Producing inflected verbs: A picture naming study

Wieke Tabak

Radboud University Nijmegen

wieke.tabak@mpi.nl

Robert Schreuder

Radboud University Nijmegen

r.schreuder@donders.ru.nl

Harald Baayen

University of Alberta, Edmonton

baayen@ualberta.ca

January 13, 2010

Corresponding author:

R. H. Baayen

Department of Linguistics

4-55 Assiniboia Hall

University of Alberta

Edmonton T6G 2E5, AB

CANADA

baayen@ualberta.ca

Abstract

Four picture naming experiments addressing the production of regular and irregular past-tense forms in Dutch are reported. Effects of inflectional entropy as well as effects of the frequency of the past-tense inflected form across regulars and irregulars support models with a redundant lexicon while challenging the dual mechanism model (Pinker, 1997). The evidence supports the hypothesis of Stemberger (2004) and the general approach of Word and Paradigm morphology (Blevins, 2003) according to which inflected forms are not derived from the present-tense stem, but accessed independently.

Introduction

Ever since Bloch (1947) and Chomsky and Halle (1968), inflected forms have been understood as being created on-line from their base words by inflectional rules in language comprehension and speech production. This conceptualization reflects how pedagogical grammars proceed from simpler to more complex forms. However, Stemberger and colleagues (Stemberger, 2002; Stemberger and Middleton, 2003; Stemberger, 2004) suggested that inflected forms are not derived in this generative sense. Their evidence, based on speech errors, suggests that in English present and past tense forms are both activated and enter into a competition process that is modulated both by the long-term probabilities of the stem vowels as well as by short-term influences exerted by vowels in preceding words.

If inflection is not driven by 'derivational' rules, this has far-reaching consequences for the dual mechanism model proposed by, e.g., Pinker and Prince (1988); Pinker (1991); Pinker and Prince (1994); Pinker (1999); Pinker and Ullman (2002) for inflection. According to the version of this model formulated in Ullman et al. (2002); Ullman (2004), irregular verbs are stored in declarative memory, while regular verbs are processed by a procedural memory encoding rules of inflection that parse inflected forms into their constituents in comprehension, or that put together inflected forms from their constituents in production. As pointed out by Stemberger (2004), a model in which present-tense and past-tense forms are in competition is also problematic for many connectionist models adopting the assumption that marked forms are derived from less marked forms (e.g. Rumelhart and McClelland, 1986; MacWhinney and Leinbach, 1991; Plunkett and Marchman, 1993; Plunkett and Juola, 1999) (but see the triangle model of Seidenberg and Gonnerman (2000) for a non-derivational connectionist approach).

On the other hand, the approach advocated by Stemberger is very much in the spirit of a line of research in linguistic morphology that has not received much attention from psychologists, word and paradigm morphology (see, e.g. Matthews, 1974; Aronoff, 1994; Beard, 1995; Blevins, 2003). Word and paradigm morphology takes the word, rather than sub-word formatives, to be the basic meaningful unit of a language, and focuses on the paradigmatic organization of inflected words. Recent experimental work has provided experimental evidence for the relevance of paradigmatic organization for lexical processing. Across a series of experiments (Baayen et al., 2006; Milin et al., 2009a,b; Tabak et al., 2005), entropy measures gauging the amount of information carried by a paradigm have been found to be robust predictors of response latencies in visual comprehension studies. Rule-based models with a sparse lexicon listing only the irreducably irregular are severely challenged by these data. If experience with inflected forms would leave no trace in declarative memory, as the dual mechanism model would have us believe, then entropy measures calculated from the probability of a word's inflected forms should not be predictive for lexical processing, contrary to what is actually observed.

Thus far, relatively little is known about the role of paradigmatic structure in speech production. Baayen et al. (2008b) reported for Dutch an inhibitory effect of inflectional entropy in a picture naming task in which subjects were presented with line drawings with one or with two exemplars of objects, and were requested to say aloud the singular or the plural form. When the singular and the plural form have more similar frequencies, the entropy is large, and response latencies become large as well. When the two forms are more dissimilar in frequency, response times are shorter.

Tabak et al. (2010) observed effects of inflectional entropy for word naming in Dutch and English, in interaction with measures of neighborhood density. When subjects were requested to name the past-tense form when presented with the present-tense form, or the present-tense form when shown the past-tense form (cross-tense naming), the paradigmatic entropy entered into a negative correlation with the naming latencies, but significantly so only for irregular verbs.

For the study of speech production, naming paradigms are suboptimal because they have a substantial comprehension component. This component is so clearly present that simple word naming is often used as a task to gauge visual comprehension, complementing the visual lexical decision task. Cross tense word naming seeks to avoid giving away what the participant has to say by asking the subject to change the tense of the stimulus. However, participants are still provided with language input, and are required to perform a metalinguistic task which, as we found out, leads to the development of task-specific strategies (Tabak et al., 2010).

A task that is not contaminated by the reading process is the picture naming task. This task engages the full production process, from conceptualization to articulation. The goal of the present study is to ascertain to what extent the probabilities of inflectional forms can be detected when this ecologically more valid task is used. Effects of the probabilities of regular inflected forms (estimated by their relative frequency) as well as effects of the paradigmatic entropy (estimated from the distribution of relative frequencies in a word's inflectional paradigm) would support the non-derivational view of inflection proposed by Stemberger (2004) in psychology and by Word and Paradigm morphology in linguistics.

The following four experiments compare picture naming for present tense forms (experiments 1 and 3) and past tense forms (experiments 2 and 4), at the same time contrasting picture naming without (experiments 1 and 2) and with (experiments 3 and 4) prior familiarization with the intended picture names. We anticipated that in naming without prior familiarization, processes related to the interpretation of the picture and the retrieval of an appropriate picture name might dominate, while in naming with prior familiarization, processes related to word form encoding might be more prominent.

Experiment 1: Unprepared present-tense picture naming

Materials

The materials consisted of photographs of a young woman enacting verbs for position and motion. Photographs were commissioned for all picturable verbs available in the set of 286 verbs studied by Tabak et al. (2005) using visual lexical decision and Tabak et al. (2010) using word naming. Photographer and actress were instructed to minimize variability between pictures by using the same background and a highly restricted set of ancillary objects. Examples are shown in Figure 1. A total of 170 photographs was obtained, of which 85 depicted regular verbs, and 85 irregular verbs. Verbs ranged in log *Lemma Frequency* (the frequency of the verb across all its inflected variants) from 4.5 to 11.3 in a 42 million word corpus (median 8.0), and ranged in *Length* (in phonemes) from 2 to 6 (median 4). The irregular verbs were of slightly higher frequency than the regular verbs 8.2 versus 7.8, ($t(153.9) = 1.9076, p = 0.06$). The complexity of the pictures, henceforth *Picture Complexity*, evaluated in terms of the size (in kilobytes) of their jpg files, differed between the regulars (mean 93.6) and irregulars (mean 71.9, $t(153.4) = -6.06, p < 0.0001$). This difference mirrors the verbal report we received from the photographer and actress that the regular verbs were more difficult to depict than the irregular verbs.

[Figure 1 about here.]

Participants

Ten women and seven men, all students at the University of Nijmegen, participated in this experiment. They all had normal or corrected-to-normal vision. Subjects received 5 Euro for each sub-experiment.

Procedure

Participants were tested individually in a noise attenuated experimental booth. They received standard picture naming instructions, specifying that they had to name the presented picture as quickly and accurately as possible. Naming latencies were registered with a Sennheiser microphone placed at a distance of 20 cm from the participant.

Each picture was preceded by *Vandaag ...* ('Today ...') displayed on the computer screen. We asked the participants to complete the prompt with a clause consisting of the verb form followed by the third person pronoun, e.g., *loopt ze* ('she walks') for the photograph depicting walking. This prompt was presented in the center of the screen for 1000 ms. After 50 ms, the photograph was shown, in portrait mode, using the full vertical dimension of the computer screen. Photographs remained on the screen for 3000 ms. A new trial was initiated 500 ms afterwards. Each participant was presented with the pictures in a different random order.

There were three short breaks during the experiment. The total duration of an experimental session was approximately 45 minutes.

Results

We removed data points with responses where the voicekey was triggered by vocalizations other than those of the onset of a real word (1.1% of the observations). The distribution of the remaining latencies was highly skewed. We reduced this skewness by means of a logarithmic transformation, which outperformed the inverse transform. We fitted a linear mixed-effects model to the data with participant and verb as crossed random effects. Factors were modeled with contrast coding. The residuals of this model showed marked deviations from normality. We therefore removed data points with absolute standardized residuals exceeding 2.5, and refitted the model. The qualitative pattern in this trimmed model is the same, but the estimates of the coefficients and their standard deviations are more precise. The same procedure was followed in the analyses of the other three experiments discussed below. For all analyses, we used a stepwise variable elimination procedure to obtain the most parsimoneous model providing a close fit to the data. Predictors that did not reach significance were removed from the model specification. Many two-way interactions, and occasionally three-way interactions, were examined. Again, only those were retained that reached significance. For the present experiments, in which many new or not well-established predictors are considered, we reasoned that a conservative strategy allowing the model becoming clogged with superfluous main effects and interactions would be counterproductive. In the tables and figures presented below, only those predictors and interactions are presented that reached significance. In other words, tables and figures represent the minimally adequate models that we fitted to the data. Tables (and figures) should therefore not be consulted as specifying for the full set of predictors and their interactions whether they were significant.

As predictors, we considered several measures in addition to *Lemma Frequency*, *Picture Complexity*, and *Regularity*. First, two control variables were included: *Trial* (the index of the item in the experimental list), and the response latency to the preceding trial (*Previous RT*). The former measure allows us to explicitly account for effects of habituation or fatigue. The second measure is necessary to account for potential dependencies between the successive trials in the experiment (see, e.g., De Vaan et al., 2007; Balling and Baayen, 2008; Tabak et al., 2010). Counterbalancing nullifies the adverse effects of non-independencies in the sequence of trials. By including these two control measures, we remove at least part of the variance associated with this non-independence from the error term, and thereby obtain more precise models.

Supplementing *Lemma Frequency*, we included as a predictor the frequency of the present-tense form. As this frequency is highly correlated with *Lemma Frequency* ($r = 0.92$), we orthogonalized it by replacing it by the residuals of a linear model regressing form frequency on lemma frequency. The residualized measure was positively correlated with the original frequency ($r = 0.35, p < 0.0001$) and represents the inflectional form's frequency in so far as that frequency cannot be predicted from lemma frequency.

A related predictor was the verb's *inflectional entropy* (Baayen et al., 2006; Milin et al., 2009b; Tabak et al., 2010), the amount of information carried by the verb's paradigm. It was estimated as

$$H_i = \sum_j p_j \log_2 p_j, \tag{1}$$

with j ranging over the different phonologically distinct inflectional variants of the verb, and p_j representing the relative frequency of that inflectional variant in its paradigm.

In addition to the factor *Regularity*, distinguishing between regular and irregular verbs, we included the factor *Sex*, distinguishing between female and male speakers, as Ullman et al. (2002) reported for English that females but not males would have representations in declarative memory for regular past-tense inflections.

As subjects received no initial instructions on what names to use for the pictures, the names actually produced included many that were not targeted. We refer to trials for which the targeted form was produced as "correct" trials (61% of responses) and to the untargeted forms as "incorrect" (39% of the responses). As a measure of uncertainty about how to name the picture, we calculated the entropy of the frequency distribution of names elicited by a given picture,

$$H_p = \sum_k p_k \log_2 p_k, \tag{2}$$

with k ranging over the different names produced, and p_k the relative frequency with which name k was produced in the experiment. In what follows, we refer to this entropy measure as *Picture Entropy*. For all unintended, "incorrect" responses, lemma frequency, present-tense form frequency, inflectional entropy, and length in phonemes were determined, so that distributional measures were always tied to the form actually produced. In our statistical analysis, we build on the robustness of linear mixed-effects models with respect to unequal numbers of observations (see, e.g., Baayen et al., 2008a): words with fewer replicates contribute less weight to the model's estimates. By including the untargeted responses, we minimize data loss.

A predictor that we also considered specifies whether a verb has a stem-final obstruent that alternates between voiced (in, e.g., the infinitive: *schrijven*, 'to write') and voiceless (in, e.g., the singular present-tense forms: *schrijft*, 'you, he, she writes'). For experimental studies addressing this alternation, the reader is referred to Ernestus and Baayen (2003, 2004, 2007). Tabak et al. (2010) observed that whether a verb has an alternating obstruent may affect response latencies in word naming. We therefore included the factor *Alternating*, with levels 'alternating' and 'non-alternating', as a predictor in our analyses.

In addition, across all four experiments, we considered whether measures of neighborhood density (the N-count measure, as well as the positional neighborhood measures explored by Bien et al. (2005)) might help explain the variance in naming latencies. However, none of these measures ever reached or approached significance, not as simple main effects, nor as participants in interactions. In what follows, they will therefore not be discussed further. Similarly, the length (in phonemes) of the target word never reached significance, and is not mentioned in the analyses.

[Table 1 about here.]

[Figure 2 about here.]

The model fitted to the data of Experiment 1 is summarized by Table 1, which lists the coefficients of the fixed-effect factors and covariates, and Figure 2. The standard deviation for the by-subject random intercepts was 0.088, that for the by-item random intercepts was 0.158, and the standard deviation for the by-observation noise was 0.241.

Panel A shows that as subjects proceeded through the experiment, they responded more slowly. Panel B illustrates that more complex pictures elicited longer latencies, as expected. As shown in panel C, subjects who responded more slowly to a preceding picture tended to take longer for responding to the current trial, exactly as observed in previous studies examining the non-independence of adjacent trials in other chronometric tasks (De Vaan et al., 2007; Balling and Baayen, 2008; Tabak et al., 2010).

The interaction of *Correctness* by *Sex* shown in panel D indicates that males required more time than females when responding with an untargeted, “incorrect” picture name. Panel E illustrates a second difference between the sexes: The effect of *Picture Entropy*, the uncertainty (or amount of information) carried by the picture was greater for females than for males.

Panel F reports a U-shaped effect for *Lemma Frequency*. Such a U-shaped frequency effect has been reported previously for language production by Bien et al. (2005) and Tabak et al. (2010).

Panel G illustrates the effect of *Inflectional Entropy*. For targeted names (‘correct’), accessing informationally more complex paradigms required more time, replicating Baayen et al. (2008b). For untargeted verbs, the pattern reversed, such that alternative names were selected more quickly as the amount of information carried by their inflectional paradigms increased. This suggests that information-rich paradigms may constitute attractors competing with the target verbs for selection in production.

Experiment 2: Unprepared past-tense picture naming

Materials

The materials were the same as those of Experiment 1.

Participants

Ten women and seven men completed Experiment 2. None of them had participated in Experiment 1. They all had normal or corrected-to-normal vision.

Procedure

The procedure was identical to that of Experiment 1, except that subjects were now asked to produce simple past-tense forms. Therefore, the prompt used in Experiment 1 (*vandaag*, ‘today’) was replaced by a prompt for past-tense forms, *gisteren*, ‘yesterday’. We asked participants to complete the phrase with the appropriate verb form, followed by the pronoun *ze*, ‘she’, as in *gisteren liep ze*, ‘yesterday she walked’.

Results

Table 2 and Figure 3 present the results of a mixed-effects model fitted to the naming latencies of Experiment 2. As for Experiment 1, untargeted verbs (41% of the trials) were included in the analysis, with predictors such as lemma frequency, past tense frequency, and inflectional entropy recalculated for these forms. Standard deviations for the by-subject random intercepts, the by-item random intercepts, and the by-observation noise were 0.117, 0.140, and 0.260 respectively.

[Table 2 about here.]

[Figure 3 about here.]

As in Experiment 1, naming latencies increased as subjects proceeded through the experiment, as shown in panel A of Figure 3. Panel B replicates the interaction of *Picture Entropy* by *Sex*. Again, the inhibitory effect of *Picture Entropy* was stronger for females than for males. Panel C indicates that *Lemma Frequency* had a U-shaped effect, exactly as in Experiment 1.

Verb forms with a greater past-tense inflectional frequency elicited shorter naming latencies, as illustrated in panel D. Furthermore, the larger the number of irregular verbs with the same rhyme (in the past tense) as the verb produced, the shorter the naming latency was (panel E; this predictor was also examined for Experiment 1, but was not significant).

Experiment 3: Prepared present-tense picture naming

Materials The materials were the same as in Experiment 1.

Participants The participants were the same as those participating in Experiment 1.

Procedure After completing Experiment 1, participants were familiarized with the intended, targeted picture names by taking them through a picture book that printed the infinitive below each photograph. All that we asked the participants to do is silently read the printed words, no overt response was elicited. Following this familiarization phase, we carried out a second naming experiment in which participants were asked to use the targeted picture names. We refer to this second version of the experiment as *prepared naming*.

Results

The statistical analysis of the data proceeded along the same lines as for Experiments 1 and 2. The familiarization phase reduced the percentage of untargeted responses from 39% (in Experiment 1) to 16%. Untargeted, incorrect responses were removed from the data set. Table 3 and Figure 4 summarize the results of a mixed-effects model fitted to the correct, targeted responses. The standard deviation of the by-subject random intercepts was 0.118. Random slopes for (centralized) *Lemma Frequency* also reached significance (standard deviation: 0.009) in a likelihood ratio test ($\chi^{(2)}_2 = 6.46, p = 0.039$). The standard deviation of the by-item random intercepts was 0.118, and that of the by-observation noise was 0.217.

[Table 3 about here.]

[Figure 4 about here.]

Table 3 presents the coefficients of the fitted model, and Figure 4 illustrates the partial effects of the predictors. *Trial* again emerged as inhibitory (panel A). As expected given Experiment 1, naming latencies increased with *Picture Complexity* (panel B). *Picture Entropy* was likewise inhibitory, this time without an interaction with *Sex* (panel C). *Lemma Frequency* emerged as inhibitory (panel D). There was significant by-subject variation in the slope of this frequency effect, as mentioned above in the discussion of the random-effects structure of the model. Inspection of the by-subject slopes revealed that all estimated slopes were positive.

The inflectional frequency of the present-tense form, which did not reach significance in Experiment 1, emerged in a significant interaction with *Sex*. As can be seen in panel E, words with greater inflectional frequency were responded to faster by males, but slower by females.

Words with final obstruents alternating with respect to their voice specification elicited longer naming latencies (panel F). Panel G depicts the facilitation from the *RhymeCount* measure: verbs that rhyme in the present tense with many other irregular verbs elicited shorter latencies.

Experiment 4: Prepared past-tense picture naming

Materials

The materials were identical to those used in Experiments 1–3.

Participants The participants were the same as those participating in Experiment 2.

Procedure After completing Experiment 2, participants were familiarized with the targeted picture names by taking them through a picture book that printed the infinitive below each photograph. Following this familiarization phase, we carried out a second, prepared naming experiment in which subjects were requested to use the targeted picture names.

Results

[Table 4 about here.]

[Figure 5 about here.]

We analyzed the data in the same way as for Experiment 3, focusing on the targeted, correct responses, and excluding the untargeted responses (15.1%) from consideration. Table 4 and Figure 4 provide an overview of the model. Standard deviations for the random intercepts for subject, item, and the by-observation noise were 0.079, 0.124, and 0.248 respectively.

The inhibitory effect of *Trial* observed for the preceding three experiments was again present in Experiment 4 (panel A). For regulars, but not for irregulars, longer latencies at preceding trials predicted longer latencies at the current trial (*Previous RT*, panel B). *Picture Entropy* (panel C) was again inhibitory, as in all three preceding experiments. No interaction with *Sex* could be observed. Panel D shows the U-shaped effect of *Lemma Frequency* familiar from Experiments 1 and 2. The frequency of the past-tense inflected form was facilitatory (panel E). There was no trace of interactions with *Sex* or *Regularity*. Panel F illustrates the interaction of *Inflectional Entropy* by *Sex*. For males, but not for females, a greater *Inflectional Entropy* afforded significantly shorter naming latencies. Finally, verbs with alternating final obstruents elicited longer naming latencies (panel G), but this difference was modulated by *Previous RT*, such that for larger preceding response latencies the effect of *Alternating* diminished.

Discussion

Table 5 presents a synopsis of the effects observed. The main patterns in this table are supported by an analysis of all four experiments jointly, summarized in Table 6. This overall analysis offers the advantage of increased power and the possibility to evaluate for which predictors and interactions elimination from the individual model specifications was too conservative, as well as for which predictors support was restricted to just a single data set. The disadvantage of the overall analysis is that the data set is somewhat unbalanced, as the data elicited for prepared naming excluded untargeted responses.

Consistent across all four experiments are the increasing naming latencies as the experiment proceeds (*Trial*), and the elongated latencies that come with more ambiguous pictures (*Picture Entropy*).

Table 5 shows that in present-tense naming, but not in past-tense naming, more complex pictures (*Picture Complexity*) elicited significantly longer latencies, according to the individual analyses. Table 6 suggests that there is an inhibitory effect for past-tense naming as well, but with a coefficient that is half that for present-tense naming. The interaction of *Picture Complexity* by *Tense* is probably due to reduced morphological processing load for present-tense compared to past-tense naming, and a concomitant reduction in the noise masking the early stages of picture interpretation.

Consistent with earlier studies, *Previous RT*, when it emerged in the individual analyses (Experiments 1 and 4), had a positive slope, indicating that there is local coherence in the speed with which subjects respond in chronometric tasks. The joint analysis supports this temporal coherence effect across all four experiments.

The analyses of the individual experiments revealed a significant effect of *Alternating* only for prepared naming. The joint analysis supports delayed latencies for words with alternating obstruents across all four experiments, in interaction with *Previous RT* (as in Experiment 4). This suggests that words with more variable morphophonology are at a disadvantage in speech production. This disadvantage is strongest when subjects are naming the pictures quickly, as indexed by *Previous RT*, and decreases where they go through the experiment more slowly. Apparently, selecting the correct phonological form slows processing only when the choice between alternatives has to be made rapidly.

When the count of rhyming irregulars was significant (Experiments 2 and 3), it was facilitating. In Experiment 2 (present-tense naming), we counted the number of irregular verbs sharing the present-tense rhyme, for Experiment 3 (past-tense naming), the *RhymeCount* is based on the number of irregular verbs sharing the same rhyme in the past tense. This effect, which in the joint analysis received full support across all four experiments, is probably best interpreted as a facilitatory neighborhood density effect.

The effect of neighborhood density in speech production is somewhat unclear. Vitevitch (2002) observed facilitation for English, but Vitevitch and Stamer (2006) were confronted with inhibition for Spanish. Bien et al. (2005) found an inverse U-shaped effect of neighborhood density in the production of compounds in Dutch, and Tabak et al. (2010) observed inhibition for neighbors differing only in their first phoneme, but facilitation for neighbors differing at later positions. In the present experiments, neighborhood density measures consistently failed to reach significance. One possible reason is that the picture naming task is dominated by conceptual processing and the retrieval of the picture name. The effect of neighborhood density would then presumably arise during later stages of phonological encoding. Another possible reason is that measures of neighborhood similarity based on a single mismatching segment may be too coarse. By contrast, the *RhymeCount* measure effectively defines neighbors on the basis of a mismatching higher unit, the onset of the stem syllable. If this interpretation is correct, it supports the hypothesis of Vitevitch (2002) that in speech production neighborhood similarity is facilitatory.

[Table 5 about here.]

In three out of four experiments, *Lemma Frequency* emerged with a U-shaped effect. In the context of experiments 1, 2, and 4, the linear inhibitory effect of *Lemma Frequency* in Experiment 3 (prepared present-tense naming) is exceptional. However, the joint model fitted to all four experiments jointly suggests that the coefficient for the quadratic term of *Lemma Frequency* for Experiment 3 does not differ significantly from the corresponding coefficients for the other Experiments. This suggests that the true *Lemma Frequency* effect in Experiment 3 is also U-shaped, and warns against overinterpreting the *Lemma Frequency* effect in Experiment 3 as special.

Following Tabak et al. (2010), who observed the same U-shaped pattern for word naming, we hypothesize that this U-shaped effect arises as a consequence of subjects optimizing their performance for verbs with the most likely, 'central', lemma frequencies. The more exceptional a lemma frequency is in the experiment, irrespective of whether it is large or small, the longer the response latencies are. In other words, if this hypothesis is on the right track, we are observing the brain's response not to a given lemma frequency as such, but to the probability of that lemma frequency. Such a higher-level response would be consistent with task-specific optimization, with optimization for general lexical processing in speech production, or with both.

[Table 6 about here.]

In order to evaluate the effect of *Inflectional Entropy*, we first note that it has been found to be facilitatory in comprehension (Baayen et al., 2006) but inhibitory in picture naming (Baayen et al., 2008b). In the naming experiments reported by Tabak et al. (2010), the effect of *Inflectional Entropy* was modulated by neighborhood density. No such modulation could be observed in the present picture naming experiments. In our picture naming experiments, the evidence for a role for *Inflectional Entropy* is somewhat fragmented across the individual experiments. The joint analysis suggests facilitation for untargeted responses, and inhibition for regular verbs. A three-way interaction in the joint analysis of *Sex* by *Inflectional Entropy* by *Preparedness*, just missing significance ($t = 1.98$), provides modest support for the interaction with *Sex* observed in Experiment 4, with stronger inhibition for females. The main trend emerging from the present data is that in picture naming *Inflectional Entropy* is inhibitory, replicating the picture naming study of plurals reported by (Baayen et al., 2008b). The reversal of this effect into facilitation for untargeted responses suggests that words with more complex inflectional paradigms serve as more powerful attractors during lemma selection.

Given the results reported by Ullman et al. (2002) for English, we explored potential differences in lexical processing between males and females. Ullman and colleagues hypothesized that the superior verbal cognitive skills of females (see Kimura, 2000, for a review) allow frequency effects for regular inflected forms to arise in females, but not in males. The present data offer tentative support for modest differences in verbal processing between females and males. However, support for a difference in sensitivity to *Form Frequency* is quite weak, as it is only in Experiment 3 that a frequency-related sex difference reached significance.

In Experiment 3, females revealed an inhibitory effect of the frequency of the present-tense form, whereas for males the slope of form frequency did not differ significantly from zero. This finding is consistent with the hypothesis that females, but not males, store regular inflected forms. However, given that in the past-tense naming experiments the effect of form frequency was significant for both sexes, and not modulated by an interaction with *Sex*, nor by an interaction with *Regularity* ($t = 0.18$ for the three-way interaction of *Sex* by *Regularity* by *Tense* in the joint analysis), it cannot be argued that males do not remember regular inflected forms. In fact, frequency effects may be weaker for males than for females, but nevertheless significantly present (see Balling and Baayen, 2008, 2009, for evidence from Danish).

Unfortunately, this frequential differences between females and males visible in the isolated analysis of Experiment 3 did not reach significance in the joint analysis ($t = 1.79$ for the interaction of *Sex* by *Tense* by *Form Frequency*). This might be due to an effect specific to prepared past-tense naming being washed out by the other three experiments. But it is equally likely that the effect in Experiment 3 is a false positive. The only robust pattern that emerges from our data, with full support from the joint analysis, is that for females the inhibitory effect of *Picture Entropy* was stronger. Females may have considered more alternatives for naming than males, consistent with the superior female verbal memory.

Conclusions

The general question addressed in this study is whether past-tense inflected forms are generated on-line from the present-tense stem during speech production. Most linguistic theories, especially those in the generative tradition (Pinker, 1999; Pinker and Ullman, 2002) claim that inflected forms are "derived" in this generative sense. Only a minority of linguists have considered non-derivational models (Bybee, 1985; Matthews, 1974; Blevins, 2003).

According to Pinker's dual mechanism model, the past-tense suffixation rule is applied only if a search for an irregular past-tense form fails. This ordering of suffixation after lexical search ensures that irregular verbs are not inflected regularly. Given this ordering, one would expect regular past-tense forms to require longer processing latencies than irregular past-tense forms. However, no clear disadvantage emerged for regular verbs. The joint analysis actually suggests that for all but the largest inflectional entropies, regular verbs were named faster than irregular verbs, instead of more slowly.

A further problem for the dual mechanism model is that in its original conceptualization (e.g., Pinker, 1991), it predicted that regular past-tense forms leave no traces in lexical memory, and hence that effects of frequency of use should not be observable for regular past-tense forms. However, frequency effects for past tense forms emerged in past-tense picture naming, irrespective of regularity. Effects of inflectional form frequency for regular inflected forms were also observed by Tabak et al. (2010) for Dutch and English in word naming tasks.

Later versions of the dual mechanism model have relaxed the claim that frequency effects should be limited to irregular past-tense forms (Pinker, 1999; Pinker and Ullman, 2002). For instance, for regular past-tense forms that are similar in form to irregular past-tense forms, a frequency effect is now posited, the idea being that storing the regular form in memory would protect this form against irregularization. However, our data do not support the hypothesis that a form frequency effect would be restricted just to regulars that are phonologically similar to irregular verbs.

Following the joint analysis of all four experiments, we observed a facilitatory frequency effect of the inflected form in unprepared and prepared past-tense naming. The slope of this frequency effect is the same for regulars and irregulars. Independently, we observed an effect of the count of rhyming irregulars, which also did not vary across regular and irregular verbs. This suggests that, at least in Dutch, the effect of form frequency in past-tense picture naming is not crucially dependent on phonological similarity to irregulars. In addition, the inhibitory effect of *Inflectional Entropy* (see also Tabak et al., 2010) provides further evidence against a sparse lexicon, as it provides evidence, albeit indirect evidence, for information in long-term memory of the likelihoods of individual inflected forms.

Finally, consider the effect of the voicing alternation (whether the stem-final obstruent alternates with respect to the feature voice) in prepared picture naming. This effect was not modulated by regularity, and presented itself for present-tense and paste-tense naming alike. According to derivational theories of this alternation, an underlying form is posited with a voiced obstruent. This voiced obstruent is visible in the past-tense forms of regular verbs, but becomes voiceless in all other forms probed in our experiments. Classical derivational theory (see, e.g., Shane, 1973) would lead one to expect an interaction of *Alternating* by *Regularity* in past-tense naming, with the forms with a voiceless final obstruent revealing the longest naming latencies (due to the application of a rule of devoicing). No such interaction was present, however — the effect of *Alternating* was observed even for present-tense naming, in which all final obstruents were realized as voiceless. While derivational theories incorrectly predict no difference for the present tense, the non-derivational theory of Ernestus and Baayen (2003, 2004, 2006, 2007) can accomodate this finding as reflecting uncertainty about the stem-final voicing across lexical paradigms. Note that *Alternating* actually represents a second layer of irregularity orthogonal to the standard distinction between regular and irregular verbs based on vocalic alternation of the stem. Just as irregularity gives rise to longer picture naming latencies, voice alternation is time-costly.

Considered jointly, these findings challenge the dual mechanism model, while fitting well with Word and Paradigm Morphology in linguistics and non-derivational approaches to inflection in psychology (Stemberger, 2002; Stemberger and Middleton, 2003; Stemberger, 2004).

The present experiments also shed some light on the consequences for lexical processing of differences in semantic density for regular and irregular verbs. According to the dual mechanism model, differences between regular and irregular verbs are restricted to phonological form. However, recent studies (Ramscar, 2002; Patterson et al., 2001) suggest that semantic and contextual factors are also relevant for understanding how regular and irregular verbs are processed. Furthermore, a lexical statistical survey (Baayen and Moscoso del Prado Martín, 2005) revealed that regulars and irregulars differ in semantic density: Irregulars tend to have denser semantic networks than regulars. For instance, when regulars and irregulars are matched for frequency, irregular verbs tend to have more meanings than regular verbs. Regular and irregular verbs also tend to have different aspectual properties, as witnessed by the non-uniform distribution of auxiliary verbs in Dutch and German. Regulars favor *hebben*, 'have', while irregulars favor *zijn*, 'be', the auxiliary marking telicity. Regular and irregular verbs are also non-uniformly distributed over Levin's verb argument alternation classes (Levin, 1993). In the analyses of the present picture naming experiments, we examined various predictors gauging differences in semantic density, such as the count of synonyms in WordNet (Miller, 1990). Whereas Tabak et al. (2005) observed this measure to be predictive for visual comprehension, it failed to reach significance for the picture naming task.

Nevertheless, the present study does provide subtle distributional evidence that regulars and irregulars differ in semantic density. The artists who made the photographs used in the picture naming experiment reported that regular verbs were especially difficult to enact. This informal observation is supported by two observations. First, the mean size of the JPG files was larger for the regulars than for the irregulars ($t(165.5) = -6.4157, p < 0.0001$). In order to depict regular verbs, more complex postures and more ancillary attributes were required. Second, irregulars were characterized by substantially smaller Picture Entropy ($t(165.5) = -4.7281, p < 0.0001$). Apparently, the artists were much more succesful in creating unambiguous pictures for irregular verbs compared to regular verbs. In hindsight, this is not surprising, as irregular verbs in Dutch tend to denote primary positions movements and actions of the body, e.g., *zitten* (sit), *staan* (stand), *liggen* (lie), *lopen* (walk), *duiken* (dive), *zwemmen* (swim), *slaan* (hit), *kijken* (look) (cf Baayen, 2007). Given this unequal distribution of regulars and irregulars across the Picture Complexity and Picture Entropy measures, and given that these two measures are invariably inhibitory across our picture naming experiments, we conclude that, at least in our data, irregular verbs had a processing advantage compared to regulars at the conceptual and semantic levels of processing.

Our experiments also call attention to the role of episodic memory in the picture naming task. In the present experiments, the costs of interpreting the picture, gauged by the compressed file size of the photograph (*Picture Complexity*) was smaller in prepared present-tense and prepared past-tense naming than for unprepared naming, unsurprisingly. Nevertheless, memories for picture names established during the familiarization with the pictures and their intended names may mask interesting differences that characterize more normal processing circumstances that do not depend on prior familiarization. For the present data, the joint analysis suggests that preparation attenuated the processing costs of *Alternating*. Furthermore, given that the *Picture Complexity* for regular verbs was greater than for irregular verbs, preparation may have been disproportionally facilitating for regular verbs verbs. Thus, for the investigation of speech production, unprepared and prepared naming both have advantages and disadvantages to offer. Unprepared picture naming may be more revealing for processes preceding form selection, while prepared picture naming may be more sensitive to form selection and subsequent lexical processing.

The role of *Regularity* as factorial predictor representing what according to Pinker (1997) would be a fundamental organizing principle of human language, is surprisingly small. In the joint analysis, regular verbs are assigned a smaller intercept than irregulars, as well as a greater slope for *Inflectional Entropy*. *Regularity* does not interact with any of the other predictors. On the one hand, this very modest role for *Regularity* may indicate a lack of sensitivity of the experimental paradigm of picture naming. On the other hand, it may also indicate that with the current array of item-bound predictors we have succeeded in capturing those differences between regulars and irregulars that are crucially involved in speech production, rendering *Regularity* as a dichotomous factor largely superfluous.

References

Aronoff, M. (1994). *Morphology by itself: stems and inflectional classes.* The MIT Press, Cambridge, Mass.

Baayen, R. H. (2007). Storage and computation in the mental lexicon. In Jarema, G. and Libben, G., editors, *The Mental Lexicon: Core Perspectives*, pages 81–104. Elsevier.

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

Baayen, R. H., Feldman, L., and 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., Schreuder, R., and Ernestus, M. (2008b). Paradigmatic structure in speech production. *Proceedings Chicago Linguistics Society 43*, 1:1–29.

Baayen, R. H. and Moscoso del Prado Martín, F. (2005). Semantic density and past-tense formation in three Germanic languages. *Language*, 81:666–698.

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

Balling, L. and Baayen, R. H. (2009). Evidence for morphologically complex words in a redundant lexicon. Manuscript submitted for publication.

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

Bien, H., Levelt, W., and Baayen, R. (2005). Frequency effects in compound production. *Proceedings of the National Academy of Sciences of the USA*, 102:17876–17881.

Blevins, J. P. (2003). Stems and paradigms. *Language*, 79:737–767.

Bloch, B. (1947). English verb inflection. *Language*, 23:399–418.

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

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

De Vaan, L., Schreuder, R., and Baayen, R. H. (2007). Regular morphologically complex neologisms leave detectable traces in the mental lexicon. *The Mental Lexicon*, 2:1–23.

Ernestus, M. and Baayen, H. (2006). The functionality of incomplete neutralization in Dutch: The case of past-tense formation. In Goldstein, L. M., Whalen, Douglas, H., and Best, C. T., editors, *Laboratory Phonology 8*, pages 27–49. Mouton de Gruyter, Berlin.

Ernestus, M. and Baayen, R. (2007). Paradigmatic effects in auditory word recognition: The case of alternating voice in Dutch. *Language and Cognitive Processes*, 22:1–24.

Ernestus, M. and Baayen, R. H. (2003). Predicting the unpredictable: Interpreting neutralized segments in Dutch. *Language*, 79:5–38.

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

Kimura, D. (2000). *Sex and Cognition*. The MIT Press, Cambridge, MA.

Levin, B. (1993). *English Verb Classes and Alternations. A preliminary Investigation*. The University of Chicago Press, Chicago.

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

Matthews, P. H. (1974). *Morphology. An Introduction to the Theory of Word Structure*. Cambridge University Press, London.

Milin, P., Filipović Durdević, D., and Moscoso del Prado Martín, F. (2009a). The simultaneous effects of inflectional paradigms and classes on lexical recognition: Evidence from serbian. *Journal of Memory and Language*, pages 50–64.

Milin, P., Kuperman, V., Kostić, A., and Baayen, R. (2009b). Paradigms bit by bit: an information-theoretic approach to the processing of paradigmatic structure in inflection and derivation. In Blevins, J. P. and Blevins, J., editors, *Analogy in grammar: form and acquisition*, pages 214–252. Oxford University Press, Oxford.

Miller, G. A. (1990). Wordnet: An on-line lexical database. *International Journal of Lexicography*, 3:235–312.

Patterson, K., Lambon Ralph, M. A., Hodges, J. R., and McClelland, J. L. (2001). Deficits in irregular past-tense verb morphology associated with degraded semantic knowledge. *Neuropsycologia*, 39:709–724.

Pinker, S. (1991). Rules of language. *Science*, 153:530–535.

Pinker, S. (1997). Words and rules in the human brain. *Nature*, 387:547–548.

Pinker, S. (1999). *Words and Rules: The Ingredients of Language*. Weidenfeld and Nicolson, London.

Pinker, S. and Prince, A. (1988). On language and connectionism. *Cognition*, 28:73–193.

Pinker, S. and Prince, A. (1994). Regular and irregular morphology and the psychological status of rules of grammar. In Lima, S., Corrigan, R., and Iverson, G., editors, *The Reality of Linguistic Rules*, pages 353–388. John Benjamins, Amsterdam.

Pinker, S. and Ullman, M. (2002). The past and future of the past tense. *Trends in the Cognitive Sciences*, 6(11):456–462.

Plunkett, K. and Juola, P. (1999). A connectionist model of English past tense and plural morphology. *Cognitive Science*, 23(4):463–490.

Plunkett, K. and Marchman, V. (1993). From rote learning to system building: acquiring verb morphology in children and connectionist nets. *Cognition*, 48:21–69.

Ramscar, M. (2002). The role of meaning in inflection: Why the past tense doesn't require a rule. *Cognitive Psychology*, 45:45–94.

Rumelhart, D. E. and McClelland, J. L. (1986). On learning the past tenses of English verbs. In McClelland, J. L. and Rumelhart, D. E., editors, *Parallel Distributed Processing. Explorations in the Microstructure of Cognition. Vol. 2: Psychological and Biological Models*, pages 216–271. The MIT Press, Cambridge, Mass.

Seidenberg, M. S. and Gonnerman, L. M. (2000). Explaining derivational morphology as the convergence of codes. *Trends in Cognitive Sciences*, 4(9):353–361.

Shane, S. A. (1973). *Generative Phonology*. Prentice-Hall, New Jersey.

Stemberger, J. P. (2002). Overtensing and the effect of regularity. *Cognitive Science*, 26:737–766.

Stemberger, J. P. (2004). Phonological priming and irregular past. *Journal of Memory and Language*, 50:82–95.

Stemberger, J. P. and Middleton, C. (2003). Vowel dominance and morphological processing. *Language and Cognitive Processes*, 18:369–404.

Tabak, W., Schreuder, R., and Baayen, R. H. (2005). Lexical statistics and lexical processing: semantic density, information complexity, sex, and irregularity in Dutch. In Kepser, S. and Reis, M., editors, *Linguistic Evidence — Empirical, Theoretical, and Computational Perspectives*, pages 529–555. Mouton de Gruyter, Berlin.

Tabak, W., Schreuder, R., and Baayen, R. H. (2010). Inflection is not "derivational": Evidence from word naming. *Manuscript submitted for publication*.

Ullman, M. (2004). Contributionss of memory circuits to language: the declarative/procedural model. *Cognition*, 92:231–270.

Ullman, M. T., Estabrooke, I. V., Steinhauer, K., Brovetto, C., Pancheva, R., Ozawa, K., Mordecai, K., and Maki, P. (2002). Sex differences in the neurocognition of language. *Brain and Language*, 83:141–143.

Vitevitch, M. S. (2002). The influence of phonological similarity neighborhoods on speech production. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 28(4):735–747.

Vitevitch, M. S. and Stamer, M. K. (2006). The curious case of competition in Spanish speech production. *Language and Cognitive Processes*, 21(6):760–770.

List of Figures

Figure 1: Stimuli for *walking* (left) and *sitting* (right).

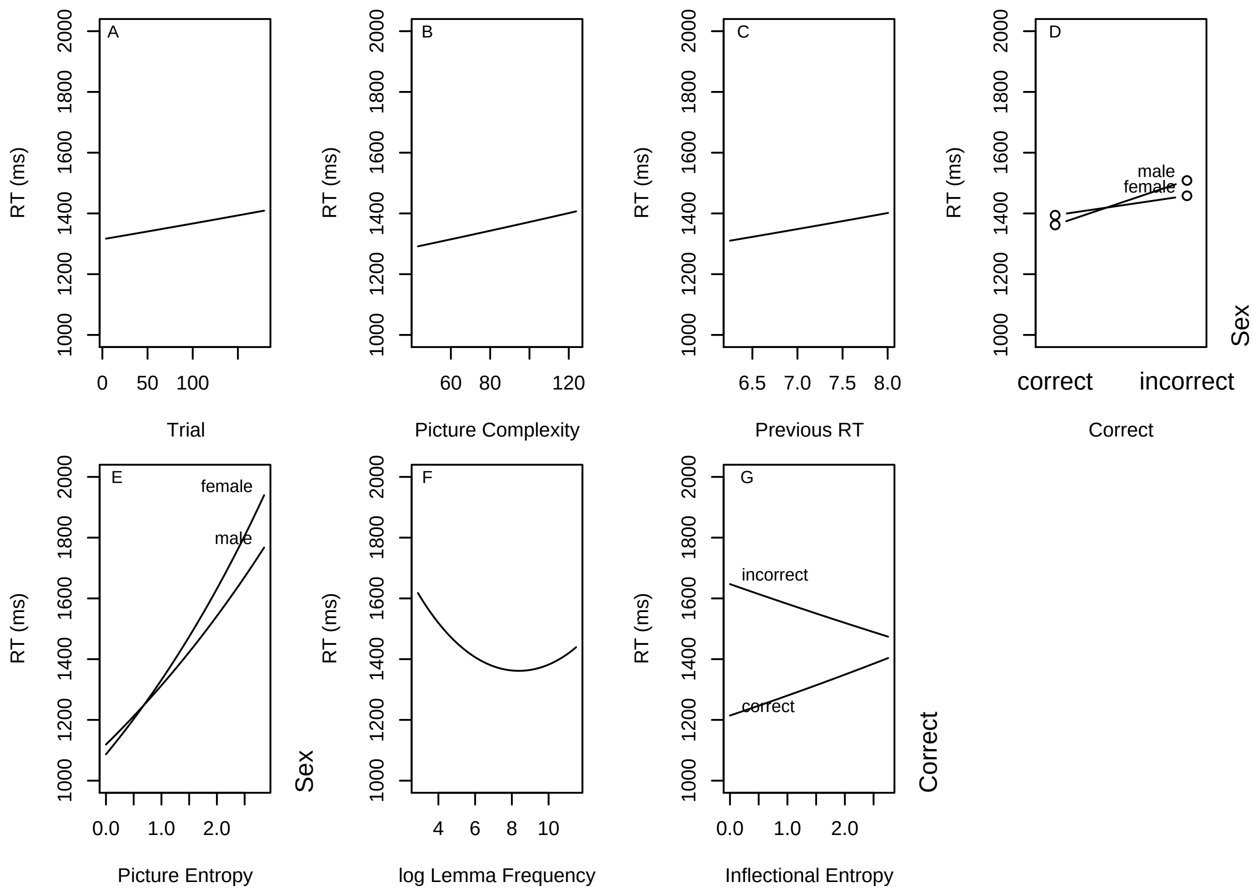

Figure 2: Partial effects of the predictors for the unprepared present-tense naming latencies in Experiment 1.

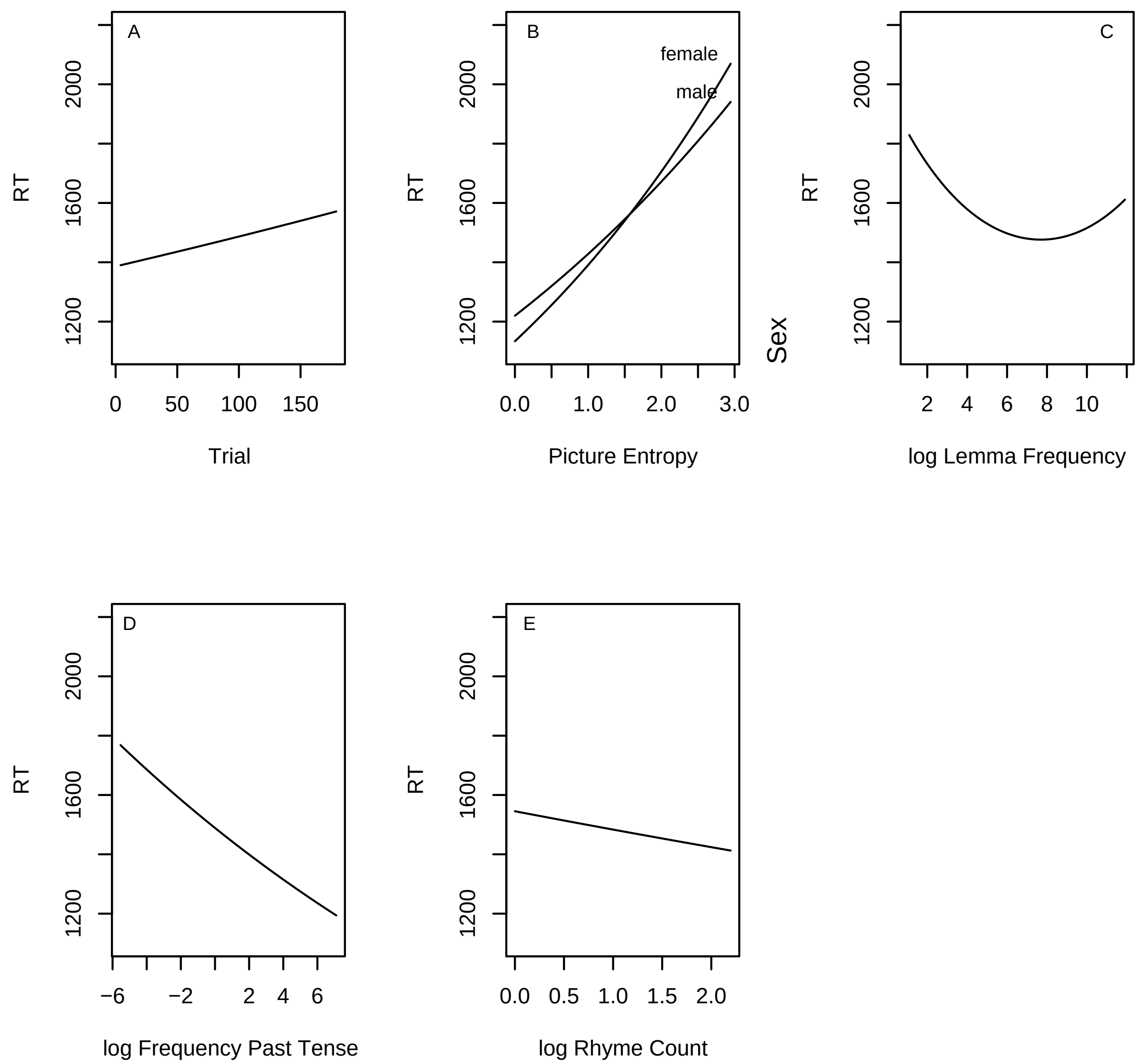

Figure 3: Partial effects of the predictors for the unprepared past-tense naming latencies in Experiment 2.

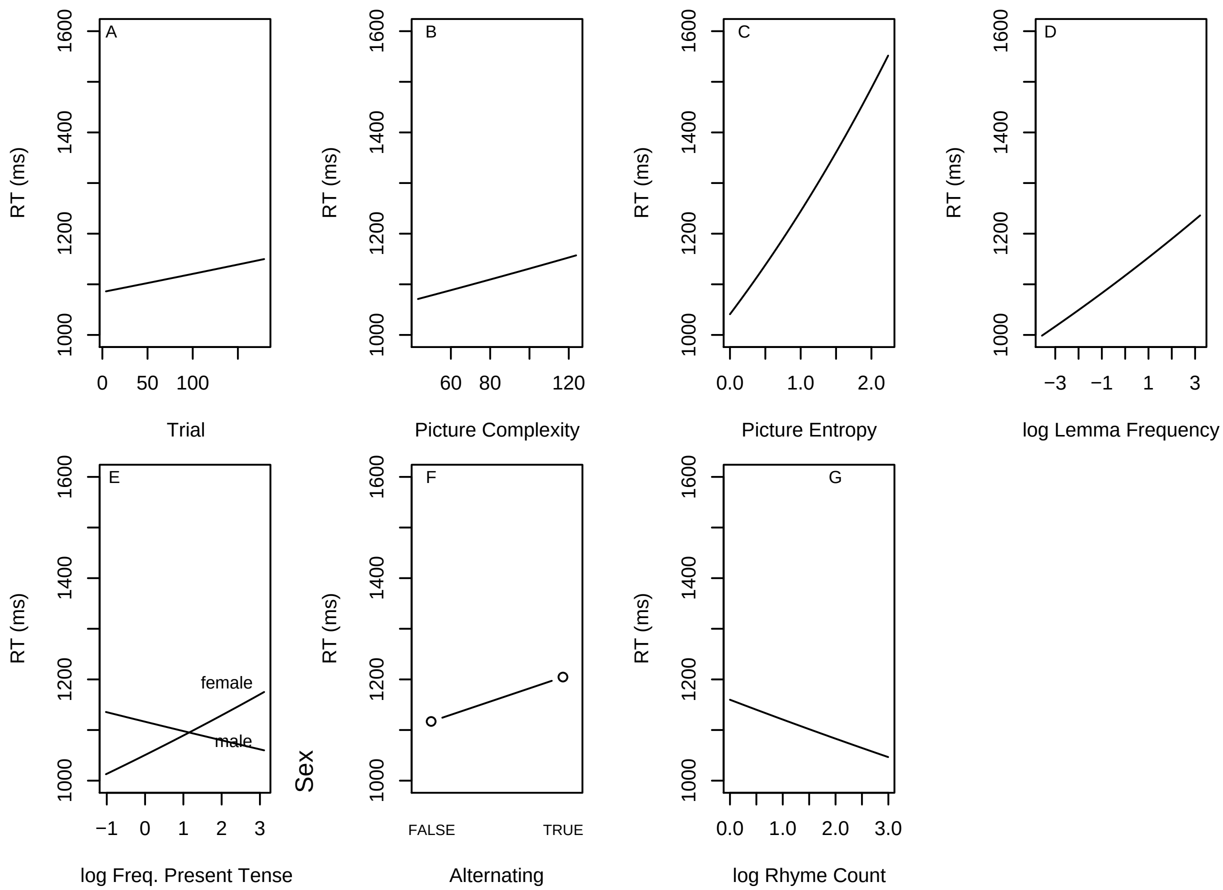

Figure 4: Partial effects of the predictors for the prepared present-tense naming latencies in Experiment 3. *Lemma Frequency* was centered.

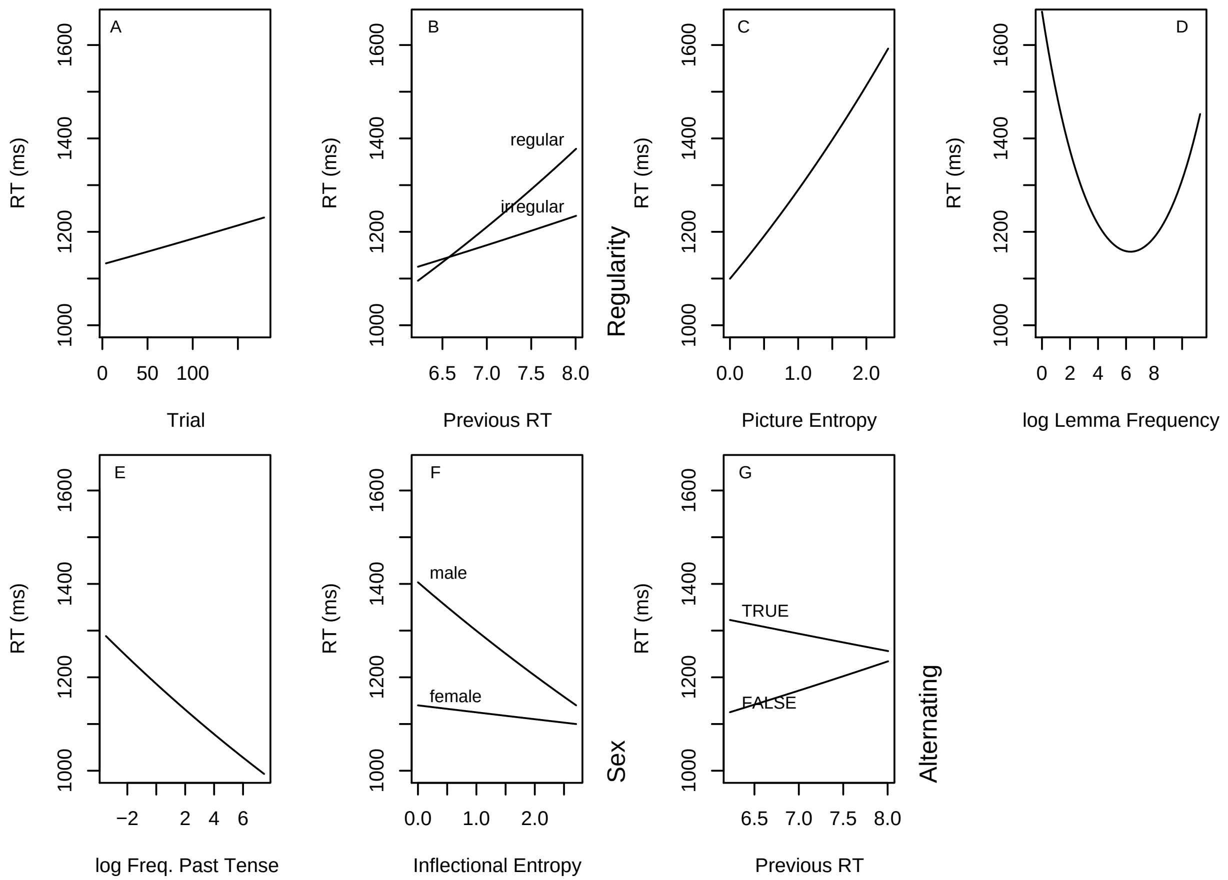

Figure 5: Partial effects of the predictors for the prepared past-tense naming latencies in Experiment 4.

List of Tables

	Estimate	lower HPD95	upper HPD95	p (MCMC)
Intercept	6.8938	6.5028	7.2891	0.0001
Trial	0.0004	0.0002	0.0006	0.0004
Previous RT	0.0384	0.0087	0.0744	0.0152
Picture Entropy	0.1603	0.1249	0.2009	0.0001
Picture Complexity	0.0011	0.0001	0.0020	0.0190
Correct=incorrect	0.3047	0.0958	0.4457	0.0030
Sex: contrast female	-0.0292	-0.1239	0.0666	0.5262
Inflectional Entropy	0.0526	-0.0178	0.0868	0.1992
Lemma Frequency (linear)	-0.0959	-0.1683	-0.0271	0.0118
Lemma Frequency (quadratic)	0.0057	0.0015	0.0103	0.0132
Correct=incorrect : Sex=female	-0.0573	-0.1048	-0.0045	0.0296
Picture Entropy : Sex=female	0.0429	0.0076	0.0732	0.0138
Inflectional Entropy : Correct=incorrect	-0.0929	-0.1534	0.0058	0.0622

Table 1: Coefficients in the mixed-effects model fit to the picture naming latencies of Experiment 1 (unprepared present-tense naming). Upper/lower HPD95: 95 percent credible intervals based on 10,000 Markov chain Monte Carlo samples from the posterior distribution of the parameters.

	Estimate	lower HPD95	upper HPD95	p (MCMC)
Intercept	7.3860	7.1978	7.6560	0.0001
Trial	0.0007	0.0005	0.0009	0.0001
Sex=female	-0.0731	-0.1897	0.0472	0.2222
RhymeCount	-0.0408	-0.0610	-0.0178	0.0004
Past Tense Frequency	-0.0311	-0.0488	-0.0148	0.0010
Picture Entropy	0.1576	0.1209	0.1907	0.0001
Lemma Frequency (linear)	-0.0758	-0.1398	-0.0336	0.0022
Lemma Frequency (quadratic)	0.0049	0.0022	0.0090	0.0018
Picture Entropy : Sex=female	0.0466	0.0146	0.0771	0.0024

Table 2: Coefficients in the mixed-effects model fit to the picture naming latencies of Experiment 2 (unprepared past-tense naming)

	Estimate	lower HPD95	upper HPD95	p (MCMC)
Intercept	6.8398	6.7356	6.9504	0.0001
Trial	0.0003	0.0001	0.0005	0.0010
Picture Entropy	0.1784	0.1428	0.2121	0.0001
Alternating=TRUE	0.0766	0.0381	0.1172	0.0002
Picture Complexity	0.0011	0.0004	0.0018	0.0038
Lemma Frequency (centered)	0.0306	0.0174	0.0442	0.0001
RhymeCount	-0.0059	-0.0104	-0.0016	0.0084
Present Tense Frequency	-0.0170	-0.0525	0.0164	0.3036
Sex=female	-0.0605	-0.1735	0.0464	0.2566
Present Tense Frequency : Sex=female	0.0518	0.0218	0.0820	0.0014

Table 3: Coefficients in the mixed-effects model fit to the picture naming latencies of Experiment 3 (prepared present-tense naming)

	Estimate	lower HPD95	upper HPD95	p (MCMC)
Intercept	7.1075	6.6139	7.5584	0.0001
Trial	0.0005	0.0003	0.0007	0.0001
Regularity=regular	-0.5060	-1.0148	-0.0129	0.0388
Sex=female	-0.2079	-0.3627	-0.0606	0.0062
Picture Entropy	0.1597	0.1250	0.1958	0.0001
Lemma Frequency (linear)	-0.1163	-0.1865	-0.0501	0.0022
Lemma Frequency (quadratic)	0.0092	0.0047	0.0137	0.0002
Alternating=TRUE	0.6652	0.1433	1.2079	0.0124
Previous RT	0.0519	-0.0021	0.1081	0.0630
Inflectional Entropy	-0.0768	-0.1502	-0.0067	0.0290
Past Tense Frequency	-0.0238	-0.0469	0.0004	0.0520
Previous RT : Alternating=TRUE	-0.0809	-0.1614	-0.0124	0.0278
Inflectional Entropy: Sex=female	0.0636	0.0082	0.1252	0.0318
Previous RT: Regularity=regular	0.0770	0.0093	0.1479	0.0214

Table 4: Coefficients in the mixed-effects model fit to the picture naming latencies of Experiment 4 (prepared past-tense naming)

Predictor	Exp. 1 present unprepared	Exp. 2 past unprepared	Exp. 3 present prepared	Exp. 4 past prepared
Trial	+	+	+	+
Picture Entropy	+ *Sex*	+ *Sex*	+	+
Picture Complexity	+ *Sex*		+	
Alternating			+	+− *Previous RT*
Previous RT	+			+ *Regularity, Alternating*
Rhyme Count		−	−	
Correctness	+− *Sex*			
Lemma Frequency	U	U	+	U
Form Frequency		−	+ *Sex*	−
Inflectional Entropy	− *Correct*			+− *Sex*
Regularity				− *Previous RT*

Table 5: Overview of predictors by experiment. +: positive slope, −: negative slope, U: U-shaped effect; interactions with *Sex*, *Regularity*, and *Correctness* are indicated where present.

	Estimate	lower HPD95	upper HPD95	p(MCMC)
Intercept	6.8780	6.6843	7.0950	0.0000
Lemma Frequency (linear)	-0.0522	-0.0859	-0.0259	0.0001
Lemma Frequency (quadratic)	0.0031	0.0015	0.0055	0.0005
Picture Entropy	0.0787	0.0657	0.0992	0.0000
Sex=female	-0.0655	-0.1324	-0.0030	0.0476
Previous RT	0.0515	0.0325	0.0714	0.0000
Picture Complexity	0.0011	0.0003	0.0019	0.0075
Inflectional Entropy	0.0224	-0.0094	0.0509	0.1857
Correct=incorrect	0.2541	0.1346	0.3628	0.0000
Trial	0.0005	0.0004	0.0006	0.0000
Alternating=TRUE	0.4598	0.1863	0.7429	0.0010
Form Frequency	-0.0165	-0.0279	-0.0062	0.0024
Present=TRUE	-0.1654	-0.2377	-0.0923	0.0000
Regularity=regular	-0.1358	-0.2449	-0.0198	0.0215
Prepared Naming=TRUE	-0.0259	-0.0636	0.0141	0.2048
Rhyme Count	-0.0200	-0.0306	-0.0103	0.0000
Picture Entropy : Sex=female	0.0340	0.0183	0.0490	0.0000
Inflectional Entropy : Correct=incorrect	-0.0789	-0.1284	-0.0234	0.0052
Alternating=TRUE : Prepared Naming=TRUE	-0.0367	-0.0620	-0.0068	0.0138
Picture Complexity : Present=TRUE	0.0011	0.0006	0.0015	0.0000
Picture Complexity : Prepared=TRUE	-0.0007	-0.0011	-0.0002	0.0046
Present=TRUE : Form Frequency	0.0264	0.0100	0.0429	0.0018
Inflectional Entropy : Regularity=regular	0.0539	0.0020	0.1048	0.0438
Previous RT : Alternating	-0.0599	-0.0982	-0.0222	0.0016

Table 6: Coefficients of a linear mixed model fit to the picture naming latencies of Experiments 1–4 jointly. P-values are based on 50,000 Markov chain Monte Carlo samples from the posterior distribution of the parameters. The standard deviations estimated for the random effects were 0.134 for Verb and 0.095 for Subject. The standard deviation of the residual error was 0.257.