explanations of why the onset latencies for inflected verbs could indeed be shorter than the onset latencies for deverbal adjectives. A first explanation might point to the fact that the production of inflections systematically involves the planning of only one phonological word, while some of the deverbal adjectives contain suffixes that are phonological words (such as -*achtig* and -*loos*). The planning of two phonological words might take longer than the planning of just one phonological word resulting in significantly shorter onset latencies for inflected verbs. If this explanation holds, the onset latencies should also differ for those deverbal adjectives that do contain a phonological-word suffix and those that do not. To test this hypothesis, I included a factor PHONOLOGICAL-WORD SUFFIX in the analysis of the deverbal adjectives. This factor was not significant.

One might further suspect that a difference in the onset latencies for the two types of morphologically complex words is caused by a difference in word length. This explanation doesn't hold. First, the difference in phonological word length between the item pools is small (means are 6.9 for the deverbal adjectives, and 5.4 for the inflected verbs). Second, phonological word length was taken along as a predictor in all analyses conducted in this thesis, but never explained a significant proportion of variance in a final model.

The item pools of deverbal adjectives and inflected verbs do, however, differ with respect to their average LEMMA FREQUENCIES (means and standard deviations in logarithmic scale: 4.2 (1.5), for the deverbal adjectives, and 6.8 (1.7) for the inflected verbs). As with the above mentioned variables, a predictor can only be assumed to cause a difference in the naming onset latencies between the groups, if that predictor is found to affect the latencies also within the groups. The LEMMA FREQUENCY was a significant predictor only for the naming onset latencies of the inflected verbs. Even under the most optimal statistical circumstances (when it was the only predictor included in the model next to the control variables), LEMMA FREQUENCY was not significant at all for the deverbal adjectives (see 3.4).

Another possible explanation is that the production of a regular inflected verb is generally an easier task than the production of a deverbal adjective, because the former is more regular.

5.2.2 A disadvantage for prefixed forms in mixed pairs

The only prefixed forms present in these two experiments are past participles. Prefixed words starting with *ge*- differ from unprefixed words in several aspects.

First, past participles contain two affixes (prefix *ge-* and suffix *-t*), rather than one. Second, given that all past participles forms share their initial diphone, one might suspect differences with respect to bigger cohorts for prefixed inflections. Next to the distribution of relative frequencies within a cohort, the size of a cohort influences its entropy. In our sample, the past participles differed from the other inflectional variants with respect to their cohort entropies. The mean logged cohort entropies H1, H2, H3 were 9.1, 9.1, 5.4 for the prefixed inflections versus 7.5, 5.4, 3.8 for the unprefixed inflections. The analysis revealed, however, that a higher cohort entropy H2 is negatively correlated with naming onset latencies, both for the deverbal adjectives and the inflected verbs. As a third difference, past participles begin with an unstressed, rather than a stressed syllable. Their longer naming onset latencies are in line with recent findings by Schiller, Fikkert, and Levelt (2004). Using picture naming, they observed longer onset latencies for stress-initial targets as opposed to stress-final targets. Finally, only one subtype of inflected verbs and none of the deverbal adjective carried a prefix, which might have created a kind of oddball out status for prefixed items in the experiment. To study whether the presence of a prefix generally elongates production onset latencies in morphologically complex words, a higher proportion of prefixes as well as different inflectional and derivational[5] prefixes should be included.

5.2.3 The stem's inflectional entropy is predictive for inflected verbs, but not for deverbal adjectives

The joint analysis reveals that the naming latency of an inflected verb form is affected by the verb's INFLECTIONAL ENTROPY. That is, the onset latencies with which inflected forms such as *miste* or *gemist* are named are a reflection of the entropy over the verb's inflectional paradigm (the frequency distribution of the inflectional variants: *mis*, *mist*, *missen*, *missend*, *miste*, *misten*, and *gemist*). In a deverbal adjective, however, the INFLECTIONAL ENTROPY of the verbal stem is no significant predictor for the latency with which the adjective (e.g. *misbaar*) is produced.

[5] While deverbal adjectives are restricted to suffixation, five of the nine types of Dutch derivations can be formed through either prefixation or suffixation (see 3.1).

This observation poses two questions. First, why is the INFLECTIONAL ENTROPY of a verb reflected in the latency with which an inflectional variant of that verb is named but not in the latency with which a derivation is named that contains the verb as stem? Remember, that the two item pools overlapped in 76% with respect to the verb stems.

While all items produced in the inflectional experiment were inflected verb forms, all items produced in the deverbal adjective experiment contained the stems in their uninflected forms. The naming of a deverbal adjective does not involve an inflection of the verbal stem. The actual production of inflected verb forms leads to a strong activation of the verbs' inflectional paradigms and to an effect of INFLECTIONAL ENTROPY. The verbal stems' inflectional paradigms play no role in the naming of deverbal adjectives. Note, that each deverbal adjective (e.g. *misbaar*) has an inflectional entropy of its own, computed over its singular and plural form. This entropy was no significant predictor of the naming onset latencies of the deverbal adjectives, all of which had been produced in their singular forms. Obviously INFLECTIONAL ENTROPY can be observed only when the inflectional paradigm is indeed relevant in performing the task.

Second, why is the effect of INFLECTIONAL ENTROPY inhibitory? The INFLECTIONAL ENTROPY of a verb is highest, when its inflectional variants are produced equally frequently. Under these circumstances, the production of a specific inflectional variant seems the hardest. Note, that there is just one entropy-value for all variants in a given inflectional paradigm, affecting the onset latencies with which higher- and lower-frequency variants are named in the same way.

An effect of inflectional entropy in production is not trivial because it suggests an influence of paradigmatic relations in the mental lexicon on the production of a selected word form. According to Levelt, Roelofs, and Meyer (1999), the spreading of activation is restricted to the levels of conceptualization and lemma selection. Once a lemma is selected, it directly activates its word form and no other word forms receive activation. The position-response association learning task I used in the present study taps in at the level of word form encoding. Next to other paradigmatic effects, the effect of INFLECTIONAL ENTROPY strongly suggests the co-activation of other word forms during production.

While an effect of INFLECTIONAL ENTROPY could be explained by an extension of the WEAVER++ model that assumes word-form representations for inflectional variants, all of which would be activated by a selected lemma, such architecture could still not explain the paradigmatic effects of neighborhood density and cohort entropy. The diversity of paradigmatic effects strongly suggests that word forms are in general connected to one-another, with stronger links within closer paradigms.

5.2.4 Inhibitory effects of the positional neighborhood N1

In Chapter 2, I had included the variable neighborhood density in the analysis of the compound naming onset latencies, to see whether I could replicate recent findings of neighborhood effects in production (Vitevitch, 2002, 2006). The first analysis of covariance reported in Chapter 2 suggested a nonlinear effect of neighborhood density. However, as the speech signal unfolds from initial to final phoneme, a whole-word count of neighborhood density seemed inadequate for production (and comprehension) research. In the later analyses, I therefore also included position-specific neighborhoods and found inhibitory effects of N1, both in the individual analysis of the deverbal adjective onset latencies and as an overall effect in the joint analysis presented earlier in this chapter.

If the mental lexicon contains many words that differ from the to-be-produced word in only the initial phoneme, production gets hard. In other words, it is difficult to produce a word that has many rhyme neighbors. Sevald and Dell (1994) had participants produce as many repetitions of sequences of overlapping words as possible within a given time. They found that it was generally easier to produce a sequence of rhyme words (such as *pick, tick*) than to produce a sequence of cohorts (such as *pick, pin*). While overlap is generally facilitative, there is an additional inhibitory component in the overlap of the initial phoneme. Sevald et al (1994) explain this inhibitory effect by the sequential cuing model, which assumes that, in such sequences, shared segments miscue the production of later sounds. Miscuing can happen in sequences such as *pin, pick*, but not in sequences such as *pick, tick*. The inhibitory effect of the positional neighborhood size of the initial phoneme found in the naming onset latencies of the deverbal adjectives and inflected verbs might add an explanation of why it is easier to repeat sequences such as *pick, tick* than it is to repeat sequences such as *pick, pin*. The positional neighborhood size effect N1 suggests that during the activation of a given word form, its rhyme neighbors are co-activated. When it is the task to produce an isolated word (as done in this dissertation), the co-activation elongates the production latency of the intended word form. When, however, it is the task to produce and repeat a sequence of rhyme neighbors (as done in Sevald et al, 1994), the co-activation is also a pre-activation that shortens the production onset latencies. Therefore, rhyme neighbors might benefit production only when they are actually produced. When they are not produced, they make the selection of the intended word form harder.

5.2.5 Non-linear effects for the lemma frequency of the morphologically complex word

For the inflected verbs, there is a non-linear effect of the LEMMA FREQUENCY of the complex word with the same shape as the COMPOUND FREQUENCY effect reported in Chapter 2. There is facilitation within the lower frequency range that turns into inhibition in the higher frequency range. At the time the Proceedings of the National Academy of Science Paper was written, I and my co-authors were cautious as to whether the trend of inhibition might be an artifact created by the model (using a quadratic polynomial), thereby actually covering a floor-effect. The same caution is appropriate when looking at the non-linear effect of LEMMA FREQUENCY presented above, though there is a more pronounced inhibition with the inflected verbs (there is relatively less facilitation in the linear component). Even if modeling has forced these LEMMA FREQUENCY effects of the complex words to appear more evenly curved than they actually are, something significant seems to be replicated here. Like with the compounds, the LEMMA FREQUENCY of an inflected verb co-affects its naming latency, with shortest onset latencies for medium frequencies.

When trying to understand this effect, it seems important to again differentiate between the LEMMA FREQUENCY of the complex word and the SURFACE FREQUENCY of the complex word. Crucially, it is an effect of the frequency of the lemma, which is the summed frequency of the inflectional variants of a word. The

SURFACE FREQUENCY, the frequency of occurrence of the articulated word form, was far from significance in all models presented in this thesis. In their localization of (morphologically simplex) word frequency effects in speech production, Jescheniak and Levelt (1994) had found an effect of lemma frequency, which was weaker than the effect of word form frequency and diminished quickly over repetition. The position association learning task used in the present study is clearly form based and the effects for the sum of frequencies of the inflectional variants are stable over repetition. Though, at this point, I can not provide a sound explanation of how these effects might arise, I want to speculate on it. Assume that inflectional variants have their own word form representations (see 5.2.3), all of which are linked to one-another and activated by a selected lemma. Imagine two low-frequent inflected verbs *(a)* and *(b)*, belonging to the inflectional paradigms *A*, and *B*, respectively. While all variants in paradigm *A* are low-frequent, paradigm *B* contains variants of higher frequency. Verb *(b)* benefits more from co-activation in its paradigm than verb *(a)* does, resulting in shorter naming onset latencies for higher LEMMA FREQUENCIES. Why, however, does facilitation turn into inhibition in the higher range of LEMMA FREQUENCIES? It seems unlikely that a high summed frequency of the inflectional variants within a paradigm is beneficial up to a certain degree after which it causes competition. It seems more likely that the co-activation in the inflectional paradigm is in general beneficial (potentially reaching a ceiling), but that something else is different about high-frequency inflectional variants. Higher frequency might lead to more well-established representations (e.g., Stemberger & MacWhinney, 1986), even in perfectly composable regular inflections. If an inflectional paradigm contains representations of high-frequency variants with tight links between their constituting morphemes, co-activation could be less beneficial than in a paradigm, in which the constituting morphemes are more distinct. Note that this effect would affect the whole inflectional paradigm as it is visible in the effect of LEMMA FREQUENCY. It would be interesting to study, whether high-frequency inflected verbs are produced less incrementally than low-frequency inflected verbs.

References

Baayen, R.H. (2007) *Analyzing linguistic data. A practical introduction to statistics*. Cambridge University Press, in press.

Baayen, R.H., Tweedie, F.J., & Schreuder, R. (2002). The subjects as a simple random effect fallacy: Subject variability and morphological family effects in the mental lexicon, *Brain and Language*, *81*, 55–65.

Bates, D.M. (2005). Fitting linear mixed models in R, *R News*, *5*, 27-30.

Bates, D.M., & Sarkar, D. (2005). lme4: Linear mixed-effects models using S4 classes, R package version 0.9975-7.

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

Kessler, B., Treiman, R., & Mullennix, J. (2002). Phonetic biases in voice key response time measurements. *Journal of Memory and Language*, *47*, 145-171.

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

Pinheiro, J.C., & Bates, D.M. (2000). Mixed-effects models in S and S-PLUS, *Statistics and Computing Series*, Springer-Verlag, New York, NY.

Schiller, N. O., Fikkert, P., & Levelt, C. C. (2004). Stress priming in picture naming: An SOA study. Brain and Language, 90, 231-240.

Sevald, C.A., & Dell, G.S. (1994). The sequential cuing effect in speech production. *Cognition*, *53*, 91-127.

Stemberger, J., & MacWhinney, B. (1986). Frequency and the lexical storage of regularly inflected forms. *Memory and Cognition*, *14*, 17-26.

Vitevitch, M.S. (2002). The influence of phonological similarity neighborhoods on speech production. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *28*, 735-747.

Vitevitch, M.S., & Stamer, M.K. (2006). The curious case of competition in Spanish speech production. *Language & Cognitive Processes, 21,* 760-770.

Appendix: The merged data set

The items are listed in alphabetical order. For details on the assignments to item sets and for details on the morphological subtypes, the reader is referred to the Appendices of the Chapters 3 and 4.

a) Deverbal adjectives (124)

beuzelachtig, beverig, brandbaar, breekbaar, broeds, broeierig, brommerig, buigzaam, deelbaar, deerlijk, denkelijk, draaglijk, draaibaar, drinkbaar, dromerig, druilerig, duldzaam, duurzaam, dweperig, flossig, grijpbaar, gruwelijk, haalbaar, handzaam, hangerig, happig, hatelijk, hebberig, heuglijk, hijgerig, hoorbaar, hopelijk, houdbaar, huichelachtig, huilerig, huwbaar, jankerig, jeukerig, kittelachtig, klaaglijk, kleverig, kloterig, kneedbaar, knorrig, krakerig, kruiperig, kwetsbaar, laakbaar, lacherig, leefbaar, leerzaam, leesbaar, lijdzaam, meetbaar, merkelijk, minzaam, misbaar, morsig, pesterig, piekerig, plaatselijk, plagerig, plakkerig, plooibaar, pronkerig, pruilerig, raadzaam, reddeloos, regenachtig, rekbaar, rillerig, roezig, schadelijk, scheidbaar, schilderachtig, schraperig, schrikkelijk, schromelijk, slaperig, sleets, smakelijk, smeerbaar, snauwerig, soezerig, spaarzaam, speels, splijtbaar, springerig, stellig, sterfelijk, strafbaar, strijdbaar, tastbaar, tekenachtig, telbaar, tobberig, toonbaar, toverachtig, twijfelachtig, vangbaar, vatbaar, vindbaar, vloeibaar, vluchtig, voedzaam, voegzaam, voelbaar, volgzaam, vreselijk, waaks, warrigweerloos, weifelachtig, weigerachtig, wendbaar, wenselijk, werkzaam, willig, woelig, zorgzaam, zweterig[6], zweverig[6], zwierig, zwijgzaam

b) Inflected verbs tested in ge-free sets (90)

bakt, beefde, beuzelden, bouw, broeden, broei, brommen, daalden, deerde, draaiend, duld, durend, dweepten, gruwen[6], gunt, hijgden, hopend, huichelt, huil, jeukten, klaagde, kleeft, kneed, kookten, krakend, lachte, leer, lenen, maak, merkt, morsten, knorren, pakkend, pestte, piekte, plaatst, plagen, pleegden, plooiend,

[6] These items were excluded from the analysis

pruilden, reddend, regent, rekte, remmen, rillend, rustte, schaatst, scheiden, scheur, schrapend, schudde, smaken, smeert, snauw, soezend, spoelde, staken, stelt, sticht, stoppend, straften, stuitte, tasten, tekent, tel, tobbend, toonden, toveren, veeg, vloekend, vluchtten, voedde[6], voegen, voelt, vulden, walgen, weerde, weifelt, weiger, wenden, wensten, werkend, woelde, zaagt, zaai, zeur, zoent, zorgden, zwetend, zwierde

c) Inflectional items tested in non-ge-free sets (36)

gehoord, brandend, droomt, geacht, gedeeld, gehaald, gehuwd, gekwetst, gemist, gepronkt[6], geraakt, geruimd, geschilderd, gespaard, gestoord, getild, gevloeid, gevreesd, gewist, gezet, hapte, haten, jankend, leeft, minnen, plak, ren, schaadden, schroomden speelden, trachtte, twijfelt, volg, wakend, zoemde, zweven

d) The regular verbal stems used in both Experiment 5 and 6

beuzel, beef, brand, broed, broei, brom, deel, deer, draai, droom, duld, duur, dweep, gruw, haal, hap, haat, hijg, hoor, hoop, huichel, huil, huw, jank, jeuk, klaag, kleef, kneed, knor, kraak, kwets, lach, leef, leer, merk, min, mis, mors, pest, plak, piek, plaats, plaag, plooi, pronk, pruil, red, regen, rek, ril, schaad, scheid, schilder, schraap, schroom, smaak, smeer, snauw, soes, spaar, speel, stel, straf, tast, teken, tel, tob, toon, tover, twijfel, vlucht, voed, vloei, voeg, voel, volg, vrees, waak, weer, weifel, weiger, wend, wens, werk, woel, zorg, zweet, zweef, zwier

SUMMARY

In a series of experiments, I studied the production of Dutch morphologically complex words, with special attention to effects of frequency, as a core issue in form retrieval. Models of speech production assume that on the way from conceptualization to articulation, word forms are accessed in the mental lexicon. Models disagree with respect to the way morphologically complex words are stored at the word form level. Full-listing approaches (e.g., Butterworth, 1983) assume that there is no qualitative difference between morphologically simplex and complex words. Fully decompositional models (e.g., Levelt, Roelofs, & Meyer, 1999) assume that the word form level contains only morphemes, and that these morphemes are accessed in the production of morphologically complex words. Clearly, the meaning of morphologically complex words must either be a predictable function of its components or it must be stored. Levelt et al (1999) assume that opaque complex words have their own lemmas, but that their constituting morphemes are nevertheless accessed at the word form level. While transparency seems to be a more gradient feature and the proportion of complex words with own lemmas might be high, there is converging evidence that the production of complex words involves access to the constituting morphemes irrespective of transparency (e.g., Zwitserlood, Bölte, & Dohmes, 2000; Roelofs & Baayen, 2002; Melinger, 2003; Dohmes, Zwitserlood, & Bölte, 2004; Gumnior, Bölte, & Zwitserlood, 2006). However, as stated by Butterworth (1983), it is possible in a full listing for all forms to have an internal structure marking morpheme boundaries.

The frequency with which a word occurs in a language has been found (e.g., Oldield & Wingfield, 1965) to correlate with its naming onset latency (i.e. the time passing from the moment a word is triggered to the start of articulation). Higher-frequency words have shorter naming onset latencies than lower-frequency words.

The word frequency effect has proven to be replicable in a wide range of tasks and has been attributed to the word form level (e.g., Jescheniak & Levelt, 1994). If morphologically complex words are fully listed at the word form level, the latency with which a morphologically complex word is produced should correlate with its frequency of occurrence as a complex word. If the production of morphologically complex word, however, involves access to the word forms of the constituting morphemes, naming onset latencies should correlate with the frequencies of occurrence of the constituting morphemes. Under the assumption of both decompositionality and incrementality, one predicts frequency effects for initial constituents only, because processing at the consecutive level starts as soon as the word form of the initial constituent has been accessed.

To measure the naming onset latencies of words, I made use of a position-response association task (Cholin, Levelt, & Schiller, 2006) in which participants first learned to associate a morphologically complex word with a visually marked position on a computer screen. Participants then had to produce the associated word in response to the appearance of the position mark and the speech onset latencies were measured. Each item was produced ten times by each participant. Testing 24 participants in each experiment, I measured 240 naming onset latencies of each item. In total, I collected naming onset latencies of 128 transparent noun-noun compounds (e.g., *wijnglas*), 124 deverbal adjectives (e.g., *grijpbaar*) and 126 regular inflected verbs (e.g., *draaiend*).

Chapter 2 presents experiments on the production of Dutch transparent noun-noun compounds. I independently varied the frequencies of the compound itself, its initial constituent (modifier), and its final constituent (head) to study their individual effects on the compound naming latency. Analyses of variance were conducted. In Experiment 1, pairs of compounds were matched in compound frequency and shared the modifier (e.g., *markt*vrouw* is as frequent as *markt*kraam*). The compound with the higher-frequency head (e.g., *vrouw*) was produced significantly faster (14ms) than the compound with the lower-frequency head (e.g., *kraam*). In Experiment 2, pairs of compounds sharing the head constituent were matched in compound frequency ((e.g., *steentijd* and *bronstijd*). Within the pairs, the compound with the higher-frequency modifier (e.g., *steen*) was produced significantly faster than the compound with the lower-frequency modifier (e.g., *brons*). The latency difference was on average 25 ms. In Experiment 3, pairs of compounds were matched in compounds frequency, while the frequency of both head and modifier was high in one compound and low in the other (e.g., *ringslang* and *roomsaus*). With an average difference of 27 ms, compounds containing the high-frequency constituents were named significantly faster than their matches. In Experiment 4, pairs of compounds were matched according to the frequencies of their modifiers and according to the frequencies of their heads, while there was a contrast in the frequency of the compound itself (e.g., *hoofdstad* and *huisvriend*). There was no significant difference in the naming latency. Taken together, the results suggest that the production latency of a compound is not influenced by its frequency of occurrence as a compound, but rather varies according to the frequencies of its constituting morphemes.

In addition to the separate ANOVAs, stepwise mixed-effects modeling was conduced on the joint data of the Experiments 1-4 (Pinheiro & Bates, 2000; Baayen, Tweedie, & Schreuder, 2002; Bates, 2005; Bates & Sarkar, 2005; Baayen, in press). The results challenge the assumption of full-listing as well as the assumption of full-decomposition and support the relevance of paradigmatic structure in the mental lexicon. For the modifier, there was an effect of the cumulative frequency of the constituent (as predicted by full decomposition) but there was an additional effect of its derivational entropy, which represents the frequency distribution of all words containing the modifier constituent. The naming latency of a compound is also affected by its head constituent. However, it is not the cumulative frequency with which the head appears in the lexicon but the entropy in the group of compounds sharing the same head constituent. The more often the head is head in other compounds and the less frequency variation there is within this compound family, the faster the compound is named. Though compounds contain two free morphemes, each of which can in principle take the place of a modifier or a head in a compound, access seems to be position-specific. In addition, there is a non-linear effect of the lemma frequency of the compound, with shortest naming onset latencies for medium frequencies. Taken together, the analyses presented in Chapter 2 suggest that the production of Dutch transparent noun-noun compounds involves access to the word forms of their constituting morphemes and is affected by paradigmatic relations between these morphemes.

With deverbal adjectives and inflected verbs, it was impossible to independently vary constituent and whole word frequencies within matched pairs. Therefore, I collected item sets that were distributed over a broad range of frequencies. Chapter 3 presents a stepwise mixed-effects modeling with subject and word as crossed random effects of the naming onset latencies of 124 Dutch deverbal adjectives. The results show that the latency with which a deverbal adjective is named is affected by the cumulative frequency of its verbal stem. The more often the verbal stem (e.g., *grijp*) occurs anywhere in the lexicon (independent or as part of a compounded, derived, or inflected word), the faster the deverbal adjective (e.g., *grijpbaar*) is named. The frequency of the deverbal adjective itself was not predictive of its naming latency, as were all other measures of frequency and entropy. The only compatible predictor was the positional frequency of the verbal stem, but the positional frequency of the stem is highly correlated with its cumulative frequency. The naming latency of the deverbal adjective was further affected by two phonological variables. The more words exist that differ from the deverbal adjective only in the initial phoneme, the longer it takes to name the adjective. While the existence of many rhyme neighbors elongates the naming latency, the existence of many other words that share the initial two phonemes, shortens the naming onset latencies. The onset latencies are shortest, when the cohort of words sharing the initial two phonemes is big and has a small variation in frequencies. These effects of neighborhood density and cohort entropy again suggest paradigmatic relations in the mental lexicon, and that the production of a word is affected by many other words in the lexicon. The finding that the production latency of a deverbal adjective is affected by the cumulative frequency of its verbal stem and unaffected by the frequency of occurrence of the deverbal adjective itself argues for decomposition and against full-listing.

Chapter 4 presents a stepwise mixed-effects modeling with subject and word as crossed random effects of the naming onset latencies of 126 regular inflected verbs in Dutch. The results show that the production of prefixed inflections (past participle forms) was significantly slower than the production of unprefixed forms. There was no effect of frequency or entropy related to either the constituting morphemes or the inflected verb itself, but there was a facilitative effect of the entropy over the cohort sharing the first two phonemes. The more words exist that start with the same two phonemes as the to-be-produced inflected verbs, the faster the verb can be named. The results presented in Chapter 4, neither fit the prediction of full listing nor do they fit the prediction of full decomposition, because there were no significant constituent of surface frequency effects. A joint analysis of the naming onset latencies of the deverbal adjectives and the inflected verbs revealed more of the underlying structure of the production of inflected verbs (Chapter 5).

The deverbal adjectives and inflected verbs were selected such as that there was a high overlap (76%) in the verbal stems used in the two studies. Chapter 5 presents a joint analysis of the naming onset latencies of the deverbal adjectives and inflected verbs to directly compare how particular variables influence the production of these two types of morphologically complex words. The facilitative effect of the cohort entropy H2 (found in both separate analysis) was significant over all items. The inhibitory effect of the position-specific neighborhood N1 (present in the separate analysis of the deverbal adjectives but not in the separate analysis of the inflected verbs) was also significant over all items. Both types of words are named the faster, the fewer rhyme neighbors they have and the more evenly frequent words exist which share their initial two phonemes. Only for the inflected verbs, there was an inhibitory effect of inflectional entropy. The higher the sum of frequencies of the inflectional variants of a verb, the longer it takes to prepare a specific inflectional variant. In the joint analysis, it became evident that prefixed verbs were not only produced slower compared to unprefixed words, but also affected the naming onset latencies of their unprefixed item set partners. For the subset of inflected verbs that neither carried a prefix nor had a prefixed item set partner, there was a non-linear effect of the lemma frequency of the inflected verbs. Yielding shortest latencies for medium frequencies, this effect seems to replicate the lemma frequency effect in transparent noun-noun compounds reported in Chapter 2).

Taken together, the naming onset latencies of the deverbal adjectives and the inflected verbs argue against full listing, as the frequency of occurrence of the complex word was far from significance in all analyses. For the deverbal adjectives, the results are in line with the assumption of decomposition, but there is no effect of the frequency of the verbal stem on the naming latency of regular inflected verbs. The non-linear effect of the lemma frequency of the inflected verb suggests that high-frequency inflectional variants might have more well-established representations and be produced less incrementally than variants of lower frequency. Neighborhood and cohort entropy effects reflect paradigmatic relations in the mental lexicon and suggest that all word forms influence the production of one word.

To summarize, the data of the present study suggest that the word form level does not contain full-listings or strictly separated morphemes but morphemes with links to other morphemes. Morphologically and phonolgically related forms influence the speed with which morphemes can be retrieved for the production of morphologically complex words.

SAMENVATTING

In een reeks van experimenten heb ik onderzoek gedaan naar de productie van morfologisch complexe woorden in het Nederlands en de rol die frequentie daarbij speelt. Spraakproductiemodellen gaan ervan uit dat bij het proces van conceptualisatie naar articulatie woordvormen worden opgehaald uit het mentale lexicon. Er zijn verschillende theorieën over de manier waarop morfologisch complexe woorden op het vormniveau zijn opgeslagen. Terwijl *full-listing* modellen (bv. Butterworth, 1983) voorstellen dat er geen kwalitatief onderscheid bestaat tussen morfologisch eenvoudige en morfologisch complexe woorden, gaat men er in een decompositioneel model vanuit dat het woordvormniveau alleen morfemen bevat, die ook gebruikt worden voor de productie van morfologisch complexe woorden (bijv. Levelt, Roelofs, & Meyer, 1999). Als de betekenis van een morfologisch complex woord niet kan worden afgeleid uit de betekenis van de individuele morfemen, is het noodzakelijk dat de betekenis apart is opgeslagen. Levelt et al. (1999) stellen voor dat opake complexe woorden een eigen lemma hebben maar, net als transparante complexe woorden, op het vormniveau uit losse morfemen worden opgebouwd. Hoewel transparantie een graduele eigenschap blijkt te zijn en het aantal complexe woorden met behoefte aan een eigen lemma daarom vermoedelijk hoog is, is er evidentie dat morfemen geactiveerd zijn wanneer morfologisch complexe woorden geproduceerd worden, onafhankelijk van hun mate van transparantie (bv. Zwitserlood, Bölte, & Dohmes, 2000; Roelofs & Baayen, 2002; Melinger, 2003; Dohmes, Zwitserlood, & Bölte, 2004; Gumnior, Bölte, & Zwitserlood, 2006). Zoals Butterworth (1983) aangeeft, zou het ook in een full-listing model mogelijk kunnen zijn dat woordvormen een interne morfeemstructuur hebben.

Onderzoek heeft aangetoond (bijv. Oldield & Wingfield, 1965) dat de frequentie van een woord (i.e., hoe vaak een woord gebruikt wordt) samenhangt met de spraak-onset-latentie van het woord (i.e., de tijd die verstrijkt vanaf het moment dat de spreker weet welk woord hij/zij wil noemen tot het begin van de articulatie). Hoogfrequente woorden hebben kortere spraak-onset-latenties dan laagfrequente woorden. Dit woordfrequentie-effect is in verschillende onderzoeksparadigmas gevonden en wordt aan het woordvormniveau toegeschreven (bijv. Jescheniak & Levelt, 1994).

Als morfologisch complexe woorden op het woordvormniveau als hele woorden opgeslagen zijn, dan zou de spraak-onset-latentie moeten samenhangen met de frequentie van het complexe woord. Maar als in plaats van een kant-en-klare woordvorm losse morfemen worden opgehaald, dan zou de latentie moeten samenhangen met de frequentie van die morfemen. Als we aannemen dat spraakproductie zowel decompositioneel als incrementeel verloopt, verwacht men alleen een frequentie effect van de eerste constituent te zien. In dat geval kunnen namelijk de volgende processen al starten, zodra de vorm van de eerste constituent is opgehaald.

Voor het meten van de spraak-onset-latenties van morfologisch complexe woorden, heb ik gebruik gemaakt van een *Symbool Positie Associatie Leertaak* (Cholin, Levelt, & Schiller, 2006). Daarbij leert een proefpersoon eerst om een woord met een visueel gemarkeerde positie op het beeldscherm te associëren, en produceert zij/hij later het geassocieerde woord als antwoord op het verschijnen van de visuele marker. Hierbij wordt de spraak-onset-latentie gemeten. Per proefpersoon heb ik voor elk woord tien latenties gemeten. Met 24 proefpersonen per experiment levert dat 240 latenties voor elk woord. In totaal heb ik de latenties van 128 transparante naamwoord-naamwoord samenstellingen (bijv. *wijnglas*), 124 deverbale adjectieven (bijv. *grijpbaar*) en 126 regelmatige werkwoorden (bijv. *draaiend*) kunnen verzamelen.

In Hoofdstuk 2 presenteer ik onderzoek naar de productie van Nederlandse transparante samenstellingen die uit twee zelfstandige naamwoorden bestaan. Om hun eigen bijdrage aan de spraak-onset-latenties te kunnen bestuderen, heb ik de frequenties van de samenstellingen en van de constituenten onafhankelijk van elkaar gevarieerd. Voor het eerste experiment heb ik paren van samenstellingen gezocht die, als geheel woord, even frequent zijn (bijv. *<u>markt</u>vrouw* komt even vaak voor als *marktkraam*) en die dezelfde modifier en daarom ook dezelfde modifier-frequentie hebben (hier: *markt*). Binnen dergelijke paren, werd de samenstelling met het frequente hoofd (bijv. *vrouw*) met een significant kortere latentie genoemd (gemiddeld 14 ms) dan de samenstelling met het minder frequentere hoofd (bijv *kraam*). Voor Experiment 2 heb ik paren van even frequente samenstellingen gezocht met eenzelfde hoofd ((bijv. *steentijd* and *bronstijd*). Weer was er een significant verschil in de latenties. De samenstelling met de frequentere modifier (bijv. *steen*) had een kortere latentie (gemiddeld 25 ms) dan de samenstelling met de minder frequente modifier (bijv. *brons*). In Experiment 3 waren de samenstellingen binnen elk paar even frequent, terwijl de modifier en het hoofd van de ene samenstelling hoger frequent waren dan die van de andere samenstelling (e.g., *ringslang* en *roomsaus*). Met gemiddeld 27 ms waren de spraak-onset-latenties van de samenstellingen met de frequentere morfemen significant korter dan de latenties van de even frequente samenstellingen met de minder frequente morfemen. De paren in Experiment 4 hadden zowel even frequente modifiers als even frequente hoofden, terwijl de ene samenstelling frequenter was dan de andere (e.g., *hoofdstad* en *huisvriend*). In dit experiment was er geen verschil in de latenties. Samengevat wijzen deze resultaten erop dat de spraak-onset-latentie van een transparante samenstelling niet samenhangt met de frequentie van de samenstelling zelf, maar met de frequentie van zijn morfemen.

Naast de aparte analyses heb ik alle in net genoemde experimenten verzamelde latenties middels een stepwise mixed-effects model geanalyseerd (Pinheiro & Bates, 2000; Baayen, Tweedie, & Schreuder, 2002; Bates, 2005; Bates & Sarkar, 2005; Baayen, in press). De resultaten hiervan zijn noch door een full-listing model noch door een strikt decompositioneel model gemakkelijk te verklaren en duiden op paradigmatische structuren in het mentale lexicon. Zoals een decompositoneel model voorspelt was de cumulatieve frequentie van het eerste morfeem (bijv. hoe vaak kom je het morfeem *markt* tegen in het Nederlands) een significante voorspeller voor de spraak-onset-latentie. Maar naast dit cumulatieve frequentie-effect, was er ook een significant effect voor de derivationele entropie van het eerste morfeem (een maat voor de frequentieverdeling binnen de groep complexe woorden waarin het morfeem *markt* voorkomt). Ook het tweede morfeem (het hoofd) had invloed op de latenties. De sterkste voorspeller was daarbij echter niet de absolute frequentie van het morfeem (bijv. *vrouw*) maar hoe veel samenstellingen er bestaan met hetzelfde hoofd (*vrouw*) en hoe de frequenties binnen die groep verdeeld zijn. Hoe groter de groep en hoe kleiner de variantie in de frequenties, des te korter was de spraak-onset-latentie. Hoewel samenstellingen uit vrije morfemen bestaan en een vrij morfeem in verschillende samenstellingen zowel modifier als hoofd kan zijn, blijkt de positie een belangrijke rol te spelen bij het ophaalproces op het woordvormniveau. Ten slotte, had ook de lemmafrequentie van de samenstelling een significant effect op de latenties. Samengevat wijzen de in Hoofdstuk 2 gepresenteerde resultaten erop dat voor de productie van Nederlandse transparante samenstellingen morfemen worden opgehaald onder invloed van paradigmatische relaties tussen de morfemen.

Voor het onderzoek naar de productie van deverbale adjectieven en vervoegde werkwoorden was het niet mogelijk om binnen paren de frequentie van de morfemen en van het complexe woord onafhankelijk te variëren. Daarom heb ik voor deze typen morfologisch complexe woorden itemsets samengesteld, die ten opzichte van belangrijke variabelen breed verdeeld zijn en hun invloed op de spraak-onset-latenties met behulp van stepwise mixed-effects modellen geanalyseerd. In het derde Hoofdstuk presenteer ik een analyse van de spraak-onset-latenties van 124 Nederlandse deverbaal adjectieven. De resultaten tonen aan dat de latentie, waarmee een deverbaal adjectief genoemd wordt, beïnvloed wordt door de frequentie van de stam. Hoe vaker de stam (bijv. *grijp*) voorkomt (zelfstandig of binnen een samenstelling, een derivatie, of een vervoeging), des te sneller begon de articulatie van het deverbale adjectief (bijv. *grijpbaar*). De frequentie van het adjectief zelf was geen voorspeller van de latentie en ook andere frequentie- en entropiematen waren geen geschikte voorspellers, met uitzondering van de positionele stamfrequentie, die echter hoog correleerde met de cumulatieve stamfrequentie. Ook fonologische maten hadden een invloed op de latenties. Hoe meer woorden enkel in het eerste foneem van het deverbaal adjectief verschilden (rijmburen), des te langer duurde het voordat het adjectief gearticuleerd kon worden. Hoe meer woorden er bestaan die met dezelfde twee fonemen beginnen, des te korter was de latentie, waarbij ook de frequentieverdeling binnen het cohort van belang was. De latenties waren het kortst voor grote cohorten met weinig variantie in de frequenties. Buurt- en cohorteffecten duiden op paradigmatische relaties in het mentale lexicon en op een invloed van alle woorden op de productie van een woord. Dat de spraak-onset-latentie van de adjectieven niet samenhangt met de frequentie waarmee het deverbale adjectief geproduceerd wordt, maar wel met de cumulatieve frequentie van de stam, spreekt tegen full-listing en voor decompositionaliteit.

In Hoofdstuk 4 heb ik de spraak-onset-latenties van 126 regelmatig vervoegde Nederlandse werkwoorden geanalyseerd. De resultaten laten zien dat de tijd voor het begin van de articulatie langer is voor geprefigeerde vormen dan voor ongeprefigeerde vormen. Net als bij de deverbale adjectieven, was er een significant effect voor de cohortentropie H2. Hoe meer woorden met dezelfde twee fonemen beginnen als het woord dat de proefpersoon noemt, des te korter is de latentie voor het begin van de articulatie. Omdat er geen significante frequentie effecten waren, sluiten de resultaten van Hoofdstuk 4 niet aan bij de voorspellingen van het full-listing model en evenmin bij de voorspellingen van het decompositionele model. Een gezamenlijke analyse van de in de Experimenten 5 en 6 verzamelde data (Hoofdstuk 5) leverde nieuwe informatie op over de productie van een vervoegd werkwoord.

In de gezamenlijke analyse van de spraak-onset-latenties van de deverbale adjectieven en de vervoegingen van de regelmatige werkwoorden heb ik onderzocht hoe bepaalde factoren de productie van deze twee typen morfologisch complexe woorden beïnvloeden. Het faciliterende effect van de cohortentropie H2 (gevonden in beide afzonderlijke analyses) was hier eveneens significant. Het inhiberende effect van de positiespecifieke buurt N1 (eerder alleen significant voor de adjectieven) was in deze analyse significant voor alle woorden. Hoe minder rijmburen ze hebben en hoe meer woorden met dezelfde twee fonemen beginnen, des te sneller worden ze genoemd. Alleen voor de vervoegingen was er een inhiberend effect van de inflectionele entropie. Hoe groter de som van de frequenties van de vervoegingen, hoe langer het duurt voordat één van de vervoegingen gearticuleerd kan worden. Verder heeft de gezamenlijke analyse duidelijk gemaakt dat geprefigeerde vormen niet alleen maar trager worden genoemd dan ongeprefigeerde vormen, maar dat ze in de gebruikte taak ook de latentie van hun setpartner beïnvloeden. Als er uitsluitend wordt gekeken naar de ongeprefigeerde vormen met een ongeprefigeerde setpartner, wordt een niet-lineair effect van de lemmafrequentie van het werkwoord zichtbaar. Met de kortste latenties voor de gemiddeld-frequente werkwoorden, lijkt dit effect op het lemmafrequentie-effect in Hoofdstuk 2.

Samengevat zijn de onderzoeksresultaten met betrekking tot de productie van de adjectieven en de vervoegingen niet gemakkelijk verklaarbaar met het full-listing model omdat de frequenties van de complexe woorden geen significant effect lieten zien. De resultaten met betrekking tot de deverbale adjectieven ondersteunen een decompositioneel model, terwijl er voor de vervoegde werkwoorden geen effect was voor de stamfrequentie. Het niet-lineare effect van de lemmafrequentie van de werkwoorden duidt erop dat hoogfrequente werkwoorden sterker gerepresenteerd zijn en minder incrementeel geproduceerd worden dan minder frequente werkwoorden. Buurt- en cohorteffecten weerspiegelen paradigmatische relaties in het mentale lexicon en wijzen erop dat alle woordvormen invloed hebben op de productie van een woord.

Samengevat duiden de resultaten van het onderzoek erop dat het mentale lexicon noch uit gehele vormen voor morfologisch complexe woorden noch uit losse morfemen bestaat, maar uit met elkaar verbonden morfemen. Morfologisch en fonologisch gerelateerde vormen beïnvloeden de snelheid waarmee de morfemen kunnen worden opgehaald voor de productie van morfologisch complexe woorden.

CURRICULUM VITAE

Heidrun Bien was born in Münster, Germany, on June 8, 1976. After graduation from the Gymnasium Arnoldium in Steinfurt in 1995, she studied Psychology at the Westfälische Wilhelms-Universität (WWU) in Münster. From 1998 to 1999 a Fulbright Scholarship enabled her to study Psychology at the University of Kansas, USA. During this time, she also worked as a research assistant at the Department for Human Development and Family Life. Back at the WWU, she continued her study while working as a research assistant in the psychology department of the WWU. In 2003 she received her MA and was awarded a scholarship from the German-Max-Planck-Gesellschaft to prepare her Ph.D. thesis at the Max Planck Institute for Psycholinguistics in Nijmegen, the Netherlands. Since August 2006 she has been working as a teacher and post-doctoral researcher in the psychology department of the WWU on the project *Neural and psychological correlates of phonological categories* funded by the Deutsche Forschungsgemeinschaft.

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 midsaggital articulography. *Niels O. Schiller*
3. Lexical access in the production of ellipsis and pronouns. *Bernadette M. 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 Oceanic language of Papua New Guinea. *Anna Margetts*
13. From speech to words. *Arie van der Lugt*
14. Simple and complex verbs in Jaminjung: A study of event categorisation in an Australian language. *Eva Schultze-Berndt*
15. Interpreting indefinites: An experimental study of children's language comprehension. *Irene Krämer*
16. Language-specific listening: The case of phonetic sequences. *Andrea Weber*

17. Moving eyes and naming objects. *Femke van der Meulen*
18. Analogy in morphology: The selection of linking elements in Dutch compounds. *Andrea Krott*
19. Morphology in speech comprehension. *Kerstin Mauth*
20. Morphological families in the mental lexicon. *Nivja H. de Jong*
21. Fixed expressions and the production of idioms. *Simone A. Sprenger*
22. The grammatical coding of postural semantics in Goemai (a West Chadic language of Nigeria). *Birgit Hellwig*
23. Paradigmatic structures in morphological processing: Computational and cross-linguistic experimental studies. *Fermín Moscoso del Prado Martín*
24. Contextual influences on spoken-word processing: An electrophysiological approach. *Daniëlle van den Brink*
25. Perceptual relevance of prevoicing in Dutch. *Petra M. van Alphen*
26. Syllables in speech production: Effects of syllable preparation and syllable frequency. *Joana Cholin*
27. Producing complex spoken numerals for time and space. *Marjolein Meeuwissen*
28. Morphology in auditory lexical processing: Sensitivity to fine phonetic detail and insensitivity to suffix reduction. *Rachèl J. J. K. Kemps*
29. At the same time...: The expression of simultaneity in learner varieties. *Barbara Schmiedtová*
30. A grammar of Jalonke argument structure. *Friederike Lüpke*
31. Agrammatic comprehension: An electrophysiological approach. *Marlies Wassenaar*
32. The structure and use of shape-based noun classes in Miraña (North West Amazon). *Frank Seifart*
33. Prosodically-conditioned detail in the recognition of spoken words. *Anne Pier Salverda*
34. Phonetic and lexical processing in a second language. *Mirjam Broersma*
35. Retrieving semantic and syntactic word properties. *Oliver Müller*
36. Lexically-guided perceptual learning in speech processing. *Frank Eisner*

37. Sensitivity to detailed acoustic information in word recognition. *Keren B. Shatzman*
38. The relationship between spoken word production and comprehension. *Rebecca Özdemir*
39. Disfluency: Interrupting speech and gesture. *Mandana Seyfeddinipur*
40. The acquisition of phonological structure: Distinguishing contrastive from non-contrastive variation. *Christiane Dietrich*
41. Cognitive cladistics and the relativity of spatial cognition. *Daniel B.M. Haun*
42. The acquisition of auditory categories. *Martijn Goudbeek*
43. Affix reduction in spoken Dutch. *Mark Pluymaekers*
44. Continuous-speech segmentation at the beginning of language acquisition: Electrophysiological evidence. *Valesca Kooijman*
45. Space and iconicity in German Sign Language (DGS). *Pamela Perniss*