

Running Head: NEOLOGISMS AND LEXICAL MEMORY

Regular morphologically complex neologisms leave detectable traces in the mental lexicon

Laura de Vaan, Robert Schreuder and R. Harald Baayen

Radboud University Nijmegen

Abstract

This study investigates whether regular morphological complex neologisms leave detectable traces in the mental lexicon. Experiment 1 (subjective frequency estimation) was a validation study for our materials. It revealed that semantic ambiguity led to a greater reduction of the ratings for neologisms compared to existing words. Experiment 2 (visual lexical decision) and Experiment 3 (self-paced reading in connected discourse) made use of long-distance priming. In both experiments, the prime (base or neologism) was followed after 39 intervening trials by the neologism. As revealed by mixed-effect analyses of covariance, the target neologisms elicited shorter processing latencies in the identity priming condition compared to the condition in which the base word had been read previously, indicating an incipient facilitatory frequency effect for the neologism.

Keywords: morphology, neologisms, memory, mixed-effects models, long-distance priming

Introduction

Studies in first language acquisition have reported that one trial often suffices for children to form quick and rough hypotheses about the meaning of a word, a phenomenon referred to by Bloom (2000) as fast mapping. The learning of new words is not restricted to childhood. Words for new concepts, persons, place names, brand names continue to be encountered throughout one's life. Among these new words, one category is especially interesting from a morphological point of view, namely, derivational neologisms. Baayen and Renouf (1996) presented a corpus-based survey showing that such complex neologisms are not uncommon in written text. The present study addresses the question of how such neologisms are processed, and whether a single exposure to such a neologism already leaves a trace in lexical memory.

In models of the mental lexicon in which words are assumed to have their own representations in lexical memory (Butterworth, 1983), the question arises how frequently a word has to be encountered for such a representation to be acquired. In distributed approaches (Rumelhart and McClelland, 1986; Seidenberg and Gonnerman, 2000), the corresponding question is how many exposures are required for the connection weights to become sufficiently well-tuned to allow a processing advantage for the whole with respect to its parts to emerge. The phenomenon of fast mapping

suggests that initial representations may already be set up after a single exposure, or, alternatively, that the tuning of connection weights proceeds extremely rapidly. However, it is not clear whether fast mapping also applies to morphologically complex words that are to a very large extent predictable from their constituents.

Two approaches can be distinguished in previous work addressing the development and existence of lexical representations. One line of research is represented by the studies of Dumay, Gaskell and Feng (2004) and Gaskell and Dumay (2003). They studied the development of representations for pseudowords in lexical memory by tracing the effects of these new representations during lexical competition in the auditory processing of existing words. Their materials consisted of novel nonsense sequences that overlapped strongly with existed words, such as *cathedruke* versus *cathedral*. They presented their materials repeatedly to their participants, and observed a facilitating effect of having heard *cathedruke* for recognizing *cathedral*. However, after five days, they observed that the recognition of *cathedral* was slowed by prior familiarization with *cathedruke*. They attribute this inhibitory effect to *cathedruke* having become a lexical competitor of *cathedral*. Dumay et al. (2004) observed similar inhibitory effects already 24 hours after the initial familiarization with the pseudowords. Thus, one night's sleep seems sufficient for a lexical representation to develop. This result is certainly consistent with the fast mapping reported

in the acquisition literature.

Several scholars working on visual comprehension, by contrast, have argued that regular complex words would not leave any traces in lexical memory (Pinker, 1991), or only traces after a great many exposures (Alegre and Gordon, 1999; Pinker and Ullman, 2002a,b). Pinker and Ullman argue that memory traces for regular complex words are in fact superfluous given that regular words can be processed adequately by rules. They argue that such memory traces would be functional only in tasks that require sensitivity to the word's physical form. The study by Alegre and Gordon is especially interesting in that it proposes a threshold frequency of 6 per million below which rules would be effective for English inflection, and above which memory traces would be active. However, the work by Alegre and Gordon has been interpreted in very different ways. Pinker and Ullman (2002a) refer to this study as indicating that only some higher-frequency words are stored, but that the majority of English inflections are processed by rule. Eddington (2002), on the other hand, argues that this study shows that the majority of English inflections are processed through memory. The problem here is that a threshold of 6 per million has very different consequences depending on whether calculations are worked out on the basis of type counts or on the basis of token counts. Inspection of the frequencies available for English in the CELEX lexical database shows that 91.6% of the types would be

processed by rule, but that no less than 72.2% of the tokens would be processed by rote.

A threshold of 6 per million is questionable, however, for several reasons. A threshold below which representations do not exist meets with the logical problem that in order to install a representation once it is sufficiently frequent, the system must have kept track of how often the word has been encountered before. But this implies the existence of some kind of memory trace, the existence of which is denied by the threshold theory. Consequently, the threshold proposed by Alegre and Gordon (1999) is probably best understood as indicating how frequent a word must be for its developing representation to emerge from the measurement error in visual lexical decision. Third, studies in Dutch have reported frequency effects far below this threshold (Baayen, Dijkstra and Schreuder, 1997; Baayen, Schreuder, De Jong and Krott, 2002), and a study of English word formation shows frequency effects below the threshold for English as well (Wurm, Aycock and Baayen, submitted).

The present study aims to provide further insight into the development of traces for complex words in lexical memory during reading by investigating how neologisms are processed when encountered for the first and second time in experimental lists as well as in coherent discourse. The neologisms that we investigated all contained the suffix *-heid* (cf. '-ness' in English). The suffix *-heid* is the most regular and

productive derivational suffix of Dutch. In addition, it is a suffix that we have studied extensively in previous work (Baayen, Schreuder, Bertram and Tweedie, 1999; Bertram, Schreuder and Baayen, 2000).

With Experiment 1, we validated that our target words were indeed neologisms, using subjective familiarity rating. Experiments 2 and 3 addressed the processing of these neologisms. Experiment 2 made use of the visual lexical decision task, and Experiment 3 of self-paced reading. In both experiments, we implemented long-distance priming with 39 intervening words and two conditions. In one condition, henceforth the stem priming condition, the stem primed the neologism (*gammel - gammelheid* 'wobbly' - 'wobbliness'). In the other condition, henceforth the identity priming condition, the neologism primed itself (*gammelheid - gammelheid*). This design offers the possibility of comparing lexical processing of a neologism that has not been encountered before (the stem priming condition) with lexical processing of a neologism that has been encountered exactly once (the identity priming condition). If Alegre and Gordon (1999) and Pinker (1991) are right, we should not observe a difference between these two conditions, as the brain would not retain memory traces for regular complex words that have been encountered only once. However, if we observe that the processing of a neologism is facilitated by prior exposure to that neologism, we have evidence that prior experience leaves traces in at the very

least working memory.

Experiment 1: Familiarity Rating

Experiment 1 was designed as a validation study for the materials of Experiments 2 and 3.

Method

Participants. Twenty two students of the University of Nijmegen took part in this experiment. All students were enrolled in the Dutch Language and Culture program and were native speakers of Dutch.

Materials. The materials consisted of 42 neologisms with the suffix *-heid* ('-ness') and 63 existing words ending in *-heid* attested in the CELEX lexical database (Baayen, Piepenbrock and Gulikers, 1995), and five practice items. We selected adjectival stems for the neologisms for which suffixation with *-heid* resulted in semantically interpretable strings. We excluded intensifiers (*piepklein* 'teeny'), color adjectives (*blauw* 'blue') and adjectives ending in *-en* (*houten* 'wooden'), as such adjectives do not lend themselves well to suffixation with *-heid* (Bertram, Baayen and Schreuder, 2000) as well as adjectives from informal registers with negative connotations (*teut* 'tipsy'). Furthermore, we avoided neologisms that would have a high-frequency co-derivative ending in *-iteit*, as such words might exert a blocking force (Booij, 2002;

Rainer, 1988). We also avoided neologisms ending in *-igheid* whenever the resulting noun denoted an act, for example, *stom* 'stupid'; *stommig* 'somewhat stupid'; *stommigheid* 'stupid act' (Booij, 2002), as the combination *-igheid* is arguably a separate independent suffix with its own semantics.

None of the neologisms were attested in the CELEX database and none had a frequency greater than ten in the Dutch section of the world wide web (as of January 2003). The existing words had a CELEX lemma frequency in the range of [0–9623], a mean of 365.60 and a median of 17 (counts based on a corpus of 42 million words). All adjectival base words had a minimum length of three characters and a maximum length of eight characters.

We randomized the materials in six different orders, resulting in six lists. Each list was preceded by five practice trials and a short instruction with three examples. The experimental materials are listed in the Appendix.

Procedure. Participants were tested as a group during a class. We asked the participants to rate on a seven point scale how often they thought they had encountered these words.

Results and Discussion

One rating value was missing and was assigned the neutral score of 3.5. The mean

ratings for the neologisms and existing words were respectively 2.54 and 4.88. In addition to the factor Status (neologism vs. existing word), we included five covariates to the data set: Surface Frequency (the string frequency of the form presented), Base Frequency (the lemma frequency of the adjectival base), Length in letters, morphological Family Size (the number of different words in the morphological family (Schreuder and Baayen, 1997)), and a measure of the number of meanings of the word's base: its Number of Synonym sets in WordNet (henceforth Synsets) (Miller, 1990; Vossen, Bloksma and Boersma, 1999), following (Baayen, Feldman and Schreuder, 2006). This measure gauges the ambiguity in the semantic interpretation of the complex word.

Because of the skewness of the distribution for Surface Frequency, Base Frequency and Number of Synsets, we applied a logarithmic transformation to these variables. Since the Surface Frequency measure is zero for all neologisms, we analyzed the existing words and the neologisms separately. A mixed-effect model of covariance with subject and word as crossed random effects (Bates and Sarkar, 2005; Faraway, 2006) and the rating scores for the existing words as dependent variable revealed main effects of Surface Frequency, Family Size, Length and Number of Synsets. Ratings increased with increasing Surface Frequency ($\hat{\beta} = 0.431, t(1535) = 8.243, p < .0001$), Family Size ($\hat{\beta} = 0.190, t(1535) = 2.354, p = .0187$) and Length

($\hat{\beta} = 0.202, t(1535) = 3.338, p = .0009$), but decreased for increasing Numbers of Synsets ($\hat{\beta} = -0.466, t(1535) = -2.169, p < .0303$). For the neologisms, a stepwise mixed-effect regression analysis resulted in a model with significant main effects of Length, and Number of Synsets. As for the existing words, ratings increased with increasing Length ($\hat{\beta} = 0.238, t(877) = 2.805, p = .0051$) and decreased for increasing Numbers of Synsets ($\hat{\beta} = -0.560, t(877) = -3.359, p = .0008$).

Recall that random effects are random variables with zero mean and unknown standard deviation. In mixed-effects models, these unknown standard deviations are estimated from the data. For the present models, the standard deviations are for Word: 0.7971 (existing words) and 0.4737 (neologisms), for Subject: 0.8031 (existing words) and 0.7107 (neologisms), and for the residual error: 1.4805 (existing words) and 1.4081 (neologisms). These standard deviations provide direct insight into the amount of variance (squared standard deviations) that can be traced to the subjects, the words, and the residual error. It is common to find that the standard deviation for Word is less than that for Subject, especially in chronometric experiments, indicating that we tend to have much less experimental control over our subjects than over our experimental materials.

A greater number of meanings for the base words of the formations ending in *-heid* led to lower subjective frequency estimates, both for existing words and for

neologisms. The coefficient for the neologisms, -0.560, was more negative than the coefficient for the existing words, -0.466. The negative sign of these coefficient suggests that a greater semantic ambiguity of the base word of a formation ending in *-heid* may render the semantic interpretation of that formation more uncertain, which would lead to lower ratings. The greater magnitude of the effect for neologisms suggests that this problem of semantic interpretation is enhanced in the case of neologisms. In order to ascertain whether the coefficient for the neologisms is significantly more negative, we fitted a model to the combined data, without including Surface Frequency as a predictor. The significance ($F(1, 2415) = 4.07, p = .0438$) of the interaction of Number of Synsets by Type of Word (existing versus neologism) revealed that the effect of semantic ambiguity was indeed more detrimental for the ratings of neologisms. The greater effect of the Number of Synsets for neologisms, and the absence of a significant effect of the morphological family size for neologisms jointly support our intuitions that the words that we labelled as neologisms are indeed unknown to our participants.

Experiment 2 proceeds to investigate how neologisms are processed under time pressure, using visual lexical decision.

Experiment 2: Visual Lexical Decision

The key question addressed with Experiment 2 is whether neologisms leave traces

in lexical memory in isolated word reading. Since our neologisms are fully regular, decompositional theories denying storage to regular complex words, such as the theory of Pinker (1991), predict that no such traces should exist. To test this prediction, we made use of a design in which one group of subjects first saw the stem and a fixed number of trials later in the experiment also had to respond to the neologism containing that stem. A complementary group of subjects first saw the neologism (instead of its stem) and the same number of trials later in the experiment had to respond to the same neologism. If a neologism leaves a memory trace, subjects who had previously responded to the neologism should be faster in their responses to the second presentation of the neologism compared to the subjects who had previously encountered only its stem. In order to make sure that each subject had to parse neologisms the same number of times, each subject was presented an equal number of times with the order stem-neologism and the order neologism-neologism. In this way we ensured that across subjects the amount of prior experience with stem and affix was controlled for the critical comparison between the presence versus absence of prior experience with the neologism.

The task that we used for Experiment 2 was visual lexical decision, which required subjects to classify neologisms such as *lobbigheid* as words and not as nonwords. This is a natural categorization, that most participants carried out without requiring any

specific instruction about how to respond to neologisms. The neologisms that we used are possible words of Dutch that do not violate any phonological or morphological constraint. They were well-interpretable, as attested by their occasional attestation on the internet. Furthermore, previous studies have shown if participants are forced to classify possible words as nonwords, excessively long response latencies ensue: Native speakers find it difficult to reject words that are well-interpretable (Coolen, Van Jaarsveld and Schreuder, 1991).

Method

Participants. Twenty-six undergraduate students at the University of Nijmegen were paid to take part in this experiment. All were native speakers of Dutch, none had participated in Experiment 1.

Materials. We used the same neologisms as in Experiment 1, with the exception of *saploosheid* 'juicelessness' and *blitsheid* 'trendyness'. We excluded *saploosheid* because of its high degree of semantic similarity to another neologism, *saprijkheid* 'juicerichness'. We discarded *blitsheid* because we suspected that *blits* 'trendy, neat', a former vogue word, is currently not well known to our subject population. In this way, we obtained a list with 40 neologisms. We created 2 master lists with these 40 neologisms. In each master list, we replaced 20 of the neologisms by their stem so

that if one master list contained the stem, the other master list held the neologism for that stem. We then added the complete list of 40 neologisms to both master lists.

To distract attention from the neologisms, we added 40 existing words ending in *-ing* and 40 ending in *-baar* as filler words. The filler words ending in *-ing* had a CELEX lemma frequency in the range of [2–46], a mean of 7.9 and a median of 5; the filler words ending in *-baar* had a CELEX lemma frequency in the range of [0–83], a mean of 13.63 and a median of 6.5.

We matched each word with a pseudoword by changing one or two letters, making sure that the resulting word did not violate the phonotactic or orthographic conventions of Dutch. The changes were always in the stems and never in the suffix. Each master list contained 20 monomorphemic pseudowords (matching 20 stems of neologisms ending in *-heid*), 20 pseudowords ending in *-heid*, 40 ending in *-ing*, and 40 ending in *-baar*. We reused 20 pseudowords ending in *-ing* and 20 pseudowords ending in *-baar* in parallel to the stems and their neologisms ending in *-heid* that appeared twice. The total number of pseudowords ending in *-baar* and *-ing* was 60 for each suffix. As a result, a master list contained in all 20 stems, 20 neologisms, 40 neologisms (corresponding with the stems and neologisms mentioned before), 40 existing filler words ending in *-ing* and 40 existing filler words ending in *-baar*, to-

gether 160 word trials, matched with 160 pseudoword trials. We pseudo-randomized both master lists in three different orders. We made sure that the number of trials between a stem and its corresponding neologism, between a neologism and its repetition, and also between a pseudoword and its repetition was held constant at 39 intervening trials. Each of the resulting six lists was preceded by 24 practice trials. The experimental materials are listed in the Appendix.

Procedure. Participants were tested individually in a noise-attenuated experimental room. They were asked to decide as quickly and accurately as possible whether the letter string appearing on the computer screen was a possible Dutch word. After the practise session, a few subjects were unclear about how to respond to the neologisms. We made it clear to them that both existing and possible words required a yes response, and that only impossible words (phonotactically legal, but meaningless) required a no response. The experiment was run without any further breaks. An experimental trial began with a fixation mark, positioned in the center of the screen for 500 ms. 50 ms later, the stimulus appeared centered at the same position. Stimuli were presented on a Nec Multisync color monitor in white lowercase, 36 point letters on a dark background and they remained on the screen for 1750 ms.

Results and Discussion

We coded no-reponses to neologisms as errors and removed these data points from the analyses of the response latencies. The by-item error scores ranged from 4% to 46%. As we are dealing with neologisms in lexical decision, these high error rates are unsurprising. We applied a logarithmic transformation to the response latencies for the prime and those for the target to reduce the non-normality of their distribution in order to avoid distortion of the statistical model by outliers with undue leverage.

The group means response latencies (RTs) across all subjects and items, 737 ms for the condition in which the base preceded the target, and 769 ms for the condition in which the neologism preceded the target, suggested a significant inhibitory priming effect, instead of the expected facilitatory priming effect. A linear mixed-effects model with subject and word as crossed random effects supported this unexpected priming effect as significant ($\hat{\beta} = 0.0301, t(830) = 2.08, p = .0376$). However, a closer inspection of the data using appropriate controls shows that this initial analysis is highly misleading.

In our full analysis we added Base Frequency, Family Size, Length (in letters), Number of Synonym sets in Wordnet and the mean number of items in a synset as covariates. As in Experiment 1, we logarithmically transformed Base Frequency, Family Size and Number of Synsets, in order to reduce the skewness in their distributions. Crucially, we also added several non-lexical, experimental control variables.

Since we expected the processing load of the prime to affect the latencies for the target, we also included the response latency of a given participant to the prime as covariate for that participant's response to the target 40 trials later in the experiment (henceforth Prime RT), a possibility afforded by mixed-effect models. Similarly, we included the lexical decision by a given participant for a given prime (word or nonword, henceforth Prime Decision) as a factor. We also registered the response latencies of a given participant to the four preceding words as potential covariates for that participant's response to the target. As these four response latencies were all mutually correlated, we used principal components analysis to orthogonalize these covariates, to which we first applied a logarithmic transformation. Of the resulting principal components, only the first turned out to be predictive for the response latencies to our targets. In what follows, we refer to this principal component as PC Preceding RTs.

We fitted a mixed-effect model of covariance to the RTs to the target, using a stepwise variable selection procedure, with subject and word as crossed random effects. We observed main effects of Base Frequency, Condition, Prime RT, Prime Decision and PC-Preceding RTs as well as an interaction of Prime RT by Prime Decision (see Figure 1). No other predictors and interactions reached significance. Inspection of the residuals revealed marked non-normality, indicating serious lack

of goodness-of-fit for this model. We therefore removed outliers with standardized residuals outside the interval $[-2.5, 2.5]$ (2.16%), and refitted the model (see, e.g., Crawley, 2002). The residuals of this trimmed model were approximately normally distributed, indicating that removal of overly influential outliers resulted in a model with superior goodness of fit.

INSERT FIGURE 1 HERE

Response latencies decreased with increasing Base Frequency, as expected for neologisms ($\hat{\beta} = -0.0104, t(807) = -2.66, p = .0080$). Response latencies were now shorter in the identity priming condition (in which participants responded to the neologism for the second time) ($\hat{\beta} = -0.0466, t(807) = -3.29, p = .0011$) than in the stem priming condition (in which the neologism was entirely new). There was no significant interaction of Base Frequency by Condition.

Furthermore, response latencies were longer for subjects who had previously rejected the prime as a word ($\hat{\beta} = 1.3837, t(807) = 3.90, p = .0001$). Response latencies also increased with increasing Prime RTs ($\hat{\beta} = 0.2353, t(807) = 7.39, p < .0001$). Prime Decision interacted with Prime RT ($\hat{\beta} = -0.1888, t(807) = -3.56, p = .0004$): For those trials for which a subject had previously rejected the prime as a word, the reaction time to the prime was not predictive for the reaction time to the target.

RTs decreased with increasing PC Preceding RTs ($\hat{\beta} = -0.0492, t(807) = -7.22, p < .0001$). As the PC Preceding RTs were themselves negatively correlated with the preceding RTs, we conclude that longer preceding RTs implied longer RTs to our target words. This finding is in agreement with the results reported by Taylor and Lupker (2001) for word naming. Unlike in their study, however, which reported small effects of around 10 milliseconds, the longitudinal effect of the processing complexity of the preceding trials emerges as a much more substantial effect in our mixed-effect analysis: The difference in the RTs, calibrated for the medians of the other predictors, between the smallest and largest values of PC Preceding RTs was of the order of magnitude of 300 ms. Hence, it is a control variable that is worth taking into account.

We complete the specification of the mixed-effect model with reporting the standard deviations. The standard deviation for the random effect of Word was estimated at 0.0297. In this model, there were three random effects involving Subject. The standard deviation for the by-subject adjustments to the intercept was 0.1164. In addition, participants turned out to be differentially sensitive to the Family Size count (log-likelihood ratio = 12.36, $p = 0.0021$). The standard deviation for the by-subject adjustments to coefficient of Family Size was 0.0505, and the estimated coefficient for the correlation of the by-subject adjustments to intercept and Family

Size was -0.156. The residual standard deviation was 0.1777.

The reversal of the priming effect, from inhibitory in a model without controls, to facilitatory in the model with appropriate controls, is brought about not by the inclusion of the lexical covariate (Base Frequency), but by the experimental control variables. Subanalyses show that each of the control variables contributes independently to this reversal. For instance, if a subject rejected a neologism as a word when the neologism was presented as the prime, and later reversed this decision for the target, this revision led to a longer response latency for the target, contributing to the greater raw group mean for the priming condition in which the neologism was repeated. In other words, the facilitatory priming effect emerges more clearly for prime-target trials where subjects accepted the neologisms at both occasions. The full model therefore provides much better insight into how in our experiment task requirements, the details of the transfer of prior processing onto later processing, and between-trial correlational structure may not only mask but completely reverse a priming effect as visible in raw group means.

We also fitted a mixed-effect model of covariance to the Prime RTs , using a step-wise variable selection procedure, with subject and word as crossed random effects. We observed main effects of PC-Preceding RTs, Base Frequency and Condition as well as an interaction of Condition by Family Size.

Response latencies to the prime decreased with increasing Base Frequency, as expected for neologisms ($\hat{\beta} = -0.0196, t(1034) = -4.19, p < .0001$). (We note that the coefficient for Base Frequency for the prime is almost twice that observed for the target ($\hat{\beta} = -0.0104$). This reduction in the Base Frequency effect may be due either to attenuation at the second exposure, to the absence of bare stems among the targets, or to both.) There was a significant main effect of Condition: Response latencies to the prime were longer in the identity priming condition (in which participants responded to a neologism) than in the stem priming condition (in which participants responded to an adjectival base) ($\hat{\beta} = 0.1500, t(1034) = 4.27, p < .0001$). There was no significant interaction of Base Frequency by Condition. Response latencies to the prime increased with increasing PC Preceding RTs ($\hat{\beta} = 0.0423, t(1034) = 5.88, p < .0001$). Inspection of the correlational structure of this control variable shows that its interpretation remains identical to that for the target words.

The effect of Condition was modulated by Family Size ($F(1, 1034) = 6.93, p = 0.0086$): For subjects who had to respond to an adjectival base, response latencies to the prime decreased with increasing Family Size ($\hat{\beta} = -0.0300, t(1034) = -2.98, p = .0030$), while Family Size was not significant in the identity priming condition ($\hat{\beta} = 0.0071, t(1034) = 0.70, p = .4853$). Recall that in Experiment 1, a larger number of

meanings (synsets) led to significantly lower ratings for neologisms than for existing words. This finding is mirrored in part by the present effect of Family Size, the count of words that are not only semantically but also formally related. For words that are not neologisms, a higher synset or a higher family count leads to higher ratings and shorter latencies, respectively. For neologisms, we find either lower ratings or no facilitation (non-significant inhibition).

We complete the specification of the mixed-effect model with reporting the standard deviations. The standard deviation for the Word random effect was 0.0575, and that for the Subject random effect was 0.0946. In addition, there was a significant random effect of Subject by Family Size (log-likelihood ratio = 14.83, $p = .0006$), indicating that the Family Size count was predictive, but correlated negatively with the RTs for some subjects, and positively with the RTs for others. The standard deviation for the by-subject adjustments to coefficient of Family Size was 0.0775, and the estimated coefficient for the correlation of the by-subject adjustments to intercept and Family Size was 0.353. The standard deviation of the residual error was 0.2285.

The analysis of the response latencies to the *prime* showed no effect of Family Size for the neologisms. The analysis of the response latencies to the *target* showed that subjects differ in their sensitivity to Family Size. For some subjects

the response latencies *decreased* with increasing Family Size, for others the response latencies *increased* with increasing Family Size. From the study of Bertram, Baayen and Schreuder (2000), we know that for well-established formations ending in *-heid* response latencies in visual lexical decision decrease with increasing Family Size. Considered jointly, these results suggest that the family size effect in lexical decision is absent for neologisms, begins to emerge already after only 1 exposure (either facilitatory or inhibitory), and ultimately becomes facilitatory.

The presence of a family size effect in the latencies for the prime that was a base word and its absence for the neologism allow us to infer that the priming effect is not driven simply by greater attentional resources being drawn to the base in the identity priming condition. If such were the case, then we would have expected a family size effect primarily for the condition in which the neologism is the prime, contrary to fact.

In summary, Experiment 2 shows that the processing of a neologism is affected by the kind of prime encountered previously. Participants *recognized* a neologism at the second exposure, allowing them to respond more quickly. The lexical decision made 40 trials previously (roughly two minutes) was also reflected in the latencies to the targets, with earlier rejection of the neologism as a word leading to prolonged RTs. In both priming conditions, the response latencies to the prime were predictive

for the processing of the target. We conclude that the primes left traces in working memory, independently of the type of prime (stem vs. neologism).

Experiment 2 made use of long-distance priming for isolated words presented in a list. The assessment in Pinker (1999) of frequency effects in lexical decision is that it is the lexical decision task that leads subjects to make use of frequency information that they would not use in reading or listening to words in normal discourse. We put this claim to the test by embedding the target words of Experiment 2 in natural discourse, and by switching from lexical decision to self-paced reading, a task with a higher degree of ecological validity. We continued to make use of long-distance priming, as both priming conditions occur in natural speech. For instance, the sequence of an adjectival stem followed by the corresponding noun further on in the discourse is not uncommon, as witnessed by example (1) (Kastovsky (1986), see also Baayen and Neijt (1997)).

(1) "... and whether our own conversation doesn't sound a little *potty*.

It's the *pottiness*, you know, that's so awful."

Similarly, the sequence of a noun ending in *-ness*, followed later in the discourse by the same noun occurs, is a common phenomenon, as witnessed by example (2).

(2) "But prime ministerial *pottiness* is ever with us. Who, doe-eyed and

dripping sincerity, invited our trust - and made himself the ultimate repository of that virtue? Tony Blair. So who now lives or dies on the red line he created? Pause, though, for *pottiness* is a three-syllable word of some complexity.”

(Preston, P. (June 21, 2004). Trust is an exaggerate commodity. *The Guardian*).

Experiment 3 therefore addressed the question of whether similar differential effects of priming arise when subjects read the same stimuli, but now presented in cohesive discourse without having to make lexical decisions, using the self-paced reading paradigm.

Experiment 3: Self-Paced Reading

Method

Participants. Thirty-two undergraduate students at the University of Nijmegen were paid to take part in this experiment. All were native speakers of Dutch. None had taken part in the preceding experiments.

Materials. We used the same neologisms as in Experiment 2. We created 40 short texts with a median length of 65 words in which we used these neologisms. For each

text, we created two variants that differed with respect to whether the prime was the base adjective or the neologism itself. Exactly 40 words later in the text, the prime was followed by the target neologism. Each text was followed by a question about its contents, in order to ensure that participants would read the texts carefully for content.

The Appendix presents an example of the parallel texts for the neologism *tembaarheid* ('tamability'), respectively the condition with stem priming and the condition with identity priming, primes and targets are shown in bold, the neologisms are also underlined. Special care was required for the construction of the parallel texts. First, since prime and target belonged to different word categories (adjective vs. noun), it was impossible to simply replace the adjectival prime by the target without effecting minimal changes in the syntax. Second, adjectives in Dutch may be inflected for person and number, in which case they are suffixed with a schwa. The texts were constructed in such a way that the number of inflected adjectival primes was kept as small as possible (6 out of 40 adjectival primes were inflected). Third, the texts were presented over several lines, such that primes and targets did not appear on the first or last position of a line.

We created two master lists. Each master list contained 20 texts with the stem as prime and 20 texts with the neologism as prime, such that a given participant read

a given prime-target pair only once. The order in which the texts were presented was randomized four times. For each random order, we used two lists that differed only in whether the base or the neologism was presented as prime. Each one of the resulting 8 lists was preceded by three practice trials; a fourth practice trial followed a short pause after the first 20 critical trials.

Procedure. Participants were tested individually in a noise-attenuated experimental room in front of a monitor and a panel with three buttons. They received standard self-paced reading instructions and were asked to read carefully and at a comfortable reading speed. The course of a trial was as follows. The participants saw a fixation point, indicating where the text would begin. By pushing the middle button they initiated the trial and the full text with dashes marking letter positions was displayed. By pushing the middle button again they saw the first word of the text. When the participants pressed the middle button again, the second word appeared and the first word was replaced by dashes. This procedure was repeated until the participants had read the whole story, at which point the word VRAAG ('question') was presented in the center of the screen, followed by a yes/no question about the text. We asked the participants to indicate their response by means of a button press (the right button for yes, the left button for no).

Results and Discussion

Inspection of the group means for the two priming conditions, averaging over subjects and words, revealed a small inhibitory effect that, unlike for Experiment 2, failed to reach significance (stem priming condition: 431 ms, identity priming condition 437 ms; $p > 0.8$).

As for Experiment 2, we added covariates to our model in order to obtain improved insight into the structure of our data. We included Base Frequency, Family Size, Length (in letters), Number of Synonym sets in Wordnet and the mean number of items in a synset as lexical covariates. Because Experiment 2 revealed an effect of the prime on the latencies of the target, we included as a first experimental control Prime RT, the reading time of a given participant for the prime, as a covariate for the participant's response to the target 40 trials later in the experiment. We also included the (orthogonalized) reading times of a given participant to the four preceding words as covariates for that participant's response to the target (henceforth, Preceding RTs) to control for between-text differences in the preceding words and syntax, and to control for reading speed.

As in Experiment 2, we logarithmically transformed the Base Frequency, Family Size, Number of Synset measures, the reading times for the prime and to the RTs for the target, as well as to the Preceding RTs to reduce the non-normality of their

distribution and to avoid distortion of the statistical model by outliers with undue leverage.

We fitted a mixed-effect model of covariance to the reading times for the targets, using a stepwise variable selection procedure, with subject and word as crossed random effects. We observed main effects of Prime RT, PC-Preceding RTs and Length as well as an interaction of Condition by PC-Preceding RTs (see Figure 2). Inspection of the residuals revealed marked non-normality, indicating a lack of goodness of fit. We therefore removed outliers with standardized residuals outside the interval $[-2.5, 2.5]$ (2.21%), and refitted the model. The residuals of this model were more normally distributed, indicating improved goodness of fit for the trimmed model.

INSERT FIGURE 2 HERE

Reading times increased with increasing Prime RT, as expected given the results of Experiment 2 ($\hat{\beta} = 0.1014, t(1233) = 5.10, p < .0001$): If the prime required longer reading time, than it is not surprising that the target, which partially or completely repeats the prime, also required longer reading time. Reading times also increased with increasing PC-Preceding RTs (the first principal component of the orthogonalized four preceding reading times, $\hat{\beta} = 0.1504, t(1233) = 16.29, p < .0001$). Inspection of this control variable shows that the

PC-Preceding RTs were positively correlated with the preceding RTs. Longer preceding RTs implied longer RTs to our target words, indicating that greater lexical and syntactic complexity of the immediately preceding context rendered the reading of the neologism more difficult. Longer words elicited longer reading times as well, as expected ($\hat{\beta} = 0.0375, t(1233) = 3.44, p = .0006$).

The effect of prime type manifested itself in the form of an interaction with PC-Preceding RTs ($\hat{\beta} = -0.0251, t(1233) = -2.24, p = .0251$): Reading times were shorter in the identity priming condition (in which participants responded to the neologism for the second time) than in the stem priming condition (in which the neologism was entirely new), proportional to the processing costs for the four preceding words as measured by means of PC-Preceding RTs. Previous experience with the neologism facilitated subsequent processing, exactly as expected for a budding frequency effect. Other predictors such as Base Frequency and Family Size were not significant. The estimates of the standard deviations of the random effects in this model were, for Word: 0.0580, for Subject: 0.1029, and for the residual error: 0.2882.

We also fitted a mixed-effect model of covariance to the Prime RTs, using a stepwise variable selection procedure, with subject and word as crossed random effects. In this case, we incorporated all four principal components. Only the first (57% of

the variance) and second (16% of the variance) principal component (respectively PC1-Preceding RT and PC2-Preceding RT) reached significance in the analysis. We also observed main effects for Condition and Family Size as well as an interaction of Subject by Condition.

Response latencies to the prime were longer in the identity priming condition (in which participants responded to a neologism) than in the stem priming condition (in which participants responded to an adjectival base) ($\hat{\beta} = 0.1538, t(1241) = 3.14, p = .0017$). Response latencies to the prime were longer for words with larger morphological families ($\hat{\beta} = 0.0355, t(1241) = 2.78, p = .0055$). Response latencies also increased with increasing PC1-Preceding RTs ($\hat{\beta} = 0.1602, t(1241) = 20.67, p < .0001$), and with increasing PC2-Preceding RTs ($\hat{\beta} = 0.0372, t(1241) = 2.83, p = .0047$). Inspection of the correlational structure of these control variables shows that their interpretation remains identical to that for the target words: greater RTs to preceding trials correlated positively with the RTs for the prime.

We complete the specification of the mixed-effect model with reporting the standard deviations. The standard deviation for the random effect of Word was estimated at 0.1590. In this model, there were three random effects involving Subject. The standard deviation for the by-subject adjustments to the intercept was 0.0709. In addition, participants turned out to be differentially sensitive to Condition (log-

likelihood ratio = 34.43, $p < .0001$). The standard deviation for the by-subject adjustments to coefficient of Condition was 0.1609, and the estimated coefficient for the correlation of the by-subject adjustments to intercept and Condition was 0.913. The residual standard deviation was 0.3177.

A comparison of the reading times for the primes and those for the targets reveals an interesting difference concerning the predictivity of the Family Size measure. For the primes, the Family Size count was inhibitory for both base words and neologisms, for the targets, the Family Size count was not significant. Recall that in visual lexical decision, Family Size is facilitatory with existing words, possibly because a lexical decision can be based on a global evaluation of the amount of lexical activation, for example, Grainger and Jacobs (1996) and Schreuder and Baayen (1997). In self-paced reading, new information has to be integrated in the current discourse representation. Apparently, this integration process is slowed down when a word has a large morphological family. Such words are more ambiguous, they require more processing resources in order to select their appropriate meaning given the preceding discourse. Interestingly, the absence of a Family Size effect for the targets suggests that this process of disambiguation and discourse integration takes place once only. When an integrated word (or its derivative ending in *-heid*) is encountered a second time, its interpretation given the discourse is retrieved (as witnessed by

the predictivity of the reading time for the prime for the reading time of the target), and its family members are no longer considered.

Compared to Experiment 2, Experiment 3 added not only a semantic context, but also a syntactic context. The syntactic context of the prime (adjective) differs slightly from that of the target (noun). In other words, the stem priming condition and the identity priming condition differed not only with respect to the presence of *-heid*, but also with respect to the local syntax. This difference in syntactic context may have attenuated the effect of Condition, but given the positive results of Experiment 2, where no difference in syntactic issue was at issue, it is clear that the effect of syntactic context cannot account for the effect of Condition all by itself.

General Discussion

The question addressed in this paper is whether neologisms leave detectable traces in lexical memory. According to Alegre and Gordon (1999) and Pinker (1991), regular morphological complex words do not have their own lexical representation. Experiment 1, a subjective familiarity rating, showed that neologisms ending in *-heid* ('-ness') were processed qualitatively differently from existing words with respect to the effect of their number of meanings, as gauged by the count of their synonym sets in WordNet. For both existing words and neologisms, a larger number of synonyms led to lower subjective frequency estimates. But for neologisms, this detrimental

effect on the ratings of semantic ambiguity was stronger. This suggests that the interpretation of a neologism becomes more difficult when the base word is more ambiguous.

Experiment 2 (visual lexical decision) and Experiment 3 (self-paced reading) showed that a neologism is processed in a qualitatively different way after even a single exposure. Both experiments made use of long-distance priming. The prime, which preceded the target neologism by 40 trials, was either the neologism's base, or the neologism itself. In visual lexical decision, neologisms elicited shorter response latencies in the identity priming condition compared to the stem priming condition. We conclude that participants recognized a neologism at the second exposure, allowing them to respond more quickly.

In self-paced reading, the difference between the two priming conditions emerged in interaction with the processing complexity of the preceding words in the discourse. We gauged the complexity of the preceding discourse by means of a measure (obtained through principal components orthogonalization) that captured a large proportion of the variance in the reading times of the four words preceding the target. The coefficient of this measure (the first principal component) was significantly more facilitatory for neologisms that had been read before, than for neologisms of which only the base had been seen previously. Given that there were

39 intervening trials in our experiments, and that the span of short-term memory for words is much shorter than this, limited to 5–9 words, these results allow us to conclude that neologisms leave traces in lexical memory.

The question that arises at this point is what the nature is of these memory traces. Do neologisms receive a lexical entry in the mental lexicon after just a single exposure? Is this entry a whole-word representation similar to the entries of monomorphemic words? Or is this entry a combinatorial trace that, given base and suffix, provides a direct link to their joint interpretation as computed when the neologism was first encountered? Is the initial memory trace qualitatively different for inflected words, derived words, and contiguous words in the syntax? At this stage of our research, we must leave questions such as these unanswered. We think that the present memory effect is not restricted to derived words, but extends both to inflected words and to (novel) collocations. And it may well be that the memory traces in our data are inherently combinatorial, as argued by Wurm et al. (submitted) for low-frequency complex words in general.

Our present results also remain silent about further consolidation in memory. Recall that Dumay et al. (2004) and Gaskell and Dumay (2003) observed that pseudowords that overlapped strongly with existed words became lexical competitors of the existing words already 24 hours after the initial familiarization (12 exposures)

with the pseudowords. Gaskell and Dumay (in press) show that the integration of a pseudoword in lexical memory crucially requires at least one night's sleep, and not just by the passing of twelve hours between training and testing.

Our experiments differ in several respects from those of Dumay et al. (2004) and Gaskell and Dumay (2003, in press). We used long-distance priming within one and the same experiment. This suggests that the consolidation in long-term memory for which a night's sleep is required did not take place. A second difference is that we observed effects of lexical learning after only one exposure, whereas Gaskell and Dumay (2003) presented their materials repeatedly to their participants. A third difference is that we studied lexical learning *directly* by investigating novel, but morphologically well-formed and semantically transparent derived words. Gaskell and colleagues, by contrast, studied lexical learning *indirectly*, by tracing the consequences of acquired nonwords on the processing of other well-established, existing words.

Given the importance of sleep for lexical consolidation, it is unlikely that full-fledged lexical entries for morphologically complex neologisms are established in long-term memory immediately after a single exposure. Research is currently in progress that addresses the importance of sleep for the consolidation of neologisms in lexical memory. What our experiments do allow us to conclude is that neologisms

do leave traces beyond immediate working memory, and that these traces facilitate subsequent processing. Hence we expect to find that even such ephemeral traces are in fact consolidated after a night's sleep.

A crucial result that supports the importance of lexical traces for language processing is that the long-distance priming effect was present not only in visual lexical decision, but also in natural discourse. Although Alegre and Gordon (1999) and Pinker (1991, 1999) have argued that especially low-frequency words would not have their own lexical representations, our results suggest that memory traces for even the lowest-frequency complex words play a functional role in discourse processing. Hence, it is unlikely that Pinker's dismissal of experimental evidence for frequency effects in visual lexical processing for regular complex words as task artifacts is justified. Our data bear witness to the relevance of previous experience with complex words also in connected discourse. In fact, we predict that the discourse context in which a neologism is first encountered leaves a trace in episodic memory, and that lexical processing of the neologism after one night's sleep will be facilitated notably when it appears in a similar context.

Our study also offers some methodological innovations, enabled by mixed-effect modeling with subject and item as crossed random effects (Faraway, 2006). First, we were able to bring at least part of the correlational structure of the latencies

for our target words and the latencies for the items preceding our target words in the experiment into our statistical models, by means of orthogonalized measures for the latencies to the preceding trials. This provided us with the necessary power to observe an effect of priming in self-paced reading, and protected us against the erroneous and theoretically unexpected conclusion suggested by the raw group means that prior experience would lead to inhibition rather than to facilitation.

Second, mixed-effect modeling also allowed us to observe that the latency to the prime, and the lexical decision to the prime, were predictive for how the target was processed 40 trials later in the experimental list. To our knowledge, priming studies have thus far only considered the effect of priming at the factor level. Mixed-effect analysis of covariance with subject and item as crossed random effects allowed us to zoom in on the effect of priming at the level of the responses of specific subjects to specific items.

References

- Alegre, M., & Gordon, P. (1999). Frequency effects and the representational status of regular inflections. *Journal of Memory and Language*, *40*, 41–61.
- Baayen, R. H., Dijkstra, T., & Schreuder, R. (1997). Singulars and plurals in Dutch: Evidence for a parallel dual route model. *Journal of Memory and Language*, *37*, 94–117.
- Baayen, R. H., Feldman, L. B., & Schreuder, R. (2006). Morphological influences on the recognition of monosyllabic monomorphemic words. *Journal of Memory and Language*, *55*, 290–313.
- Baayen, R. H., & Neijt, A. (1997). Productivity in context: a case study of a Dutch suffix. *Linguistics*, *35*, 565–587.
- Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database (CD-ROM)*. Linguistic Data Consortium, University of Pennsylvania, Philadelphia, PA.
- Baayen, R. H., & Renouf, A. (1996). Chronicling The Times: Productive Lexical Innovations in an English Newspaper. *Language*, *72*, 69–96.
- Baayen, R. H., Schreuder, R., Bertram, R., & Tweedie, F. J. (1999). The semantic

- functions of the Dutch suffix *-heid*: evidence from lexicography, lexical statistics, and psycholinguistics. In: Nenonen, M., & Jarvikivi, J. (Eds.), *Languages, minds, and brains* (pp. 1–29). University of Joensuu, Joensuu.
- Baayen, R. H., Schreuder, R., De Jong, N. H., & Krott, A. (2002). Dutch inflection: the rules that prove the exception. In: Nootboom, S., Weerman, F., & Wijnen, F. (Eds.), *Storage and Computation in the Language Faculty* (pp. 61–92). Kluwer Academic Publishers, Dordrecht.
- Bates, D. M., & Sarkar, D. (2005). The lme4 library. [*On-line*], Available: <http://lib.stat.cmu.edu/R/CRAN/>.
- Bertram, R., Baayen, R. H., & Schreuder, R. (2000). Effects of family size for complex words. *Journal of Memory and Language*, *42*, 390–405.
- Bertram, R., Schreuder, R., & Baayen, R. H. (2000). The balance of storage and computation in morphological processing: the role of word formation type, affixal homonymy, and productivity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 489–511.
- Bloom, P. (2000). *How Children Learn the Meanings of Words*. MIT Press, Cambridge, MA.

- Booij, G. (2002). *The Morphology of Dutch*. Oxford University Press, Oxford.
- Butterworth, B. (1983). Lexical representation. In: Butterworth, B. (Ed.), *Language production (Vol. II): Development, Writing and Other language Processes* (pp. 257–294). Academic Press, London.
- Coolen, R., Van Jaarsveld, H. J., & Schreuder, R. (1991). The interpretation of isolated novel nominal compounds. *Memory and Cognition*, 19, 341–352.
- Crawley, M. J. (2002). *Statistical computing. An introduction to data analysis using S-plus*. Wiley, Chichester.
- Dumay, N., Gaskell, M. G., & Feng, X. (2004). A day in the life of a spoken word. In: Forbus, K., Gentner, D., & Regier, T. (Eds.), *Proceedings of the twenty-sixth annual conference of the cognitive science society* (pp. 339–344). Erlbaum, Mahwah, NJ.
- Eddington, D. (2002). Spanish diminutive formation without rules or constraints. *Linguistics*, 40, 395–419.
- Faraway, J. J. 2006. *Extending Linear Models with R: Generalized Linear, Mixed Effects and Nonparametric Regression Models*. Chapman & Hall/CRC, Boca Raton, FL.

- Gaskell, M. G., & Dumay, N. (2003). Lexical competition and the acquisition of novel words. *Cognition*, *89*, 105–132.
- Gaskell, M. G., & Dumay, N. (in press). Sleep-associated changes in the mental representation of spoken words. *Psychological Science*.
- Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review*, *103*, 518–565.
- Kastovsky, D. (1986). Productivity in word formation. *Linguistics*, *24*, 585–600.
- Miller, G. A. (1990). Wordnet: An on-line lexical database. *International Journal of Lexicography*, *3*, 235–312.
- Pinker, S. (1991). Rules of language. *Science*, *153*, 530–535.
- Pinker, S. (1999). *Words and Rules: The Ingredients of Language*. Weidenfeld and Nicolson, London.
- Pinker, S., & Ullman, M. (2002a). Combination and structure, not gradedness, is the issue: Reply to McClelland and Patterson. *Trends in the Cognitive Sciences*, *6* (11), 472–474.
- Pinker, S., & Ullman, M. (2002b). The past and future of the past tense. *Trends in the Cognitive Sciences*, *6* (11), 456–462.

- Rainer, F. (1988). Towards a theory of blocking: the case of Italian and German quality nouns. In: Booij, G. E., & Van Marle, J. (Eds.), *Yearbook of Morphology I* (pp. 155–185). Foris, Dordrecht.
- Rumelhart, D. E., & McClelland, J. L. (1986). On learning the past tenses of English verbs. In: McClelland, J. L., & Rumelhart, D. E. (Eds.), *Parallel Distributed Processing. Explorations in the Microstructure of Cognition. Vol. 2: Psychological and Biological Models* (pp. 216–271). The MIT Press, Cambridge, Mass.
- Schreuder, R., & Baayen, R. H. (1997). How complex simplex words can be. *Journal of Memory and Language*, 37, 118–139.
- Seidenberg, M. S., & Gonnerman, L. M. (2000). Explaining derivational morphology as the convergence of codes. *Trends in Cognitive Sciences*, 4 (9), 353–361.
- Taylor, T. E., & Lupker, S. J. (2001). Sequential effects in naming: A time-criterion account. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 27, 117–138.
- Vossen, P., Bloksma, L., & Boersma, P. (1999). *The Dutch WordNet (CD-ROM)*. European Language Resources Association (ELRA), Luxembourg.
- Wurm, L. H., Aycocock, J., & Baayen, R. H. (submitted). *Lexical dynamics for*

low-frequency complex words: a regression study across tasks and modalities.

Manuscript submitted for publication.

Appendix

Experimental materials of Experiments 1 and 2:

kortafheid; lobbigheid; labielheid; pitloosheid; saprijkheid; saploosheid*; summierheid; tactvolheid; tembaarheid; tilbaarheid; visrijkheid; pipsheid; blitsheid*; enormheid; jofelheid; koketheid; onwelheid; ovaalheid; riantheid; antiekheid; banaalheid; basaalheid; dementheid; erkendheid; gammelheid; ludiekheid; ondiepheid; royaalheid; aftandsheid; beschuttheid; bezweetheid; contentheid; coulantheid; bebrildheid; markantheid; beuheid; blusbaarheid; geurloosheid; kalkrijkheid; onattentheid; onbelastheid; ontroerdheid

Example of experimental materials of Experiment 3:

Stem - Neologism:

Als je vraagt naar een dier dat **tembaar** is, dan denkt bijna iedereen direct aan het paard. Paarden worden al sinds mensenheugenis ingezet als lastdieren en als werkdieren. Voordat ze ingezet kunnen worden, moeten de dieren zadelmak gemaakt worden. Hiervoor heb je veel geduld nodig. Ook de tembaarheid van een dier speelt mee. Bij vurige dieren duurt het nu eenmaal langer voor je ze zadelmak hebt, en je ze dus in kunt zetten als lastdier of als werkdier, dan bij rustigere dieren.

If you ask for a **tamable** animal, everyone will immediately think of a horse. Horses have been used as pack animal and as working animal since time immemorial. Before they can be used, the horses have to be trained. You need to be patient for this job. The **tamability** of an animal plays also a role. It will take more time to train high-spirited horses than to train calm animals. Hence, it will take more time before you can use a high-spirited horse as pack animal or as working animal.

Neologism - neologism:

Als je vraagt naar de **tembaarheid** van dieren, dan denkt bijna iedereen direct aan paarden. Paarden worden al sinds menscheugenis ingezet als lastdieren en als werkdieren. Voordat ze ingezet kunnen worden, moeten de dieren zadelmak gemaakt worden. Hiervoor heb je veel geduld nodig. Ook de **tembaarheid** van een dier speelt mee. Bij vurige dieren duurt het nu eenmaal langer voor je ze zadelmak hebt, en je ze dus in kunt zetten als lastdier of als werkdier, dan bij rustigere dieren.

If you ask for the **tamability** of animals, everyone will immediately think of horses. Horses have been used as pack animal and as working animal since time immemorial. Before they can be used, the horses have to be trained. You need to be patient for this job. The **tamability** of an animal plays also a role. It will take more time to train high-spirited

horses than to train calm animals. Hence, it will take more time before you can use a high-spirited horse as pack animal or as working animal.

* Excluded in Experiments 2 and 3

Author's address:

Laura de Vaan, Interfaculty Research Unit for Language and Speech Behaviour,
P.O. Box 310, 6500 AH, Nijmegen, The Netherlands. E-mail:laudev@mpi.nl

Figure 1. Partial effects for response latencies in visual lexical decision. In the upper left panel, the dotted line represents prime-target pairs for which the prime was incorrectly classified as a pseudoword, and the solid line the prime-target pairs for which the prime was accepted as a word.

Figure 2. Partial effects for self-paced reading latencies. The dashed line in the lower left panel represents the identity priming condition, the solid line the stem priming condition. (The effect of PC1 was linear in log RT, but is nonlinear for RT in ms.)

Figure 1.

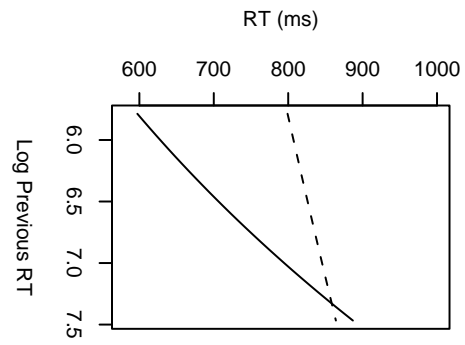
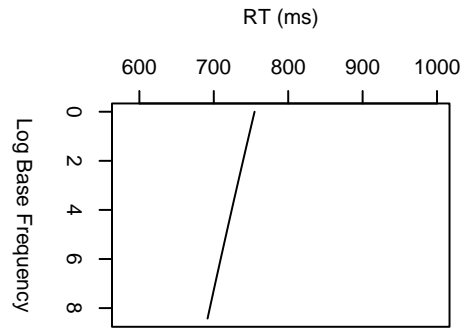
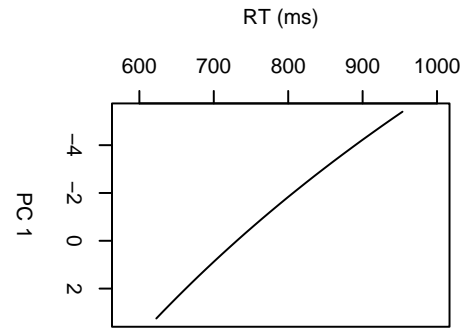
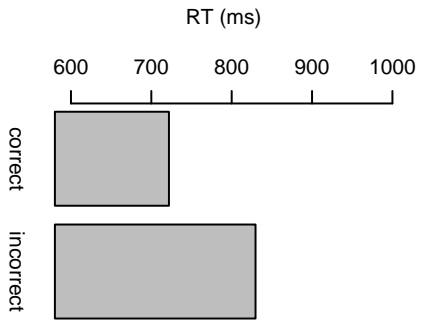
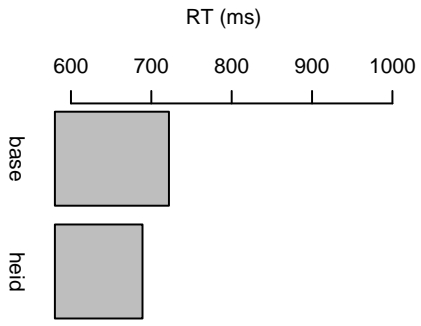


Figure 2.

