Frequency effects in the production of Dutch deverbal adjectives and inflected verbs

Heidrun Bien[1], R. Harald Baayen[2], Willem J. M. Levelt[3]

[1] Institute for Psychology II, University of Muenster, Germany

[2] University of Alberta, Edmonton, Department of Linguistics, Canada T6G 2E5

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

Short title: Frequency effects in morphologically complex words

Address for correspondence:

Heidrun Bien

Institute for Psychology, Fliednerstrasse 21, 48149 Münster, Germany

Tel. +49 251 8331432

Fax: +49 251 8334104

E-mail: heidrun.bien@uni-muenster.de

Abstract

In two experiments, we studied the role of frequency information in the production of deverbal adjectives and inflected verbs in Dutch. Naming latencies were triggered in a position-response association task and analyzed using stepwise mixed-effects modeling, with subject and word as crossed random effects. The production latency of deverbal adjectives was affected by the cumulative frequencies of their verbal stems, arguing for decomposition and against full listing. However, for the inflected verbs, there was an inhibitory effect of inflectional entropy, and a non-linear effect of lemma frequency. Additional effects of position-specific neighborhood density and cohort entropy in both types of words underline the importance of paradigmatic relations in the mental lexicon. Taken together, the data suggest that the word-form level does neither contain full forms nor strictly separated morphemes, but rather morphemes with links to phonologically and – in case of inflected verbs – morphologically related word forms.

Acknowledgements: We are grateful to Pienie Zwitserlood for valuable discussion.

Introduction

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 level of word forms. Full-listing approaches (e.g., Butterworth, 1983; Janssen, Bi, & Caramazza, 2008) assume that there is no qualitative difference between morphologically simplex and complex words. Both have entries at the word form level. Fully decompositional models (e.g., Levelt, Roelofs, & Meyer, 1999; Levelt, 2001) propose that the word form level contains only morphemes, and that these morphemes are accessed in the production of a morphologically complex word. Since the meaning of morphologically complex words must either be stored or be a predictable function of the meaning of the constituents, Levelt et al. (1999) assume that opaque complex words have their own lemmas (an abstract form containing syntactic information, intermediate between the semantic level and that containing word-forms). At the word form level, all morphologically complex words (opaque or not) access their constituting morphemes. There is converging evidence that the production of complex words involves access to the constituent morphemes, irrespective of transparency (e.g., Zwitserlood, Bölte, & Dohmes, 2000; Ayala & Martin, 2002; Roelofs & Baayen, 2002; Melinger, 2003; Dohmes, Zwitserlood, & Bölte, 2004; Mondini et al., 2005; Gumnior, Bölte, & Zwitserlood, 2006; Koester, & Schiller, 2008; but see Janssen et al., 2008). However, as stated by Butterworth (1983), with full listing, it is possible for all forms to have an internal structure marking morpheme boundaries, and that it is by no means a tautology that everything available via a rule must be available only in that way.

High-frequency words tend to have shorter naming latencies than low-frequency words (e.g., Oldfield & Wingfield, 1965). This 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, but see Caramazza et al, 1998). If a morphologically complex word is fully listed at the word form level, its production latency should correlate with its frequency of occurrence as a complex word. If the production of a complex word, however, involves access to the word forms of the constituent morphemes, latencies should correlate with the frequencies of occurrence of these morphemes. Assuming both decompositionality and incrementality, frequency effects are expected for initial constituents only. Rightward incremental word-form encoding has been observed in production studies by Roelofs (1996) and Cholin, Schiller, and Levelt (2004), the former using Dutch verb-particle utterances, the latter using Dutch bisyllabic words. But there is also evidence against strict incrementality. Though the effect of the initial constituent was stronger, Bien, Levelt, and Baayen (2005) found frequency effects of both constituents in the production of Dutch noun-noun compounds. Studies on the acoustic realization of complex words (Koester, Gunter, Wagner, & Friederici, 2004; Kemps, Ernestus, Schroeder, Baayen, 2005; Kemps, Wurm, Ernestus, Schreuder, Baayen, 2005) suggest that the planning of the articulation of initial constituents is to some extent dependent on the presence of a second one. Stems appearing as the initial constituents of complex words (be it compounding, inflection or derivation), tend to have shorter durations and tend to be produced with a different intonation contour than the same stems pronounced in isolation.

Based on the full-listing versus full-decomposition debate, the main question is whether the naming latency of a morphologically complex word (e.g., *grijpbaar*, touchable) can be predicted by its own frequency of occurrence, or by the frequency of its first and / or second constituent (e.g., *grijp*, to touch; -*baar*, -able). Given the additional information revealed by multilevel analysis of covariance in a closely related study on Dutch compounds (Bien et al., 2005), considering just the frequencies of the word and its constituents may be an oversimplification. There, factorial analyses revealed frequency effects of both constituents, but not of the compound itself, suggesting composition during production. Additional stepwise regression analyses revealed a superior predictivity of frequency and entropy measures[1] that challenge full decomposition and suggest a role for paradigmatic structure in speech production. More precisely, it suggests a lexicon in which word forms are interlinked and organized in paradigms, such that they influence each-other in production. In the current study we will explore whether such a paradigmatic account (we will refer to this account as 'structured storage' as well) is viable for the production of derivations and regular inflections or whether the traditional full-listing or decomposition models are more viable options. In order to do so, we will examine the predictivity of morphological and frequency variables on word production latencies. For both inflected and derived words, strict decomposition predicts effects of constituent frequency, full-listing models predict effects of surface frequency only, and a paradigmatic account would be supported by effects of lemma and positional frequency as well as effects of entropy measures. We will explain these hypotheses more clearly in the following sections. In addition to this, we will also explore whether there are phonology-based connections between word forms by examining the predictivity of variables that tap into different levels of phonological similarity.

We analyze the naming latencies of a wide range of derived and inflected Dutch words, including the constituent and whole-word frequency variables, along with other predictors, in a stepwise mixed-effects modeling with Participant and Word as crossed random effects (e.g., Pinheiro & Bates, 2000; Baayen, Tweedie, &

---

[1] **Next to a u-shaped effect of compound frequency there were facilitative effects of both the Left and the Right Positional Entropy (i.e., Shannon's entropy in a set of compounds sharing the left (or right) constituent). We will explain these entropy measures further in the below.**

Schreuder, 2002; Bates, 2005; Bates & Maechler, 2009; Baayen, 2008; Baayen, Davidson and Bates, 2008). With this type of analysis, we hope to capture multifaceted effects in the production of complex words, involving form frequency as well as phonological variables and, of course, morphological variables.

Derivation and inflection can be defined as distinct processes of word formation. At the same time, a vast body of literature suggests more of a continuum with some operations falling in between the prototypical derivation and inflection. Affixation, the attachment of bound morphemes to stems, is the formal operation subserving both derivation and inflection. Derivation forms new words which can be of the same or a different class than their base word. Here, we study the production of Dutch deverbal adjectives (e.g., *lees-baar,* readable), one type of derivation that is restricted to suffixation. While some derivational affixes (such as '*zaam*', '*baar*', 'erig') are phonological words, this is not true for inflectional affixes. The outcome of inflection is a syntactically appropriate variant of the same word. Here, we study the production of inflected regular Dutch verbs. All but one inflection (the past participle involves a circumfix) is formed through suffixation.

Before turning to the details of the two production studies on derivation and inflection, we present information about the variables relevant to both studies. These concern *control* variables, *frequency* variables, *morphological* variables, and *phonological* variables. A first control variable relates to acoustic characteristics of a word's initial phoneme. As latencies are recorded via microphone, some prevoicing might not be detected by the voice key (Kessler, Treiman, & Mullennix, 2002), leading to longer latencies for plosive initial words. In all, we tried to reach a fair distribution of initial phonemes and features over the distribution of other variables. We included the variables INITIAL PHONEME (plosive, fricative, other) and VOICE (voiced or unvoiced initial phoneme) in the analyses. Another control variable was REPETITION, as we measure ten latencies per word per participant, and participants might become faster (due to practice) or slower (due to fatigue) within the sequence.

We were mainly interested in the frequency variables, the morphological variables, and the phonological variables. Frequency variables were 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 take a closer look at these variables, the motivation to study them and the expectations concerning their effects on the production latencies of morphologically complex words.

Frequency Variables

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. The production latency of a complex word should relate to its own frequency of occurrence (the SURFACE FREQUENCY) as observed for Chinese compounds in Janssen et al. (2008). Stemberger and MacWhinney (1986) suggest that only high-frequency regular variants are stored as complex words, while low-frequency regular inflections are composed from their constituents.

In contrast, decompositional models of speech production assume that all morphologically complex words are assembled from their constituting morphemes at the form level. 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. We 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 role of incrementality. Assuming both decompositionality and incrementality, Levelt et al. (1999), predict frequency effects for initial morphemes only.

An intermediate position between fully decompositional and non-decompositional models of speech production is structured storage, in which morphemes are stored as separate entities with information about their composability. In such a model, the frequency in which a morpheme occurs as a first constituent in complex words might be a better predictor of the production latency of one of these words than its frequency as an independent word, or as a constituent in *any* position. We computed the POSITIONAL FREQUENCY, defined as the sum of the frequencies of all members in the constituent family (Krott, Baayen, & Schreuder, 2001). For *grijp* in *grijpbaar*, it is the sum of the frequencies of all complex words that contain *grijp* as first constituent. The LEMMA FREQUENCY is defined as the summed frequencies of a word's inflectional variants. For deverbal adjectives, it is the frequency of the singular form plus the frequency of the plural form. For inflected verbs, it is the sum of the frequencies of all inflections, leaving aside the frequency of occasional singular nouns that are homophonous to the first person singular form of some verbs in Dutch, as in English (such as in *wens,* to wish). Note that we are at the level of word forms, adding up form frequencies of inflectional variants. While effects of frequency can be observed at higher levels in speech production, the robust effects have been attributed to the form-level (e.g., Jescheniak & Levelt, 1994). Furthermore, the task we use in the present studies taps in at the word-form level. Thus, our LEMMA FREQUENCY represents potentially connected word-forms. An effect of LEMMA FREQUENCY is not predicted by either full-form or decompositional models and would support the assumption of connected and co-activated word-forms. If complex words are stored with structural information, that is, with links between constituent morphemes that reflect the probabilities with which these morphemes tend to be combined, paradigmatic variables should outperform both morpheme and full-form frequencies, as observed previously for compounds.

Morphological Variables

The INFLECTIONAL ENTROPY is based on a word's inflectional paradigm. Rather than summing up the frequencies of the inflectional variants (as done to compute the lemma frequency), Shannon's entropy (Shannon, 1948; Shannon & Weaver, 1949) is a token-weighted count of types (Moscoso del Prado Martín, Kostic, & Baayen, 2004; Baayen, Feldman, & Schreuder, 2006). A high inflectional entropy indicates that a given stem is actually used in many or all of its inflected forms with similar frequencies. In that case, the production of a specific inflected form might be harder than when the forms in the inflectional paradigm are few and / or of different frequencies (Baayen, Levelt, Schreuder, & Ernestus, 2008). Chances of observing an effect are higher in a richer paradigm with more variation. In addition, the production of various inflected forms during the Experiment might lead to a stronger activation of the whole paradigm. Potentially, entropy effects can be observed only when the paradigm is relevant in performing the task. As a consequence, there are good reasons to expect different results with the derivations and inflections used in the present study due to differences in the richness and relevance of their inflectional paradigms. For deverbal adjectives, the inflectional variants are the singular and plural form and only singular forms are produced in Experiment 1. The paradigm of a verb contains seven inflectional variants, all of which are actually produced in Experiment 2. Independent of paradigmatic richness, effects of INFLECTIONAL ENTROPY are not predicted by either full-form or decompositional models of production and would indicate a co-activation of multiple word-forms with paradigmatic structure. The operationalization as frequency or entropy cannot differentiate between existing production models but may aid in the development of future models.

The DERIVATIONAL ENTROPY of *grijp* in *grijpbaar* reflects the frequency distribution of all morphologically complex words that share *grijp* as constituent, independent of its position within the word. In decompositional models, the shared constituent is accessed as a separate node, the more frequently, the faster. Thus, the frequencies with which complex words are used influence the speed with which a shared constituent can be accessed, but their relative distribution is irrelevant. Under the assumption of full-storage, the frequencies of other complex words should not matter at all.

Constituent families are the basis for calculating both the positional frequency presented above, and the POSITIONAL ENTROPY, defined as Shannon's entropy estimated by the relative frequencies of the constituent family members. 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 families with either a few members, or a large variation in frequency. The POSITIONAL ENTROPY of *grijp* in *grijpbaar*, thus, reflects the frequency distribution of all complex words that share *grijp* as first constituent. An effect of

POSITIONAL ENTROPY is not predicted by production models that assume either individual nodes for full-forms or nodes for constituents with no connections between them. The operationalization of constituent families in added frequencies or computed entropies cannot differentiate between existing models (as present models of production do not incorporate structured storage), but suggest what connections between form nodes may look like.

Phonological Variables

As speech unfolds over time, the time required for articulating as well as perceiving a word increases with its number of phonemes. If a word is fully planned before the onset of articulation, PHONOLOGICAL WORD LENGTH will not only affect the time required for articulation but also the speech onset latency (e.g., Meyer, Roelofs, & Levelt, 2003). In speech comprehension, a word may be recognized before its last phoneme has been perceived, depending on the location of the uniqueness point (e.g., Marslen-Wilson, 1990; Balling & Baayen, 2008). Similarly, length effects might not necessarily be related to the whole word, if the planning unit of articulation is shorter.

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 (i.e., a word's NEIGHBORHOOD DENSITY) have been encountered in both comprehension and production studies, 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). According to the Neighborhood Activation Model (Luce et al., 1998), effects of word frequency are directly tied to the number and relative frequency of phonologically similar words activated by a stimulus input, such that a high-frequency word with many or high-frequency neighbors might be harder to recognize than a low-frequency word with few or low-frequency neighbors. Scarborough (2004) showed that vowel-to-vowel co-articulation is more likely in words with sparse phonological neighborhoods. Vitevitch (2002) found words with dense neighborhoods to be produced more quickly than words with sparse neighborhoods. He attributed this finding to a co-activation of phonologically similar words that increases the activation of the target word. Neighborhood Density effects in speech production are crucial with respect to 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 eventually be selected and activate its word form with no interference by irrelevant, non-selected 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. They cannot inform us with respect to the question of full-storage versus full-decomposition, given that long or morphologically complex words tend to have few or no neighbors. We can only include neighborhood densities for the constituents.

The idea to take into consideration the phonological neighbors of a word originally comes from reading studies with whole words presented at once. As the speech signal unfolds over time, both in speaking and listening, words are processed from beginning to end. Therefore, we looked at the influence of phonological neighbors in an additional, more detailed and hopefully suitable way, separately counting the number of neighbors exchanging the first, second, third, etc. phonemes (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. We included the specific neighborhoods of the first three positions (N1-N3), because some morphemes consist of only three phonemes. If words enter into competition in an incremental process, a cohort effect is expected with maximum competition for the initial position. The greater the number of words compatible with the initial segment is from which a target has to be selected, the longer the naming latency.

A second group of variables based on the notion of incremental speech processing concern COHORT ENTROPIES (H1, H2, H3, etc.), defined as the entropy of all 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, 2006) found greater reduction of segments with small information loads in their production cohort, compared to segments with large information loads. This strongly suggests influences of other word forms on the production of the intended one. Cohort entropies and position-specific neighborhoods tend to be negatively correlated.

In stepwise covariance modeling, a number of variables can be used simultaneously to predict the response latencies. 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 in trying to explain as much variance as possible, using as few predictors as possible. Note that sometimes a variable is taken out even though it has some predictive value when other related predictors are removed from the model specification. In the presence of a correlated, stronger predictor, the first predictor simply does not provide sufficient information, and only the stronger one will appear in the final model. In such cases, discussing the remaining variables with an eye on the excluded ones can be very informative and enhance the general understanding.

Experiment 1 – Deverbal adjectives

In Dutch derivational morphology, new words (lexemes) are formed through affixation. The derived word (the output word) can be of the same or of a different word class than its base word (the input word). Booij (2002) presents examples for nine input-output word class relations in Dutch. While all examples in Figure 1 involve the attaching of a suffix, five of the nine input-output combinations can be also formed through the attaching of a prefix (e.g., *ver-slaaf*, to addict to). In what follows, we concentrate on deverbal adjectives (e.g., *lees-baar,* readable), the materials used here.

INSERT FIGURE 1 ABOUT HERE

Dutch deverbal adjectives are formed by attaching an adjectival suffix to a verbal stem, changing the word class from verb to adjective. The following nine suffixes form deverbal adjectives in Dutch: *-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 as two phonological versions of the same underlying suffix. As *-elijk* and *-lijk* differ in the number of syllables, we treat them as different suffixes.

Method

From the CELEX lexical database (Baayen, Piepenbrock, & Gulikers, 1995, CD-rom), 124 Dutch deverbal adjectives (e.g., *drinkbaar*, drinkable) were selected based on the following selection criteria. All suffixes should be represented, but each verbal stem should occur only once. The CELEX frequencies (based on a corpus of 42 million words) of both the deverbal adjectives and the verbal stems had to be greater than zero. For individual items, the CELEX frequency, a Google frequency and an average familiarity rating based on 27 participants (mean rating for the selected material was 4.3 on a 7-point scale, with a standard deviation of 2.0) had to be correlated. Last but not least, both the deverbal adjectives and their verbal stems had to cover a wide range of frequencies (Figure 2).

INSERT FIGURE 2 ABOUT HERE

Table 1 lists the absolute and relative CELEX-frequencies of the deverbal suffixes, next to their number and percentage of items used in the experiment. Equal-sized groups of items per suffix were unfeasible, given the selection criteria and the huge variation in frequency of occurrence of the deverbal suffixes. Compared to their CELEX token-frequencies, lower-frequency suffixes were overrepresented and higher-frequency suffixes were underrepresented in the material. All selected items are listed in the Appendices.

TABLE 1

The selected adjectives were assigned to 61 pairs, with the constraint that the adjectives within a pair had minimal phonological overlap and were semantically unrelated. Based on this first basic list of 61 pairs, we created three additional basic lists, balancing the order of adjectives within the sets. This was important as (see task description) the first adjective in a pair was presented *first* and had to be associated with an icon on the *left* side of the screen, while the second adjective was presented *second* and associated with an icon on the *right* side of the screen. By balancing the order of adjectives between lists (that is, between participants), potential position effects were cancelled out. In basic list 2, the order was reversed within all pairs, in lists 3 and 4 it was flipped in half of the pairs. For each basic list, we then constructed 6 randomizations of set order, creating a total of 24 lists. As the experiment had to be divided over two sessions, 31 experimental pairs were presented in the first session, 30 in the second, both of which were preceded by three practice pairs, which contained adjectives of similar structure and frequency.

Position-Response Association Task

We used a position-response association task (Cholin et al., 2004; Bien et al., 2005; Cholin, Levelt, & Schiller, 2006), a variant of the implicit priming paradigm (Meyer, 1990, 1991), which allows us to prompt the production of specific words without presenting them in written letters or pictures. Participants learn to associate auditorily presented words to one of two positions on the computer screen. The position mark is then used to prompt the production of the associated word. In previous studies by Cholin and colleagues and Bien and colleagues, this method has been shown to be sensitive to frequency effects. Its independence of picturability and orthographic effects makes it especially attractive for the production of derived and inflected words.

Participants were tested individually in a dimly lit sound-attenuated booth. Wearing headphones, they were comfortably seated in front of a CRT computer screen, a Sennheiser microphone and a button box. On average, a single session lasted 70 min. Participants who wanted to complete both sessions on one day had to take a break of minimally 90 min in between. The experimental procedure consisted of alternating learning and test phases. Both 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 next phase. In the learning phase, participants were presented with two spoken words over headphones.

Simultaneously with hearing the first word, they saw an icon of a loudspeaker appearing on the left side of the screen. Simultaneously with hearing the second word, 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) word. In the immediately following test phase, participants were repeatedly presented with the left or right icon as a prompt to name the associated word. Prompting was pseudo-randomized with maximally 4 consecutive repetitions of the same target word. Each word was prompted ten times. In the test phases, we included distractor trials to make it difficult for participants to prepare one of the target words and to avoid consecutive productions of the exact same word. 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. The voice key was activated simultaneously with the presentation of a prompt. Naming latencies longer than 1500 ms were counted as time-outs. The experimenter monitored the responses online, registering incorrect naming, hesitations and voice key errors.

Participants

From the subject pool of the Max Planck Institute for Psycholinguistics, 24 native speakers of Dutch (4 males) were recruited and received € 15 for completing both sessions. Most participants (17) completed both sessions within one day.

Results

Due to computer problems, the onset latencies of one participant were not recorded correctly and excluded from the analysis. Also, two items were excluded, because they are phonological neighbours (*zweterig* and *zweverig*). Of the remaining 28060 experimental trials (23 participants producing 122 items, each ten times), 959 (3%) time out trials (latencies >1500 ms), hesitations, wrong namings and voice key errors were removed prior to analyses. Naming latencies were analyzed in a mixed-effects regression analysis, with subject and item as crossed random effects (Pinheiro & Bates, 2000; Baayen et al., 2002; Bates, 2005; Bates & Maechler, 2009; Baayen, 2008; Baayen et al., 2008). 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.

The final model incorporated two random-effect factors: random intercepts for stem (STD = 0.034) and for subject (STD = 0.12), and the residual error (STD = 0.22). Table 2 summarizes the fixed-effect structure of the final model, including beta weights, standard errors, t- and p-values. There were four numerical predictors (REPETITION, POSITIONAL NEIGHBOURHOOD N1, COHORT ENTROPY H2, and CUMULATIVE STEM FREQUENCY) and the two factorial predictors (levels in parentheses, with the reference[2] level in italics) VOICED (*voiced*, unvoiced) and INITIAL PHONEME (*fricative*, other, plosive). Panels 1 to 6 of Figure 3 illustrate the partial effects of each predictor, adjusted for the effects of the other covariates at their medians.

TABLE 2

[2] In factorial variables, one level is modeled to lie on the intercept. The table lists the adjustment(s) for the other level(s)

Participants started relatively fast within the test phases and slowed down towards their ends. The inhibitory effect of REPETITION ($\beta = 0.0052$, $t(26500) = 10.46$, $p < 0.0001$) is shown in panel 1 of Figure 3.

INSERT FIGURE 3 ABOUT HERE

Words with unvoiced initial segments were named faster than words with voiced initial segments ($\beta = -0.0188$, $t(26500) = -2.16$, $p = 0.0309$, panel 2, for VOICE) and plosive-initial words elicited longer latencies than words beginning with non-plosives ($F(2, 26500) = 5.76$, $p = 0.0032$, for plosive, illustrated in panel 3, where label f denotes fricatives and label o denotes other initial phonemes).

There was an inhibitory effect of N1, the POSITION-SPECIFIC NEIGHBORHOOD DENSITY of the initial phoneme ($\beta = 0.0025$, $t(26500) = 2.25$, $p = 0.0247$, panel 4) and a facilitatory effect of the COHORT ENTROPY for the first two phonemes ($\beta = -0.0055$, $t(26500) = -2.31$, $p = 0.0207$, panel 5). The model further revealed a facilitatory, linear effect of the CUMULATIVE STEM FREQUENCY ($\beta = -0.0040$, $t(26500) = -2.09$, $p = 0.0367$). 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).

Discussion

The statistical model that best predicts the naming latencies of 124 Dutch deverbal adjectives 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, INITIAL PHONEME, and VOICE).

Within the sequences of test trials, participants gradually slowed down towards their end. It might be argued that, if anything, repeated naming of the same word should lead to shorter, rather than longer latencies. However, the response-association task seems to be rather boring and causes a decrease in motivation or alertness from the first naming to the last repetition. The disadvantage for plosive-initial words is most likely an artifact of the voice key.

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 (N1), whereas the non-significant densities N2 and N3 showed trends of facilitation. In the presence of the position-specific neighborhood counts, the overall NEIGHBORHOOD DENSITY dropped out of the model. As mentioned previously, variables falling out during stepwise analysis of covariance are not necessarily unpredictive. 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.

The results underline the importance of NEIGHBORHOOD DENSITIES for speech production latencies. The inclusion of neighborhoods computed for specific positions within the word was fruitful. N1 was a stronger predictor than the overall density count. A high neighborhood density of the initial phoneme slows the naming latency of the adjective. The more rhyme neighbors (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 overlap is generally facilitative, there is an inhibitory component in overlapping initial phonemes. In their sequential cuing model, Sevald et al. propose that 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 results of the present study suggest that when N1-neighbors are produced in sequence (as in pick, tick), they also co-activate each other as N1-neighbors, turning the disadvantage of a big N1-neighborhood into an advantage. Without a preceding production of rhyme words, bigger N1-neighborhoods mean more co-activation, making it harder to select the to-be-produced initial phoneme.

While rhyme neighbours make the production of a deverbal adjective more difficult, words starting with the same two phonemes facilitate production. Naming is the fastest (i.e. the COHORT ENTROPY H2 is the highest) when the number of words in the cohort sharing the initial two phonemes is high and these words have little variation in frequency. In speech corpora studies analyzing the relative length of segments within words, the information load in a production cohort was found to be negatively correlated with the amount of reduction of segments. The present study suggests that the frequency distribution in a production cohort does not only affect the length with which word segments are produced, but also the time it takes to plan the word.

There was no effect of INFLECTIONAL ENTROPY for the deverbal adjectives, suggesting that either there was too little variation in the paradigms (just the singular and plural form), the paradigms were not sufficiently activated (only singulars were produced), or they simply do not matter in production.

Of the frequency variables, only the CUMULATIVE STEM FREQUENCY remains in the final model. The absence of a suffix-frequency effect may have several reasons. First, it 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. Second, the number of different suffixes used (nine) might be too small to show an effect. Third, the repeated usage of the deverbal suffixes might have masked actual frequency differences.

Because of the importance of the debate on the presence or absence of full-form frequency effects (e.g., Caramazza et al., 1998; Jescheniak, Meyer, & Levelt, 2003; Bien et al., 2005; Janssen et al., 2008) we want to stress that the whole-word frequency was not disregarded simply because it could not explain a significant proportion in the presence of better predictors. It was actually non-significant 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 Initial Phoneme), none of the following variables yielded a significant fixed effect: Surface Frequency (β = -0.0019, $t(26504)$ = -0.80, p = 0.4237); Lemma Frequency (β = -0.0024, $t(26504)$ = -0.96, p = 0.3371); Neighborhood Density (β = 0.0008, $t(26504)$ = 1.73, p = 0.0836). The only competitor for the Cumulative Stem Frequency was the Positional Frequency of the stem (β = -0.0062, $t(26504)$ = -2.69, p = 0.0072). 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 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. Indication of paradigmatic relations between morphemes is weak as the positional frequency was predictive but outperformed by the cumulative frequency of the stem.

Experiment 2 - Inflected verbs

Predictors

As in Experiment 1, we included four kinds of predictors: control variables, frequency variables, morphological variables, and phonological variables. The control variables were REPETITION (naming one to ten), INITIAL PHONEME (plosive, fricative, other) and VOICED (voiced versus unvoiced initial phoneme). In the frequency domain, we looked at the LEMMA FREQUENCY of the verb (the summed frequency of all inflectional variant), 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).

We further included the INFLECTIONAL ENTROPY, with a high entropy indicating that a particular verb stem is used in many or all of its inflectional variants, and that these inflections occur with similar frequency. Under these circumstances production of a specific inflected form might be harder. INFLECTIONAL ENTROPY was not significant in Experiment 1, but there are good reasons to expect a different result in Experiment 2 as the verbal inflectional paradigms are not only substantially richer but also activated throughout the experiment. As only one of the seven types of Dutch verbal inflections, the past participle, contains a circumfix (*ge-t/d*), we added the two-level factor CIRCUMFIX (circumfixed versus uncircumfixed) next to the PHONOLOGICAL WORD LENGTH, which has been found to affect naming latencies in Meyer et al. (2003), but was not a significant predictor in Experiment 1.

As with the deverbal adjectives, we computed the number of phonological neighbors of the inflected verbs (NEIGHBORHOOD DENSITY), the POSITION-SPECIFIC NEIGHBORHOODS N1, N2, and N3, and the COHORT ENTROPIES H1, H2, and H3. Recall that we observed significant effects for a POSITION-SPECIFIC NEIGHBORHOOD (N1) and a COHORT ENTROPY (H2) in Experiment 1.

Method

We selected 126 Dutch verb stems from the CELEX lexical database (Baayen et al., 1995), reusing as many verb stems as possible that had been used in Experiment 1 to increase comparability. We could not reuse all of the verbal stems because we restricted the inflectional study to regular verbs. Irregular inflections differ from regular inflections in several aspects. Some irregular inflections undergo vowel changes. Furthermore, inflectional variants of irregular verbs and regular verbs often differ in the number of syllables. We decided not to mix regular with irregular verbs in the present study, but to focus on the role of frequency information in the inflection of regular verbs. Replacing all irregular verb stems in the item pool of Experiment 1 with regular verbs, the final overlap was 76%. All selected verbs had CELEX stem frequencies greater than zero. The 126 selected verbal stems were evenly distributed over the seven types of Dutch verbal inflections (Table 3). The assignment to a particular type of inflection was pseudo-randomized with two restrictions: First, the frequency of the inflected form had to be greater than zero and in line with both their Google frequencies and additionally collected familiarity ratings from 46 participants. For the selected item pool, mean familiarity was 4.4, with a standard deviation of 1.8 on a 7-point scale. Second, initial phonemes had to be distributed over the seven inflections, especially with respect to the features plosive and voice. A complete list of items, as paired in the position-response association task, is provided in Appendix B. Figure 4 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.

TABLE 3

FIGURE 4

We constructed 24 experimental lists, one list for each participant. First, the 126 inflected verbs were assigned to 63 sets of two (items 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 be semantically unrelated. As in Experiment 1, four basic lists, each with 6 randomizations, were created, split and complemented by three practice sets.

Position-Response Association Task

The same position-response association task was used as in Experiment 1. Each participant took part in two sessions lasting 70 min on average with a break of minimally 90 min in between. 16 participants took part on different days.

Participants

24 native speakers of Dutch (3 male) were recruited from the subject pool of the Max-Planck Institute for Psycholinguistics, Nijmegen, and received € 15 for completing both sessions.

Results

One participant did not complete the experiment and none of his latencies were analysed. We further excluded all latencies of the experimental item *gruwen*, as many participants seemed to have problems recognizing it over headphones, causing an unusual high number of hesitations, time outs trials and false naming responses. All latencies of the experimental item *voedde* were excluded, because it accidentally appeared in two sets, once instead of the intended item *gepronkt*. Of the remaining 28290 experimental trials, only those that were named correctly within a latency of 1500 ms were included in the analyses. A total of 1307 (4.6%) time-out trials (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. As a rule of thumb, the number of item-based predictors should not exceed the number of items divided by 15, to avoid overfitting. Thus, given the number of items, no more than eight item-based parameters should be included in the model. 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 summarizes the fixed-effects statistics, including beta weights, standard errors, t-values and p-values. Figure 5 pictures the partial effects of the significant predictors, each adjusted for the effects of the other covariables.

TABLE 4

As in Experiment 1, the control variables REPETITION and INITIAL PHONEME affected the naming latencies. Naming was fast at the beginning of the test phases and slowed down towards their end. This inhibitory effect of REPETITION ($\beta = 0.0055$, $t(26699) = 8.24$, $p<0.0001$) interacted with INITIAL PHONEME in such a way that it showed strongest for fricative-initial (f) words and, to a lesser degree, the plosive-initial (p) words, while there was hardly any slowing down for other (o) initials (panel 1 of Figure 5).

Naming latencies of circumfixed inflections (panel 2) were slower than those of uncircumfixed inflections (β = 0.0941, t(26699) = -3.80, p=0.0002). The only Dutch verbal inflection carrying a circumfix is the past participle form (*ge-…-t*). Finally, as in Experiment 1, we observed a facilitative effect of the COHORT ENTROPY H2 (β = -0.0080, t(26699) = -3.32, p<0.0001). 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).

INSERT FIGURE 5 ABOUT HERE

Discussion

Triggered by previously associated visually presented symbols, 24 native speakers of Dutch repeatedly produced 126 Dutch inflected verb forms. The variance in the naming latencies is best modeled using the following information: how often the item has been produced previously in the experiment (REPETITION), whether or not the item contains a circumfix (CIRCUMFIX), and the size of the entropy in the production cohort H2. There were no frequency effects of the inflected verb or its stem.

As in Experiment 1 (and in Bien et al., 2005), the inhibitory effect of REPETITION suggests that the test-phases failed to maintain the initial alertness over all 40 trials (20 namings of words plus 20 namings of numbers). Alternatively, the two experimental items within a test phase might become stronger competitors with increasing repetition. After all, the task is predictable with respect to the *moment* at which the trigger of an experimental item is presented (unpredictable is its *position*) and participants might prepare both items while awaiting the trigger. Slowing down was most apparent for fricatives and plosive-initials. Potentially, repetitive articulations of fricatives and plosives were relatively more tiring.

Information on the PHONOLOGICAL WORD LENGTH or NEIGHBORHOOD DENSITIES (overall or position-specific) did not (help to) explain a significant proportion of variance in the naming latencies of the inflected verbs. Also not among the best predictors were the SURFACE FREQUENCY of the inflected verb, the LEMMA FREQUENCY (the summed frequency of all its inflectional variants), and the INFLECTIONAL ENTROPY estimated by the relative frequencies of a these inflectional variants. Neither were there STEM FREQUENCY effects (cumulative or positional). Note that the LEMMA FREQUENCY is part of the CUMULATIVE STEM FREQUENCY. In inflected verbs, the CUMULATIVE STEM FREQUENCY is little more than the LEMMA FREQUENCY, and these frequency measures are highly correlated (r = 0.85 in the item pool of Experiment 2, Figure 6).

INSERT FIGURE 6 ABOUT HERE

In multivariate regression analyses, highly correlated predictors cause high colinearity. To test whether this correlation is responsible for the absence of any frequency effects, we computed a new variable, OTHER FREQUENCY, representing the unique part of the CUMULATIVE STEM FREQUENCY that does not overlap with the LEMMA FREQUENCY. 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 collinearity between the predictors, we 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 was true for the POSITIONAL STEM FREQUENCY, the LEMMA FREQUENCY and SURFACE FREQUENCY. The only variables that were significant in the absence of other non-control predictors, were the NEIGHBORHOOD DENSITY ($\beta = 0.0012$, $t(26701) = 2.36$, $p=0.0183$), as well as the POSITIONAL NEIGHBORHOOD N1 ($\beta = 0.0033$, $t(26701) = 3.29$, $p=0.0010$).

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 latencies Experiment 1. 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 that carries a circumfix. Thus, it is mostly the distribution of frequencies within the cohorts that determines their entropies. A cohort with little frequency variation speeds up the production latency of any member of the cohort. As discussed before, the probability distribution in a production cohort, thus, not only affects the acoustic length with which segments are produced within the word, but also the naming latency of the word.

Back to the question how regularly inflected verbs are represented in the mental lexicon. The surface frequency of the inflected verb does not affect the speed with which the verb is named, challenging models that assume a full listing of complex words. But neither does the frequency of the stem correlate with the naming latencies, challenging decompositional models. Even as the only predictor next to the control variables, no significant proportion of variance in the naming latencies was explained by either the frequency of the stem or by the frequency with which the surface form occurs. The latencies with which the regularly inflected verbs were named in Experiment 2, solely reflect the frequency distribution within the cohort sharing the initial two phonemes.

Motivated by the observation of elongated latencies for circumfixed forms, we performed an additional stepwise multivariate analysis of covariance, excluding all (36) items that were either circumfixed themselves or were tested with a circumfixed partner. Circumfixed forms differ from the other inflections in various ways. First, their affix (*ge- –t/d*) is placed around the stem, rather than attached to its end. Second, given that all past-participles forms share the initial diphone, their cohorts might be exceptionally large. The size of a cohort influences its entropy, next to the distribution of the frequencies. In our sample, the mean logged COHORT ENTROPIES H1, H2, H3 were 9.1, 9.1, 5.4 for the circumfixed inflections versus 7.5, 5.4, 3.8 for the others. Both for the deverbal adjectives and the inflected verbs, however, H2 was negatively correlated with naming latencies. Third, past participles begin with an unstressed, rather than a stressed syllable. Their longer naming latencies are in line with findings by Schiller, Fikkert, and Levelt (2004). Using picture naming, they observed longer latencies for stress-initial targets as opposed to stress-final targets. Finally, only one subtype of inflected verbs carried a circumfix, possibly creating a kind of oddball out status for these items in the experiment. Figure 7 presents the partial effects for those (90) inflectional items, which had been presented in circumfix-free pairs (separate lists of items are provided in the Appendix B).

INSERT FIGURE 7 ABOUT HERE

Also in this subset of items, the facilitatory effect of the COHORT ENTROPY H2 (Panel 3 of Figure 7) is significant, indicating that it cannot be attributed to the shared circumfix of the past participles. The inhibitory effect of REPETITION was independent of INITIAL PHONEME (Panel 1). Intriguingly, the exclusion of the circumfix-sets revealed three additional predictors. There was an inhibitory effect of the POSITION-SPECIFIC NEIGHBORHOOD N1 (Panel 4). The same effect was present Experiment 1. Furthermore, there was an inhibitory effect of INFLECTIONAL ENTROPY (Panel 2), and a non-linear effect of LEMMA FREQUENCY (Panel 5). With shortest latencies for medium frequencies, this non-linear effect is very similar to the compound-frequency effect described in Bien et al. (2005). Note that the LEMMA FREQUENCY is computed over verbs only. 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 the frequency of the homophonous noun is added up along with the frequencies of the verb's inflectional variants, predictivity vanishes. We will come back to this in the general discussion.

To study whether the presence of a prefix generally elongates production latencies, and differs from the production of other morphologically complex words, future research could make use of derivational prefixes (deverbal adjectives are restricted to suffixation, but five of the nine types of Dutch derivations can be formed through either prefixation or suffixation). Given that the final models explain about one third of the total variance in the naming latencies, there is certainly room for so-far neglected factors to play a role.

General Discussion

Taken together, there is mixed evidence with respect to the decompositionality and paradigmatic relations of morphologically complex words in speech production. In Experiment 1, the latencies of deverbal adjectives were predicted by the frequency of the verbal stem, arguing for decomposition. Other than surface frequency (which was not a significant predictor even when entered as the only variable next to control variables), the positional frequency of the stem was predictive but outperformed by the cumulative frequency. There is no indication of a paradigmatic structure, in which morphemes are tightly linked, but strong evidence for the co-activation of phonologically related word forms. The latencies with which the regularly inflected verbs were named in Experiment 2 were neither related to the frequency of the stem nor to the surface frequency of the complex word, but reflect paradigmatic relations between inflectional variants and cohort words. In the following, we discuss three aspects in more detail. First, why is there a paradigmatic effect of inflectional entropy that is predictive for inflected verbs, but not for deverbal adjectives, why is it inhibitory, and what are the implications of such an effect in production? Second, what do inhibitory effects of the POSITION-SPECIFIC NEIGHBORHOOD N1 imply? And finally, how can we interpret a non-linear effect of lemma frequency of a morphologically complex word?

Why does the INFLECTIONAL ENTROPY affect the latency with which one of the inflectional variants is named for verbs but not for deverbal adjectives? As discussed previously, the paradigm of a Dutch regular verb contains seven inflectional variants, all of which were actually produced in Experiment 2. For an adjective, inflectional entropy is computed over just two forms, singular and plural, and only singulars were produced in Experiment 1. In a richer paradigm, there is a higher chance of observing an effect. On top of that, the actual production of inflected forms throughout the experiment might lead to a stronger activation of the paradigm. Task relevance might even be a prerequisite for observing an effect of the inflectional paradigm. It is thus not surprising for INFLECTIONAL ENTROPY to turn out significant in Experiment 2, and non-significant in Experiment 1. With the inflections and derivations used, it is most likely a consequence of differences in both the activation and richness of the inflectional paradigms.

Why is the effect of INFLECTIONAL ENTROPY inhibitory? Note that there is just one entropy-value for all variants in a given inflectional paradigm, affecting the latencies with which higher- and lower-frequency variants are named in the same way. A high INFLECTIONAL ENTROPY is indicative of an inflectional paradigm with a higher information load. Selecting an inflected variant from such a paradigm is more difficult, because the paradigm has more members, because the members have more similar probabilities, or both (e.g., Baayen, Levelt, Schreuder, & Ernestus, 2008). For Dutch regular verbs, entropy is highest when all seven inflectional variants are produced equally frequently. Under these circumstances, the production of a specific inflectional variant seems the hardest. 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 et al. (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 that we used, taps in at the level of word form encoding. Next to other paradigmatic effects, the effect of INFLECTIONAL ENTROPY, thus, 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 (Roelofs, 1997) 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.

Both in the naming latencies of the deverbal adjectives and inflected verbs (analyses without circumfix-sets), we saw an inhibitory effect of POSITION-SPECIFIC

NEIGHBORHOOD N1. It is difficult to produce a word that has many rhyme neighbors. Their 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 latencies. Crucially, the effect of N1 indicates competition between word forms in production.

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 Bien et al. (2005), with shortest latencies for medium frequencies. Crucially, it is an effect of the frequency of the lemma, which is the summed frequency of the inflectional variants of a word. 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 task used in the present study is form based, and the effects for the sum of frequencies of the inflectional variants are stable over repetition. One possible explanation for the anti-frequency effect observed for higher-frequency lemmas is that the inflected variants for such lemmas develop their own full-form representations (e.g., Stemberger & MacWhinney, 1986; Bybee & Scheibman, 1999), which would then compete with the stem representations during lexical access. Alternatively, the U-shaped effect of lemma frequency might be due to response optimization, with shorter latencies for the more likely lemma frequencies and longer latencies as lemma frequencies become more extreme and less probable in the experiment as argued by Tabak, Schreuder, and Baayen (2010) for a similar U-shaped frequency effect observed in a picture naming study, in which participants produced past-tense forms of Dutch regular and irregular verbs. There, the authors hypothesize that the effect arises as a consequence of response optimization for the most *likely* lemma frequencies in the experiment. According to this explanation, production latencies are proportional not to lexical probability (as gauged by lemma frequency), but proportional to the probability of these probabilities The more exceptional a lemma frequency is, the longer the response latencies are, optimizing not the highest frequencies but the most common ones.

Taken together, the naming latencies of the deverbal adjectives and the inflected verbs argue against full listing (e.g., Butterworth (1983); Janssen et al., 2008), 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 (e.g., Levelt, et al., 1999; Levelt, 2001), but there is no effect of the frequency of the verbal stem on the naming latency of regular inflected verbs. The effect of INFLECTIONAL ENTROPY challenges pure decomposition. The non-linear effect of the LEMMA FREQUENCY of inflected verbs might suggest a less incremental production of high-frequency variants (e.g., Stemberger & MacWhinney, 1986; Bybee & Scheibman, 1999), or following Tabak et al. (2010), a response optimization for the most common LEMMA FREQUENCY. Neighborhood (e.g., Sevald & Dell, 1994) and COHORT ENTROPY effects (Van Son & Pols, 2003; Van Son & Van Santen, 2005; Kuperman et al., 2006) reflect paradigmatic relations in the mental lexicon and suggest that all word forms influence the production of one word. In all, our data suggest that the word-form level does not contain full listings or strictly separated morphemes but morphemes with links to phonologically and – in case of inflected verbs - morphologically related word forms.

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. (2008) *Analyzing Linguistic Data. A Practical Introduction to Statistics*. Cambridge University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. Journal of Memory and Language, 59(4), 390-412.

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., Levelt, W. J. M., Schreuder, R., & Ernestus, M. (2008). Paradigmatic structure in speech production. *Chicago Linguistics Society*, *43(1)*, 1-29.

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.

Balling, L., & Baayen, R.H. (2008). Morphological effects in auditory word recognition: Evidence from Danish. *Language and Cognitive Processes*, *23*, 1159-1190.

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

Bates, D.M., & Maechler, M. (2009). lme4: Linear mixed-effects models using S4 classes, R package version 0.999375-32.

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.

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

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.

Bybee, J., & Scheibman, J. (1999). The effect of usage on degrees of constituency: reduction of don't in English. *Linguistics*, *37*, 575-596.

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.

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.

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.

Janssen, N., Bi., Y., & Caramazza, A. (2008). A tale of two frequencies: Determining the speed of lexical access in Mandarin Chinese and English compounds. *Language and Cognitive Processes*, 23(7), 1191-1223.

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: Comparatives and agent nouns in Dutch and English. *Language and Cognitive Processes*, *20*, 43-73.

Koester, 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.

Koester, D., & Schiller, N. O. (2008). Morphological priming in overt language production: Electrophysiological evidence from Dutch. *NeuroImage*, *42*, 1622-1630.

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., and Baayen, R. H. (2006). Morphological predictability and acoustic salience of interfixes in Dutch compounds. Journal of the Acoustical Society of America 51, 2018-2024.

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.

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. (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.

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.

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.

Pinker, S. (1999). *Words and Rules: The Ingredients of Language.* New York: Basic Books.

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

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*.

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

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

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

Shannon, C.E. (1948). A Mathematical Theory of Communication, *The Bell System Technical Journal*, *27*, 379-423.

Shannon, C. E., & Weaver, W. (1949). *The Mathematical Theory of Communication*. University of Illinois Press, Urbana.

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

Tabak, W., Schreuder, R., & Baayen, R. H. (2010). Producing inflected verbs: A picture naming study. *The Mental Lexicon*. *5(1)*, 22-46.

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.

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

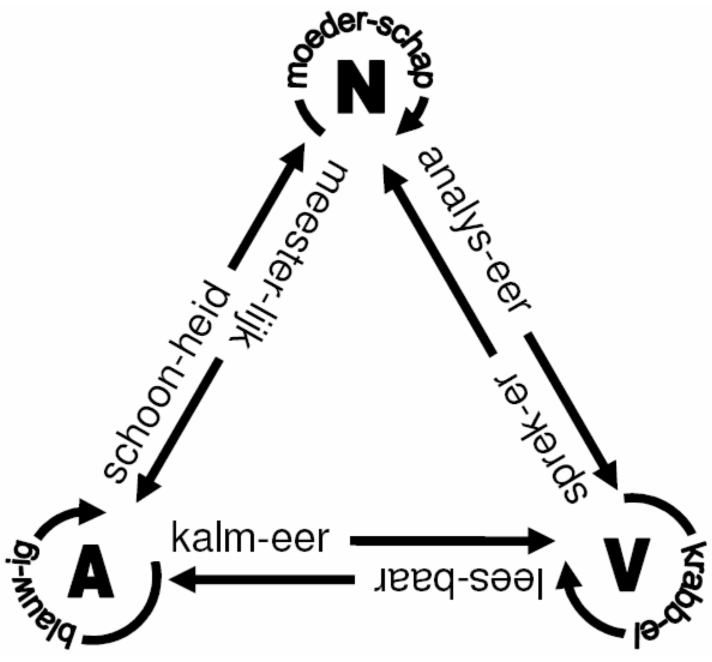

Figure 1: Dutch derivational morphology knows nine input-output word class combinations (N for noun, A for adjective, V for verb). The examples are taken from Booij (2002).

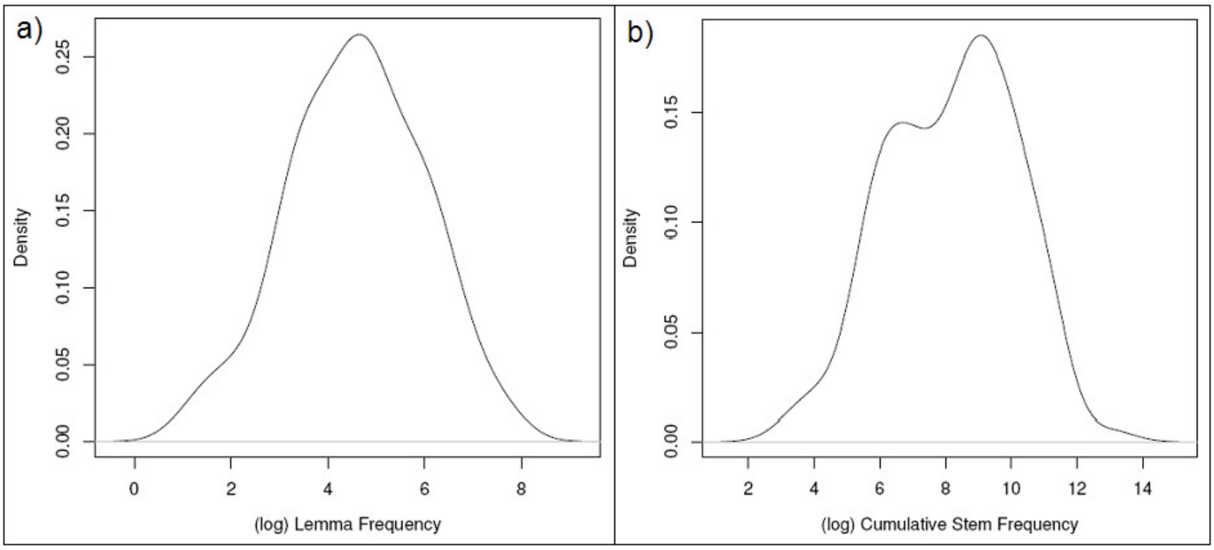

Figure 2: The selected deverbal adjectives were distributed over a wide range of both the lemma frequencies of the deverbal adjectives (a) and cumulative stem frequencies (b).

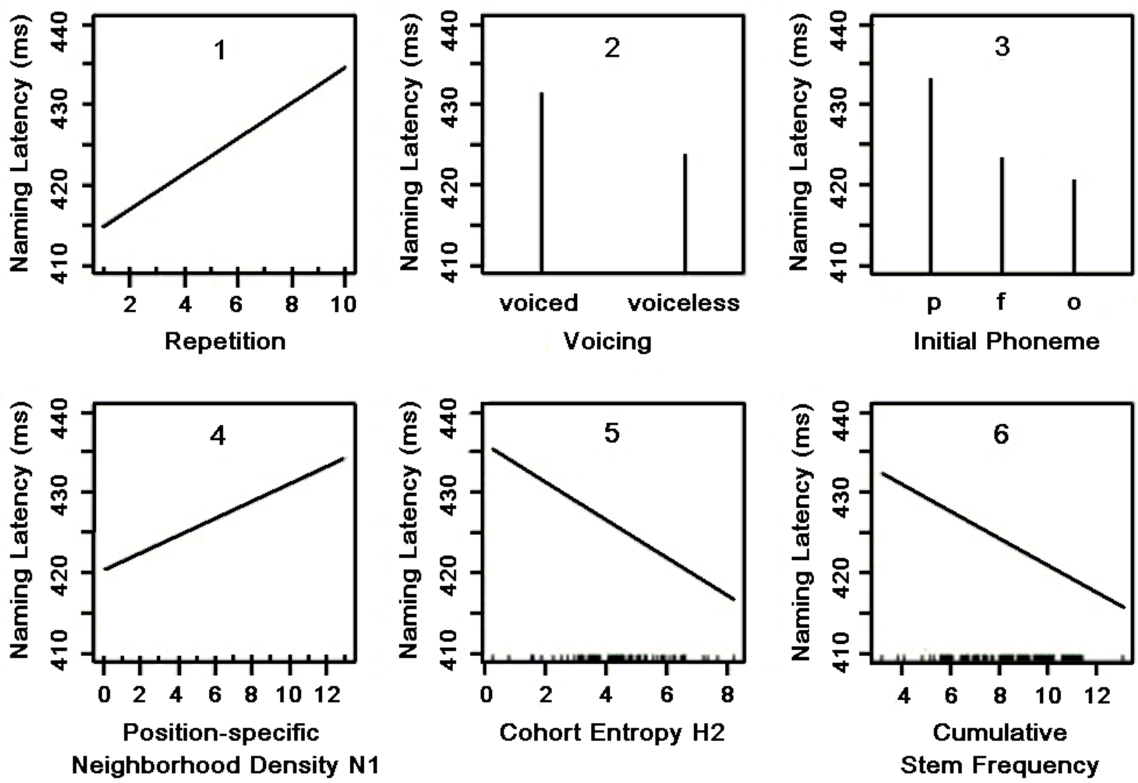

Figure 3: Deverbal adjectives: Partial effects of the predictors adjusted for the effects of the other covariables (panel 3: p=plosives, f=fricatives, o=other initial phonemes).

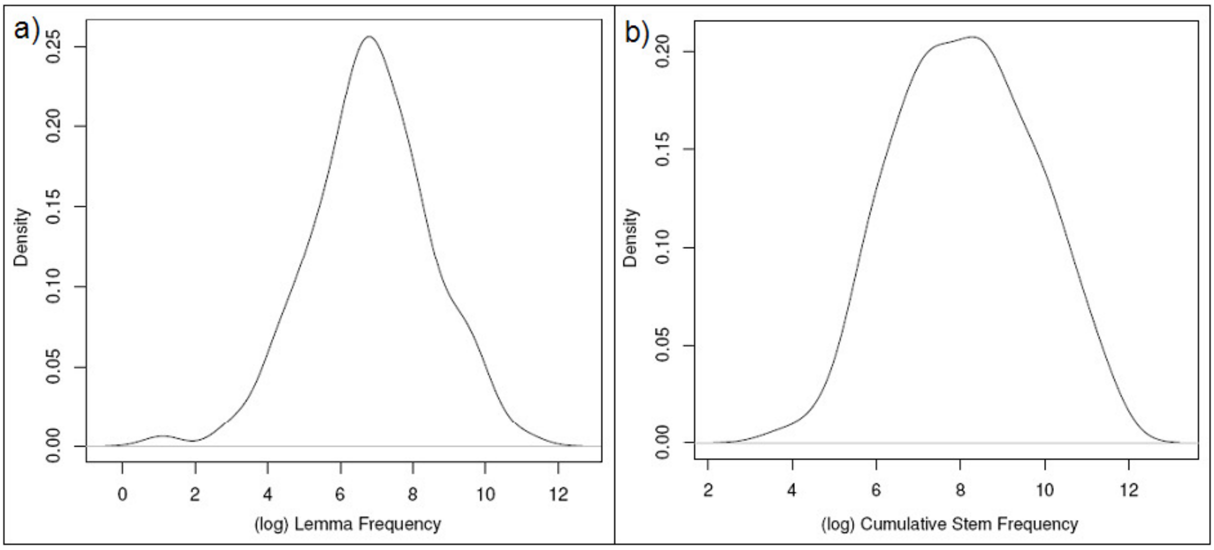

Figure 4: The distribution of LEMMA FREQUENCIES (a) and CUMULATIVE STEM FREQUENCIES (b) in the material.

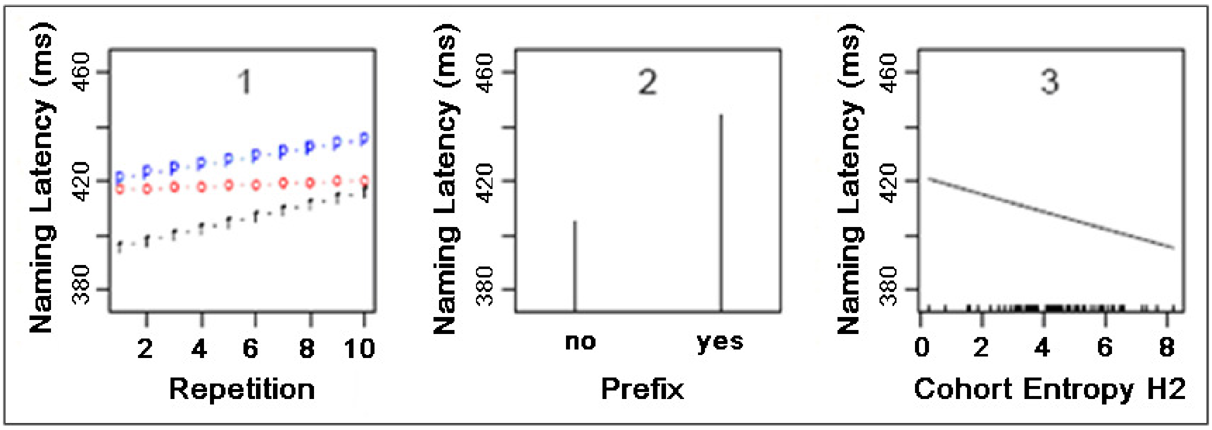

Figure 5: The partial effects of each predictor, adjusted for the effects of the other ones. Panel 1 shows the interaction of REPETITION and INITIAL PHONEME with plosive (p), fricative (f), other (o)).

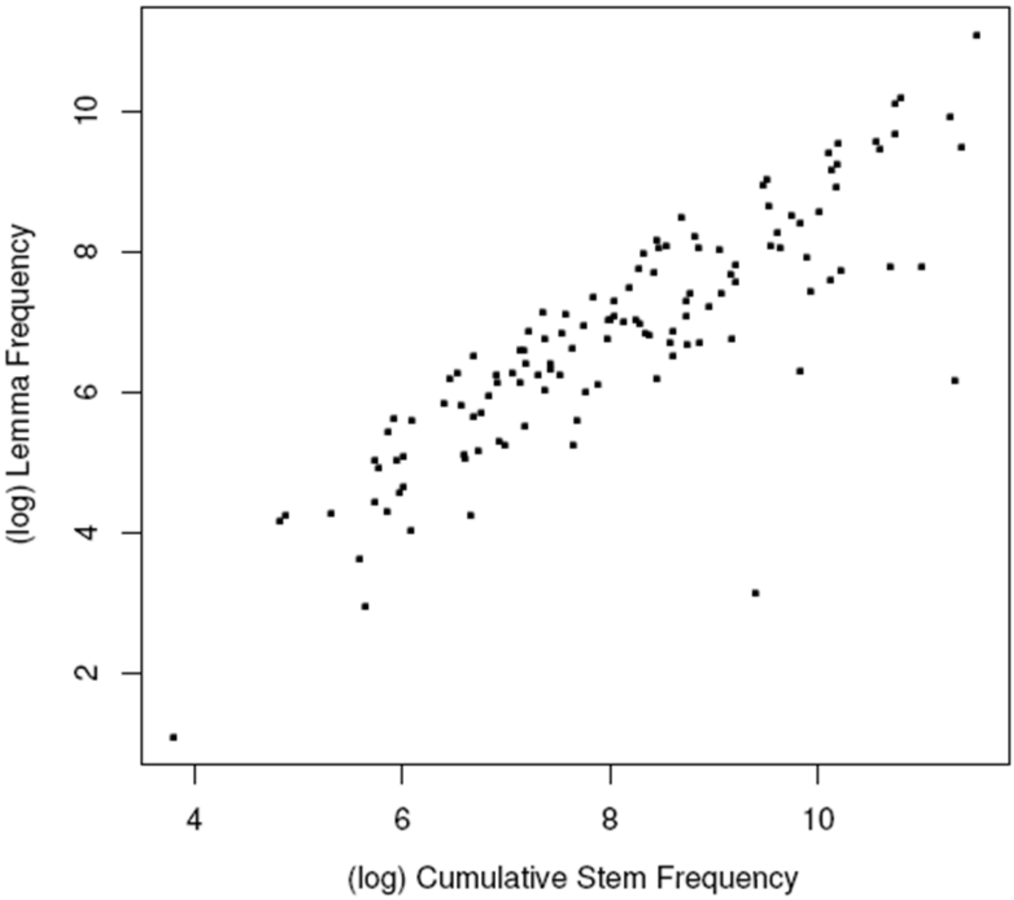

Figure 6: In Dutch inflected verbs, the LEMMA FREQUENCY and the CUMULATIVE STEM FREQUENCY are highly correlated (r = 0.85 in the present sample).

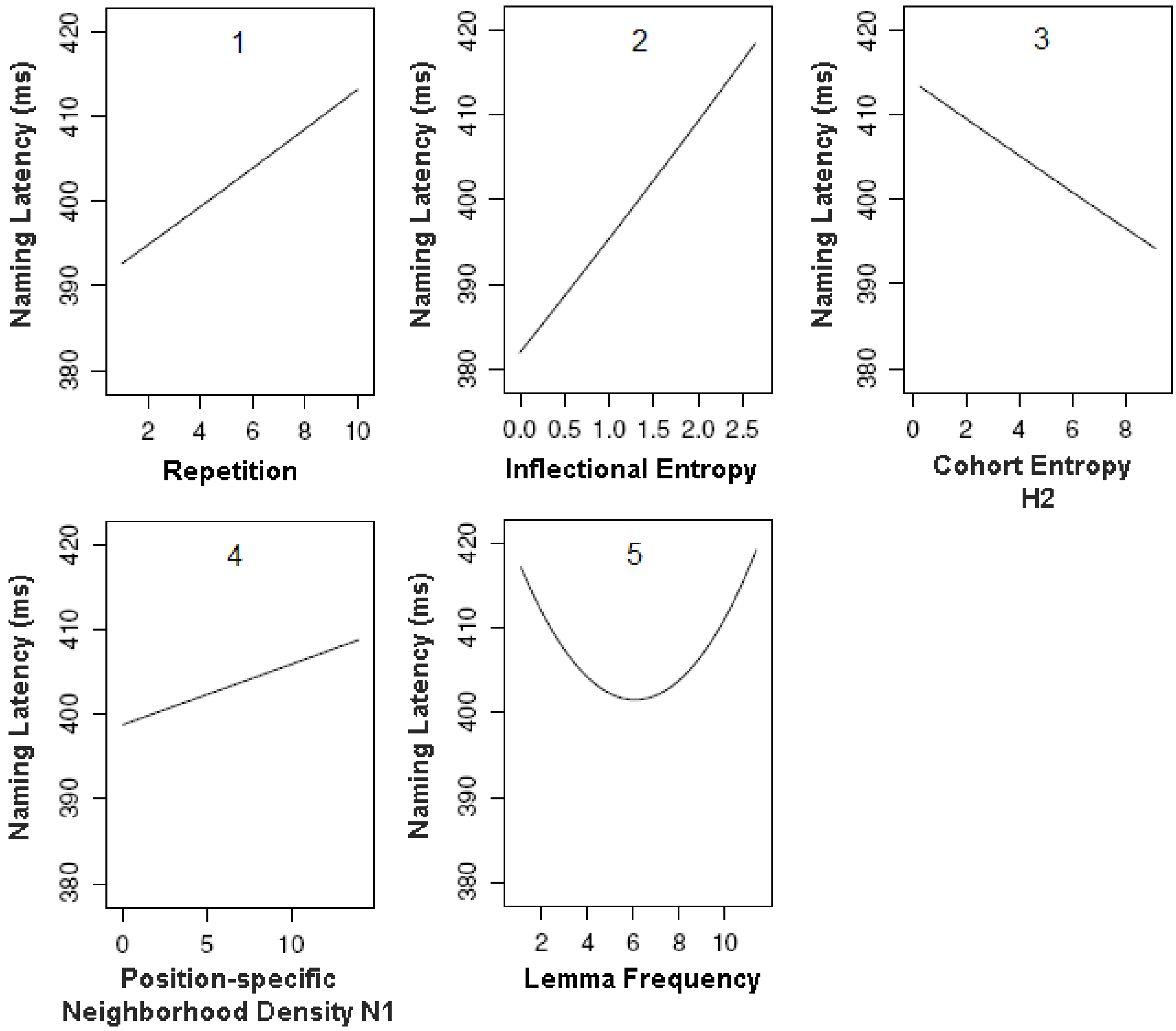

Figure 7: The partial effects for the subset of those 90 inflected verbs, which had been produced in prefix-free pairs in the position-learning association task.

Table 1: The nine suffixes forming deverbal adjectives in Dutch, their occurrence in the CELEX lexical database and in Experiment 1

deverbal suffix		**CELEX**				**Experiment**	
		Frequency		**%**		**Number**	**%**
		Token	**Type**	**Token**	**Type**	**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%**

Table 2: Fixed effects of the multilevel regression model, with subject and stem as crossed random effects of the deverbal adjective naming latencies.

		Estimate	Std.Error	DF	t.value	p.values
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
INITIAL PHONEME	**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

Table 3: 126 regular verbal stems were pseudo-randomly assigned to the seven types of Dutch verbal inflections (1-7), 18 items per type of inflection. (items marked with * were excluded from the analyses, see results-2)

(1) present 1st sg	**(2) present 2nd, 3rd sg**	**(3) present pl, infinitive**	**(4) present participle**	**(5) past sg**	**(6) past pl**	**(7) past participle**
affixation: -Ø	**affixation: -t / -d**	**affixation: -en**	**affixation: -end**	**affixation: -de / -te**	**affixation: -den / -ten**	**affixation: ge- … -d / -t**
bouw	bakt	broeden	brandend	beefde	beuzelden	geacht
broei	droomt	brommen	draaiend	deerde	daalden	gedeeld
duld	gunt	gruwen *	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 *
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 *	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

Table 4: The 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 are modelled to 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

Appendix A: Deverbal adjectives

The 124 deverbal adjectives assigned to 62 item pairs 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)).

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

Appendix B: 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 pairs for presentation in the position-response association task.

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

a) The regular verbal stems used in both Experiment 1 and 2

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

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

bakt, beefde, beuzelden, bouw, broeden, broei, brommen, daalden, deerde, draaiend, duld, durend, dweepten, gruwen[2], 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, 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[2], voegen, voelt, vulden, walgen, weerde, weifelt, weiger, wenden, wensten, werkend, woelde, zaagt, zaai, zeur, zoent, zorgden, zwetend, zwierde

c) Inflected verbs tested in prefix-sets (36)

gehoord, brandend, droomt, geacht, gedeeld, gehaald, gehuwd, gekwetst, gemist, gepronkt[2], 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