The lifespan of lexical traces for novel morphologically complex words

Laura de Vaan[1], Mirjam Ernestus[1,2] and Robert Schreuder[3]
[1]Center for Language Studies, Radboud University Nijmegen / [2]Max Planck Institute for Psycholinguistics, Nijmegen / [3]Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen

This study investigates the lifespans of lexical traces for novel morphologically complex words. In two visual lexical decision experiments, a neologism was either primed by itself or by its stem. The target occurred 40 trials after the prime (Experiments 1 & 2), after a 12 hour delay (Experiment 1), or after a one week delay (Experiment 2). Participants recognized neologisms more quickly if they had seen them before in the experiment. These results show that memory traces for novel morphologically complex words already come into existence after a very first exposure and that they last for at least a week. We did not find evidence for a role of sleep in the formation of memory traces. Interestingly, Base Frequency appeared to play a role in the processing of the neologisms also when they were presented a second time and had their own memory traces.

Keywords: lexical representation, neologisms, lifespan, sleep, morphology

Introduction

In daily life, we are quite often confronted with a novel regular morphologically complex word (see e.g., Baayen & Renouf, 1996). Usually, the meaning of such a neologism is transparent. That is, its meaning is derivable from the meanings of its existing morphemes (e.g., *kortaf+heid* 'abrupt'+'ness'). The question arises whether such a novel word is stored in the mental lexicon after just one encounter. Pinker (1991) argued that this is not the case. Nevertheless, De Vaan, Schreuder, and Baayen (2007) provided experimental evidence that a novel Dutch transparent derivation may leave traces in lexical memory after its very first occurrence. The present paper addresses the lifespan of these lexical traces.

De Vaan and colleagues (2007) investigated the lexical traces of completely regular derivational neologisms using a priming paradigm. They presented the

neologisms visually as isolated words (visual lexical decision) or in short stories (self paced reading). Half of the neologisms were primed by their stems (stem priming), the other half were primed by themselves (identity priming). The prime and target words were separated by a fixed number of trials. In both experimental paradigms, response latencies 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). De Vaan and colleagues interpreted the repetition effect as evidence for the formation of a trace in lexical memory for the neologism. Since the target word was presented very close to the prime (after 39 intervening trials, equal to approximately 1.5 minutes), the question arises whether the observed traces were long lasting or only present for approximately 1.5 minutes. Does such a trace still exist after a couple of hours or even a week? One would expect that the lexical trace loses its strength with time and that its effect is thus less strong after, for instance, twelve hours, than after several minutes.

Dumay, Gaskell, and Feng (2004) and Gaskell and Dumay (2003) investigated the lifespan of lexical traces for novel pseudowords that are morphologically simple. These researchers studied the development of lexical representations by tracing the effects of these new representations on the auditory processing of phonologically similar existing words. They observed that listeners recognize a word such as cathedral more quickly if they have just heard 12 tokens of a phonologically similar pseudoword (e.g., *cathedruke* for *cathedral*). However, five days after the familiarization, listeners recognized the already existing words more slowly. Gaskell and colleagues attributed this inhibitory effect to the pseudowords having become lexical competitors for the existing words. They thus concluded that the formation of a lexical representation takes some time.

Tamminen and Gaskell (2008) studied the effect of new words on the recognition of phonologically similar words over a longer period of time. Like Gaskell and Dumay (2003), they found that lexical competition effects emerge only after a delay. In addition, Tamminen and Gaskell demonstrated that new words remain lexical competitors up to eight months after exposure.

Gaskell and Dumay (2007) investigated whether sleep affects the formation of lexical representations. They observed that if participants learned words at 8 p.m., they did not show (inhibitory) competition effects immediately, but only after a 12 hour interval including a night's sleep. Participants who learned the words at 8 a.m. showed competition effects neither immediately nor after 12 hour of wakefulness, but did so after 24 hours, that is, again after a night's sleep. Thus, sleep is a necessary condition for competition effects to show up. Further support for the role of sleep in establishing traces in lexical memory comes from the study by Davis, Di Betta, Macdonald, and Gaskell (2008). They observed that participants show

lexical competition and faster recognition times only for novel words they learned the previous day, in contrast to words they learned the same day. The two groups of novel words also showed differences in an fMRI study: cortical activation was of equivalent magnitude for novel words the participants had not seen before and novel words learned on the day of scanning, while cortical activation was significantly reduced for novel words learned on the previous day. Hence, fMRI results also suggest effects of sleep in the formation of lexical memory traces. The role of sleep in consolidating memory in general has been extensively documented, see for instance, Stickgold and Walker (2005) and Walker (2005).

The present study investigates the lifespan of regular morphologically complex neologisms (instead of the morphologically simplex pseudowords studied by Dumay et al., 2004; Gaskell & Dumay, 2003, 2007; Tamminen & Gaskell, 2008), including the role of sleep. In addition, we investigated whether the results of De Vaan et al. (2007) based on neologisms created by the combination of existing words with one single suffix (-*heid* '-ness') generalize to neologisms ending in two other Dutch suffixes.

Bertram, Schreuder, and Baayen (2000) argued that the balance between parsing and storage depends on several characteristics of the affix. Words with an unproductive affix tend to be stored in the mental lexicon, whereas words with a productive affix tend to be parsed. In addition, words with an affix that has a productive rival affix tend to be stored, whereas words without such a rival affix tend to be parsed. Finally, words with affixes that do not cause a change in the meaning of the base tend to be parsed, others are either parsed or stored. The three suffixes we investigated are productive, do not have a rival affix, and do not change the meaning of the base. This decreases the likelihood that neologisms with these suffixes are stored in the mental lexicon.

We tested Dutch novel, morphologically well-formed and semantically transparent derived words ending in the suffixes -*heid* 'ness', -*baar* 'able', and -*ing* (e.g., *saprijk+heid* 'juice-rich'+'ness', *ontkleur+baar* 'decolour'+'able', *onthars+ing* 'deresinate'+'ing'). The suffix -*heid*, which was also tested in De Vaan et al. (2007), is the most regular and productive derivational suffix of Dutch. It attaches to adjectival stems and turns them into nouns. The deverbal suffix -*baar* only attaches to transitive verbs, at least in its productive use, which we investigated here. This restriction is directly related to its meaning contribution 'being able to be *Verb*-ed' (Booij, 2002). The suffix turns verbal bases into adjectives. The suffix -*ing* makes action nouns from verbs, and is particularly productive for morphologically complex verbs (Booij, 2002; Haeringen, 1971). In the present study, it is always attached to morphologically complex stems, as for instance in *ont+hars+ing* 'de+resinate+ing'. Suffixation of a stem with -*baar* or -*heid* does not lead to changes in the pronunciation of the stem. Suffixation with -*ing*, in contrast, leads to

resyllabification of the stem-final obstruents to the following syllable formed by the suffix.

We carried out two experiments using a design in which participants were tested in two sessions. Participants were tested for half of the target words in the same session as the presentation of the primes (short priming distance) and for the other half in a different session, which took place 12 hours after the first session (Experiment 1) or one week later (Experiment 2; both long priming distance). The prime was either the stem (stem priming) or the neologism itself (identity priming). The target was always the neologism. To test whether sleep played a role in the formation of lexical traces, we tested half of the participants in Experiment 1 over the course of one day, and the other half with an intervening night's sleep.

Our experimental design provides the possibility to investigate another interesting issue. The words in the Short Priming Distance condition occur either three times in the full experiment (twice in the first part and once in the second part) or twice (once in the first part, primed by their stems, and once in the second part). This implies that we can test whether neologisms are recognized faster (and therefore have stronger lexical representations) if the reader has read them twice instead of only once. We tested this hypothesis in Experiments 1 and 2.

Finally, since this study focuses on neologisms consisting of existing morphemes, the frequencies of these morphemes may affect processing times. This would be in line with most models of morphological processing since they assume that the processing of a morphologically complex word involves the activation of its morphemes, even if the word is processed via its whole word trace in the mental lexicon (e.g., Balling & Baayen, 2008; Burani & Caramazza, 1987; Giraudo & Grainger, 2001; Plaut & Gonnerman, 2000; Schreuder & Baayen, 1995). We included this prediction in our analyses of the data.

Experiment 1: a 12 hours interval

As mentioned above, Experiment 1 was a visual lexical decision task run in Dutch. We used a design in which we crossed Prime Type (stem priming versus identity priming) with Priming Distance (short priming distance versus long priming distance), as illustrated in Table 1. In the case of stem priming, participants first performed lexical decision for the stem. Then they had to respond to the neologism, which was presented either 40 trials later in the same session of the experiment and 12 hours later in the second session of the experiment or only in the second session (stem priming in the short priming distance condition versus stem priming in the long priming distance condition). In the case of identity priming, participants responded only to the neologism, never to its base. The second token

Table 1. Design of Experiment 1. Words in italics form the second exposure, words in bold form the third exposure (further explanation in text).

Session 1:			Session 2:		
		Short Priming Distance		Long Priming Distance	
Prime	Intervening Trial	Target	Intervening Time	Target	Prime Type
Stem ('wobbly')	39 trials	*Neologism* ('wobbliness')	12 hours	**Neologism** ('wobbliness')	Stem Priming
Neologism ('wobbliness')	39 trials	*Neologism* ('wobbliness')	12 hours	**Neologism** ('wobbliness')	Identity Priming
Stem ('wobbly')			12 hours	*Neologism* ('wobbliness')	Stem Priming
Neologism ('wobbliness')			12 hours	*Neologism* ('wobbliness')	Identity Priming

appeared 40 trials after the first token or 12 hours later in the second session of the experiment (identity priming in the short priming distance condition versus identity priming in the long priming distance condition). If the second token occurred after 40 trials, a third token was presented in the second session.

We included the stem priming condition as a baseline in order to investigate whether priming in the identity priming conditions results from the full neologism or just its stem. If it is just the stem of the neologism that induces priming, the stem and the neologism should show similar priming effects. If, in contrast, participants stored the full neologisms, the neologisms should show more priming than just their stems.

Participants were asked to classify regular morphologically complex neologisms consisting of existing morphemes as words. They performed the second session of the experiment either after 12 hours of awakeness or after 12 hours with a night's sleep.

Method

Participants. Twenty-four native speakers of Dutch were paid to take part in this experiment. Six of them were male and eighteen were female. The participants had at the time of the experiment a mean age of twenty one, in a range of 18 to 37. Twenty-two of them were undergraduate students at the Radboud University of Nijmegen, two participants were already graduated.

Materials. The target materials consisted of 120 Dutch, morphologically well-formed and semantically transparent derived neologisms: 40 neologisms with the suffix -heid (these neologisms were taken from De Vaan et al. (2007), e.g., *pitloos+heid* 'seedless' + 'ness'), 40 neologisms with the suffix -*baar* (e.g., *bewapen+baar* 'arm' + 'able'), and 40 neologisms with the suffix -*ing* (e.g., *beboter+ing* 'butter' + 'ing', see the Appendix for all neologisms). The suffix -*heid* was attached to adjectival stems and turned them into nouns, the deverbal suffix -*baar* was attached to transitive verbs and turned them into adjectives, while the suffix -*ing* was attached to morphologically complex verbs and turned them into action nouns.

All 120 neologisms had a different stem. Some stems are monomorphemic words, other stems are complex words, but all stems are existing words and are listed in the celex database (Baayen, Piepenbrock, & Gulikers, 1995). In contrast, the neologisms are not listed in the CELEX database. Moreover, none of the neologisms with the suffix -*heid* appeared over ten times in the Dutch section of the world wide web (as of January 2003). For the neologisms with the suffixes -*baar* and -*ing*, 96% had a frequency below ten in this part of the world wide web (as of June 2006).

For session 1 of the experiment, we created 4 master lists. A given target neologism was represented in each list in a different condition and each condition was equally represented in each master list as was each suffix. So, we had 30 stems, 30 stems followed by their 30 corresponding neologisms after 39 intervening trials, 30 neologisms, and 30 neologisms repeated by themselves after 39 intervening trials. This makes 120 prime words and 60 target words for session 1.

To distract participants' attention from the neologisms (which might trigger specific strategies), we added existing filler words with different stems and suffixes than these in the target words, 40 ending in -*achtig*, 40 ending in -*eer*, and 40 ending in -*schap*. The filler words all appear in celex with a lemma frequency of maximally 14, and have a mean frequency of 1.88 and a median of 0. We added these filler words to the master lists in the same four conditions as the neologisms: 30 stems, 30 stems followed by their 30 corresponding complex words after 39 intervening trials, 30 complex words, and 30 complex words repeated by themselves after 39 intervening trials.

Finally, we added 360 pseudowords that follow the phonotactic and orthographic conventions of Dutch in the same conditions as the targets words. Each resulting master list then also contained 60 pseudo stems, 60 pseudo stems followed by their 60 corresponding pseudo complex words after 39 intervening trials, 60 pseudo complex words, and 60 pseudo complex words repeated by themselves after 39 intervening trials. The pseudowords differed from the existing words and neologisms in the experiment by one or two letters. The differences always occurred in the words' stems. In order to exclude the possibility that the pseudo-

words primed our target neologisms, we checked whether three other participants could guess the words from which the pseudowords were derived. The participants were able to do so for only 6% of the pseudowords.

We randomized all master lists six times, which resulted in 24 lists. We ensured that the number of trials between the words sharing their stems was held constant at 39 intervening trials. Each of the resulting lists was preceded by 24 practice trials (six stems, six words, six pseudo stems, and six pseudo complex words). All practice words differed in their stems from each other and from the other words in the experiment.

For session 2 of the experiment, we created a master list with all 120 neologisms. Half of these neologisms were primed once in session 1 of the experiment (by their stems or by themselves), the other half were primed twice in session 1 of the experiment (half of the neologisms primed twice were primed once by their stems and once by themselves, the other half were primed twice by themselves).

To distract attention from the neologisms, we added 40 existing complex words (filler words) ending in *-aar*, 40 ending in *-er*, and 40 ending in *-lijk/-elijk*. These filler words all appear in celex and have a CELEX lemma frequency in the range of [0–10], a mean of 2.44, and a median of 2.

We matched each filler word and each neologism with a pseudoword by changing one or two letters in their stems, making sure that the resulting pseudowords did not violate the phonotactic or orthographic conventions of Dutch. To exclude the possibility that the pseudowords primed our target neologisms, we checked also for these pseudowords whether the three participants could guess the words from which the pseudowords were derived. The participants were able to do so for only 7% of the pseudowords. The resulting list contained 240 word trials and 240 pseudoword trials.

We randomized this list six times. Each of the resulting lists was preceded by 24 practice trials. Half of the practice trials were words, half were pseudowords; each of the six suffixes (*-baar, -heid, -ing, -aar, -er, -lijk/-elijk*) were presented an equal number of times. The practice trials or their stems did not appear in the experimental part of the experiment.

Procedure. Participants were tested individually in a noise-attenuated experimental room. They first read the instructions on the computer screen. These instructions made clear that both existing and possible Dutch words required a "yes" response, and that only words with non-existing morphemes required a "no" response. The first session consisted of two blocks separated by a break. Target words were part of the same block as their prime words.

In both sessions of the experiment, an experimental trial began with a fixation mark, positioned in the center of a Nec Multisync color monitor for 500ms. After

50ms, the stimulus appeared centered at the same position. Stimuli were presented in black lowercase, 36 point letters, on a white background and they remained on the screen for 1750 ms.

Half of the participants were tested in the morning and in the evening of the same day, the other half were tested during the evening and the following morning. The interval between the two sessions of the experiment was fixed at 12 hours. The testing hours lay between 8am and 10am and between 8 pm and 10 pm. The experimental setting and the experimenter were the same in both sessions of the experiment.

Results and Discussion

We first analyzed the response latencies to the target neologisms that were primed only once (in italics in Table 1). We coded no-responses as errors and removed these data points from the analyses of the response latencies. The percentage of errors for a neologism ranged from 4% to 58%. As we are dealing with neologisms in lexical decision, these high error rates are unsurprising. Moreover, we removed all data points of the nine words with error rates over 50% (*annoteerbaar* 'annotatable', *bebrildheid* 'spectacledness', *begiering* 'feeding with liquid manure', *bekonkeling* 'hatching', *betralieing* 'barring', *föhnbaar* 'blow-dryable, *omkeiling* 'knock flying', *ontmugging* 'disinsected' and *verboersing* 'becoming rustic'). Finally, we left out of consideration all trials in which participants did not react or reacted only after 1750 ms. We then applied a logarithmic transformation to the remaining response latencies to reduce the skewness of the distribution.

The dependent variable in our analysis is response latency (RT) to the target neologism. We added Prime Type (stem priming versus identity priming), Priming Distance (short priming distance versus long priming distance) and Sleep (yes or no) as our main predictors. We also added Base Frequency as covariate (the lemma frequency of the stem of a neologism) as obtained from CELEX. We logarithmically transformed Base Frequency in order to reduce the skewness in the distribution. Finally, we also included as predictors for a given participant's response, that participant's response latency to the prime of that target word (henceforth Prime RT) and the lexical decision for that prime (word or nonword, henceforth Prime Decision). We added the latter predictor because participants erroneously classified the primes as non-existing words in 24% of the trials.

We fitted a mixed-effect model of covariance, using a stepwise variable selection procedure, with participant and word as crossed random variables (e.g., Baayen, Davidson, & Bates, 2008; Bates & Sarkar, 2005; Faraway, 2006). Inspection of the residuals revealed marked non-normality, indicating serious lack of

Table 2. Results of the Stepwise Multilevel Regression Model (df = 2191) fit to the Reaction Times of Experiment 1 for target words primed only once. Prime Type provides the effect of identity priming relative to stem priming. Distance provides the effect of long priming distance relative to short priming distance. Decision provides the effect of an incorrect decision to the prime relative to a correct decision to the prime. The standard deviation for the random effect of Word was estimated at 0.0489, and that for the Participant random effect at 0.1191. In addition, there was a significant random effect for Participant by Base Frequency (log-likelihood ratio = 11.72, $p = 0.003$, standard deviation = 0.0116, estimated coefficient for the correlation = −0.599). The residual standard deviation was 0.1925.

	$\hat{\beta}$	t-value	p-value
(Intercept)	5.3487	32.65	< .0001
Prime Type	−0.0535	−6.25	< .0001
Distance	0.5340	2.74	0.0062
Sleep	−0.0635	−1.52	0.13
Decision	1.6067	6.50	< .0001
Prime RT	0.2058	8.69	< .0001
Base Frequency	−0.0165	−4.19	< .0001
Decision by Prime RT	−0.2200	−6.10	< .0001
Distance by Prime RT	−0.0711	−2.46	0.01
Sleep by Decision	−0.1087	−4.58	< .0001

goodness-of-fit for this model. We therefore removed outliers with standardized residuals outside the interval of [−2.5, 2.5] (1.67%), 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 a better goodness of fit. This final model is listed in Table 2.

Long priming distance generally resulted in longer reaction times to target neologisms than short priming distance and participants reacted less quickly if they had made an error on the prime. Furthermore, response latencies increased with increasing Prime RTs, but only if the participant had accepted the prime as a word. This positive effect of Prime RT was smaller after long than after short priming distance.

More importantly for our research questions, we observed an effect of Prime Type: Response latencies to the target word were shorter in the identity priming condition than in the stem priming condition. Prime Type did not interact with Prime Distance suggesting that the priming effect is as strong after long priming distance as after short priming distance. In other words, the lexical trace for a neologism stays in the mental lexicon for at least 12 hours.

These results are very similar to the results of De Vaan et al. (2007). Also in their experiment, participants responded more quickly to a neologism if primed by itself than if primed by its stem. De Vaan and colleagues only tested neologisms ending in the suffix *-heid*. Since we tested three suffixes, we can conclude that the effect reported by De Vaan and colleagues is not restricted to the suffix *-heid*: Lexical traces are also formed after the single presentation of morphologically complex words ending in other suffixes, and all these traces are long lasting, surviving at least several hours.

Interestingly, our results did not show a significant main effect for sleep, but only a significant interaction of Sleep by Prime Decision. Participants generally reacted more slowly to a target neologism if they had rejected the prime as a word, but this effect was smaller after a night's sleep. Participants were thus less influenced by their initial decision after sleep.

Since sleep did not interact with Prime Type, sleep appears not to affect the formation of lexical traces. In contrast with previous research by Gaskell and Dumay (2007), who have shown that sleep is crucial for the formation of lexical traces for novel morphologically simple pseudowords, our results and the results of De Vaan et al. (2007) indicate that if a novel word is morphologically complex, sleep is not crucial. We will return to this issue in the general discussion.

The effect of Prime Type strongly suggests that the neologisms, at their second presentation, were processed via their full form representations. Nevertheless, we also observed that the response latencies from most participants decreased with increasing Base Frequency (we observed a fixed effect of Base Frequency and a random effect by participant for Base Frequency). Since there was no significant interaction of Base Frequency by Prime Type, we can conclude that these participants processed the neologisms through their stems, at both the first and second presentation. Apparently, most participants activated the morphemes forming the neologisms as well as the lexical traces formed by the neologisms.[1]

We now address the question whether participants recognized a neologism more quickly if they had seen the neologism twice before instead of only once the neologism and once its stem. We fitted a mixed-effect model of covariance to the RTs of neologisms in the second session of the experiment with two primes (in bold in Table 1). We again made use of a stepwise variable selection procedure, and entered the same independent variables as in the first analysis. In the present analysis, however, we did not only add the Prime RT, but the RT to the second exposure as well (Prime RT 2) as covariates and we also included as predictor the Decision to this second exposure (Prime Decision 2). Because of a correlation of Base Frequency with Prime RT 2, we orthogonalized these covariates by using the residuals of a model predicting Base Frequency as a function of Prime RT 2 instead of the raw values of Base Frequency.[2] In this analysis, we also removed

Table 3. Results of the Stepwise Multilevel Regression Model (df = 1111) fit to the Reaction Times of target words primed twice in Experiment 1. Decision and Decision 2 provide the effects of an incorrect decision to the first and second primes, respectively, relative to a correct decision. The standard deviation for the random effect of Word was estimated at 0.0458 and of Participant at 0.1011. The residual standard deviation was 0.1940.

	$\hat{\beta}$	t-value	p-value
(Intercept)	5.7810	36.23	< .0001
Base Frequency	−0.0174	−5.32	< .0001
Decision	0.0684	4.36	< .0001
Distance 2	0.0785	3.67	< .0001
Prime RT 2	0.1271	5.33	< .0001

outliers with standardized residuals outside the interval [−2.5, 2.5] (1.93%), and refitted the model. The final model is listed in Table 3.

The results show no effect for Prime Type. This suggests that there was no difference in the strength of memory traces for neologisms that participants had seen only once and for neologisms that they had seen twice in the first session of the experiment. The other predictors, including Base Frequency, showed effects that could be expected given the analysis of the RTs for the tokens of the target words primed only once.

In summary, Experiment 1 shows that a lexical trace is formed after a single exposure for Dutch morphologically complex neologisms ending in *-heid*, *-baar*, and *-ing*. These traces last for at least 12 hours and are not affected by a night's sleep. The strength of a memory trace (after 12 hours) is also not affected by the repetition of the neologism within 1.5 minutes after the first exposure. Finally, participants activated the memory trace in the processing of the second (and third) presentation of the neologism in addition to activating its constituent morphemes.

Experiment 2: a week interval

We observed in Experiment 1 that a memory trace for a morphologically complex neologism is as strong after 12 hours as it is after 39 intervening trials. The key question addressed in Experiment 2 is whether a memory trace for a neologism remains in the mental lexicon for more than 12 hours. Will there still be a memory trace after one week? Will this trace be as strong as it is after 39 intervening trials?

Method

Participants. Twenty-four undergraduate students at the Radboud University of Nijmegen were paid to take part in this experiment. Six of them were male and 18 were female, and at the time of the experiment, they all had a mean age of 20, in the range of 18 to 24. All were native speakers of Dutch, and none had participated in Experiment 1 or in similar experiments using similar materials.

Materials. We used the same materials and design as in Experiment 1.

Procedure. The procedure was identical to the procedure of Experiment 1, except that session 2 of the experiment was presented to the participants one week, instead of 12 hours, after the first session.

Results and Discussion

We followed the same procedure for analyzing the data as for Experiment 1. We coded no-responses to the target neologisms as errors and removed these data points from the analyses of the response latencies. We also removed all data points of the eight neologisms with error rates over 50% (*annoteerbaar* 'annotable', *begiering* 'feeding with liquid manure', *bepoeiering* 'dusting', *betralieing* 'barring', *omkeiling* 'knock flying', *ontmugging* 'disinsected', *onttuiging* 'stripping down', *verboersing* 'becoming rustic'). Finally, we left out of consideration all trials in which participants did not react or reacted only after 1750 ms.

Table 4. Results of the Stepwise Multilevel Regression Model (df = 2241) fit to the Reaction Times to the target words primed only once in Experiment 2. Prime Type provides the effect of identity priming relative to stem priming. Distance provides the effect of long priming distance relative to short priming distance. Decision provides the effect of an incorrect decision to the prime relative to a correct decision to the prime. The standard deviation for the random effect of Word was 0.0575, and that for the Participant random effect was 0.1137. The residual standard deviation was 0.2070.

	$\hat{\beta}$	*t*-value	*p*-value
(Intercept)	5.5376	33.64	< .0001
Prime Type	−0.0446	−4.97	< .001
Distance	0.6975	3.43	< .0001
Decision	0.1057	8.35	< .0001
Prime RT	0.1723	7.13	< .0001
Base Frequency	−0.0145	−4.12	< .0001
Distance by Prime RT	−0.0932	−3.08	< .001

We again fitted first a mixed-effect model of covariance to the RTs for the targets primed only once, using a stepwise variable selection procedure. We considered participant and word as crossed random variables and Prime Type, Priming Distance, Base Frequency, Prime RT, and Prime Decision (participants erroneously judged the primes as non-words in 24% of the trials) as fixed predictors. The RTs and Base Frequency were again logarithmically transformed. We removed outliers with standardized residuals outside the interval [−2.5, 2.5] (1.45%), and refitted the model. The final model is listed in Table 4.

As in Experiment 1, we observed a main effect for Prime Type. Response latencies were shorter in the identity priming condition than in the stem priming condition. We observed no significant interaction of Prime Type by Priming Distance, which indicates that the magnitude of the effect of Prime Type is similar after long priming distance as after short priming distance. In other words, the lexical trace for a neologism appears to stay in the mental lexicon for at least a week, without losing strength.

Furthermore, we observed similar main effects for Prime Decision and Base Frequency as in Experiment 1. Participants reacted more quickly if they had accepted the prime as a word and to neologisms with more frequent stems. Finally, we observed main effects and an interaction for Priming Distance and Prime RT: Similar to the results of Experiment 1, participants reacted more quickly after short than after long distance priming, and this difference was smaller the more quickly the participant reacted to the prime.[3]

Table 5. Results of the Stepwise Multilevel Regression Model (df = 1177) fit to the Reaction Times of target words primed twice in Experiment 2. Decision provides the effect of an incorrect decision to the prime relative to a correct decision to the prime. The standard deviation for the random effect of Word was estimated at 0.0585, and that for the Participant random effect at 0.1230. In addition, there was a significant random effect for Participant by Base Frequency (log-likelihood ratio = 22.96, $p < .0001$, standard deviation = 0.0121, estimated coefficient for the correlation = −0.780). The residual standard deviation was 0.1983.

	$\hat{\beta}$	t-value	p-value
(Intercept)	5.2624	22.45	< .0001
Decision	0.9507	3.02	0.0026
Prime RT	0.1497	5.31	< .0001
Prime RT 2	0.0669	2.80	0.0053
Base Frequency	−0.0119	−2.47	0.0138
Distance by Prime RT	−0.1326	−2.88	0.0041

We then addressed the question whether participants recognized neologisms more quickly if they had seen the neologism twice before instead of only once the neologism itself and once its stem. We fitted a mixed-effect model of covariance to the RTs of the neologisms in the second session of the experiment with two primes. We again made use of a stepwise variable selection procedure, and entered the same independent variables as in the analysis of Experiment 1, namely Prime Type, Prime RT, Prime RT 2, Prime Decision, Prime Decision 2, and Base Frequency. Also in this analysis, we removed outliers with standardized residuals outside the interval [−2.5, 2.5] (1.41%) and refitted the model. The final model is listed in Table 5.

As in Experiment 1, the results show no effect for Prime Type. This supports the suggestion from Experiment 1 that there is no difference in the strength of memory traces for neologisms that participants had seen only once and for neologisms that they had seen twice in the first session of the experiment. The other predictors, including Base Frequency, showed effects that we also saw in the analyses presented above.

In conclusion, lexical memory traces for regular morphologically complex neologisms last for at least one week in the mental lexicon, and appear as strong after a week as after 39 intervening trials. Furthermore, as in Experiment 1, participants showed evidence both for lexical traces and for activation of the constituent morphemes, even when they read the neologisms for the second or third time.

General Discussion

The questions addressed in this paper are: (1) Do regular morphologically complex neologisms ending in a variety of suffixes and presented just once leave detectable traces in lexical memory? (2) Does a memory trace remain in the mental lexicon for at least 12 hours or even a week? (3) Does sleep play a role in the formation of such lexical traces?

De Vaan et al. (2007) showed that regular Dutch morphologically complex words ending in the productive suffix *-heid* are represented in the mental lexicon after the very first exposure. Experiments 1 and 2 in the present paper, both priming experiments in which the prime was either the stem of the target neologism or the neologism itself, shows that also regular morphologically complex words ending in the productive suffixes *-baar* and *-ing* are stored in the mental lexicon after the very first exposure. These affixes are productive, do not have a rival, and do not change the meaning of the base. According to Bertram et al. (2000), neologisms ending in these suffixes are therefore not likely to be stored. Nevertheless, we observed clear lexical traces for such neologisms.

Moreover, Experiments 1 and 2 show that these traces are still detectable after 12 hours (Experiment 1) and even after a week (Experiment 2) and that they do not appear to lose strength. This result is important because even native adult speakers often encounter novel complex words and they have encountered a huge number of existing complex words only once or twice (see for lexical statistic evidence e.g., Zipf, 1935). We conclude that the mental lexicon has many more entries than has traditionally been assumed (e.g., Alegre and Gordon, 1999). An alternative explanation of our priming results may be that participants did not so much form a lexical trace for the neologism itself but rather remembered how they morphologically analyzed the neologism when they first encountered it, which would be in line with a decomposition-first model (Rastle, Davis, & New, 2004; Taft, 2004). We think, however, that these two accounts are in fact two sides of the same coin. If readers remember how they analyzed a word, then they must remember the word itself as well, which implies that there is a memory trace for the word.

It may be that the participants just showed strategic, task-dependent effects. We would like to note that, if so, this does not affect our interpretation of the results. Participants may have decided to store the neologisms because they noticed that some of these words re-occurred later in the experiment. Whether they were able to store these neologisms (for whatever reason) and whether these traces stay in memory for more than a few minutes form exactly the research topics of the present study. Further, since De Vaan et al. (2007) found the same priming results in a self-paced reading experiment, in which the neologisms were presented in short sentences, we believe that these priming effects are not task specific and therefore are not likely to be purely strategically based.

The question arises whether the lexical trace formed by the neologism is an episodic trace, and thus contains information on indexical properties of the word token (e.g., the font of the word token, whether it appeared in a cold room, whether the participant had to sneeze after the word token), or whether the neologism was stored in the form of an abstract lexical representation, specifying the word's form only in abstract linguistic units (e.g., graphemes, phonemes, morphemes). This is a question that we are currently addressing in a new series of experiments. In these experiments, we investigate, among other things, whether the lexical trace is modality specific, for instance, whether the occurrence of a neologism in a spoken text facilitates the visual comprehension of the same word in visual lexical decision experiments carried out a week later. If so, the memory trace is likely to be an abstract lexical representation rather than a pure episodic memory trace.

Gaskell and colleagues (Dumay et al., 2004; Gaskell & Dumay, 2003; Gaskell & Dumay, 2007; Tamminen & Gaskell, 2008) also investigated the formation of lexical traces for neologisms. They observed that participants needed at least 12 exposures to learn neologisms and moreover that the formation of a lexical rep-

resentation required an incubation-like period crucially associated with sleep. Especially the consolidation stage would rely on the process of sleep (Stickgold & Walker, 2005; Walker, 2005). In contrast to these findings, our results show that one exposure may be sufficient for the formation of a lexical trace and that sleep is not a prerequisite. An explanation for this discrepancy in results may be found in two differences in the materials and designs between their and our experiments.

First, whereas Gaskell and colleagues tested morphologically simplex neologisms, such as *cathedruke*, we focused on fully regular morphologically complex neologisms consisting of morphemes that are already present in the mental lexicon. Therefore, our neologisms can be morphologically linked to other lexical representations, and may consequently be better embedded in the mental lexicon. This may lead to faster and more robust storage. It is also possible that combinations of existing morphemes have a different kind of lexical trace than monomorphemic words: Rather than forming independent representations by themselves, they may purely consist of the information that the relevant morphemes may be combined to form a word. For example, the word *bewapenbaar* ('armable') may be represented as just the possibility to combine the morphemes *be*, *wapen*, and *baar* (Balling & Baayen, 2008; Bybee, 1985).

Second, we studied the formation of lexical traces directly by comparing response latencies to neologisms when they had already been presented before and when they had only been primed by their stems (i.e., identity priming versus stem priming). Gaskell and colleagues, on the other hand, studied lexical storage more indirectly by tracing the consequences of acquired neologisms on the processing of well-established existing words. Their indirect test may be less sensitive than our priming procedure.

Base Frequency showed a facilitatory effect in Experiment 1 as well as in Experiment 2. A morphologically complex neologism is, necessarily, processed through morphological decomposition. Surprisingly, however, for most participants the effect of Base Frequency was independent of whether these participants had seen the neologism before. This suggests that morphological decomposition played a role in these participants' processing of both the first and the second token of the neologism. Thus, our results indicate that most participants relied for the processing of the second token both on the lexical trace of the neologism and on morphological decomposition. These results are in line with the conclusions presented by Balling and Baayen (2008) that the direct route and the parsing route in the dual route model (see, e.g., Bertram et al., 2000; Schreuder & Baayen, 1995) are not independent (De Jong, Schreuder, & Baayen, 2003). Both routes eventually activate the same semantic representation in the lexicon and this lexical activation resonates through the whole system. The effect of Base Frequency is also in line with the supra-lexical hypothesis (Giraudo & Grainger, 2001), which states that

the activation of a word may lead to the activation of its constituents, and with the model of Plaut and Gonnerman (2000), which assumes independent routes for morphologically simple and complex words, ascribing morphological effects to orthographic, phonotactic, and semantic similarities.

In conclusion, our study has clearly shown that the first occurrence of a neologism consisting of existing morphemes leaves a long lasting trace in the mental lexicon. The recognition of the later occurrences of such a neologism proceeds via this trace while at the same time it also activates its constituents.

Notes

1. A logistic mixed-effect model of covariance on the errors (a binomial variable) with participant and word as crossed random variables showed significant $\hat{\beta}$s for Prime RT ($\hat{\beta}=1.61751$, $z=3.780$, $p=0.0002$), Prime Decision ($\hat{\beta}=10.23143$, $z=3.251$, $p=.0012$), Base Frequency ($\hat{\beta}=-0.14840$, $z=-3.275$, $p=.0011$), and for the interaction of Prime Decision with Prime RT ($\hat{\beta}=-1.22406$, $z=-2.685$, $p=.0073$). Participants made fewer errors if they had reacted more quickly on the primes, if they had judged these primes correctly as words (especially if this had taken some time), and if the neologism had a higher base frequency. The error analysis showed no effects for Prime Type or Priming Distance, and no interactions of Priming Distance with Prime RT, or Sleep with Prime Decision. Our data thus show no evidence for speed-accuracy trade-offs.

2. No correlation was present for the predictors in the analysis of the RTs to the second exposure, analyzed above.

3. A logistic mixed-effect model of covariance on the errors (a binomial variable) with participant and word as crossed random variables showed significant $\hat{\beta}$s for Prime Type ($\hat{\beta}=-0.49488$, $z=-3.699$, $p=0.0002$), Base Frequency ($\hat{\beta}=-0.13809$, $z=-2.939$, $p=0.0033$), Prime RT ($\hat{\beta}=0.84946$, $z=2.715$, $p=0.0066$), and Prime Decision ($\hat{\beta}=1.85377$, $z=13.061$, $p<.0001$). A negative $\hat{\beta}$ indicates a reduction in errors. Participants made fewer errors for neologisms primed by these neologisms themselves than for neologisms primed by their stems, for neologisms with stems of a higher frequency, if they had reacted more quickly on the corresponding primes, and if they had correctly classified these primes as words. The error analysis showed no effect for Priming Distance or an interaction of Priming Distance with Prime RT. Our data thus show no evidence for speed-accuracy trade-offs.

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

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.

Balling, L., & Baayen, R. (2008). Morphological effects in auditory word recognition: Evidence from Danish. *Language and Cognitive Processes, 23*(7/8), 1159–1190.

Bates, D. M., & Sarkar, D. (2005). *The lme4 library*. [On-line], Available: http://lib.stat.cmu.edu/R/CRAN/.

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.

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

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

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

Crawley, M. J. (2002). *Statistical computing. An introduction to data analysis using S-plus*. Chichester: Wiley.

Davis, M. H., Di Betta, A. M., Macdonald, M. J. E., & Gaskell, M. G. (2008). Learning and consolidation of novel spoken words. *Journal of Cognitive Neuroscience, 21*(4), 803–820.

De Jong, N. H., Schreuder, R., & Baayen, R. H. (2003). Morphological resonance in the mental lexicon. In R. H. Baayen, & R. Schreuder (Eds.), *Morphological structure in language processing* (pp. 65–88). Berlin: Mouton de Gruyter.

De Vaan, L., Schreuder, R., & Baayen, R. H. (2007). Regular morphologically complex neologisms leave detectable traces in the mental lexicon. *The Mental Lexicon, 2*(1), 1–24.

Dumay, N., Gaskell, M. G., & Feng, X. (2004). A day in the life of a spoken word. In K. Forbus, D., Gentner, & T. Regier (Eds.), *Proceedings of the twenty-sixth annual conference of the cognitive science society* (pp. 339–344). Mahwah, NJ: Erlbaum.

Faraway, J. J. (2006). *Extending Linear Models with R: Generalized Linear, Mixed Effects and Nonparametric Regression Models*. Boca Raton, FL: Chapman & Hall/CRC.

Gaskell, M. G., & Dumay, N. (2003). Lexical competition and the acquisition of novel words. *Cognition, 89*, 105–132.

Gaskell, M. G., & Dumay, N. (2007). Sleep-associated changes in the mental representation of spoken words. *Psychological Science, 18*(1), 35–39.

Giraudo, H., & Grainger, J. (2001). Priming complex words: Evidence for supralexical representation of morphology. *Psychonomic Bulletin and Review, 8*, 127–131.

Haeringen, C. B. v. (1971). Het achtervoegsel -ing: Mogelijkheden en beperkingen [The suffix -ing: Possibilities and restrictions]. *De Nieuwe Taalgids, 64*, 449–468.

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

Plaut, D. C., & Gonnerman, L. M. (2000). Are non-semantic morphological effects incompatible with a distributed connectionist approach to lexical processing? *Language and Cognitive Processes, 15*(4/5), 445–485.

Rastle, K., Davis, M. H., & New, B. (2004). The broth in my brother's brothel: Morpho-orthographic segmentation in visual word recognition. *Psychonomic Bulletin & Review, 11*(6), 1090–1098.

Schreuder, R., & Baayen, R. H. (1995). Modeling morphological processing. In L. B. Feldman (Ed.), *Morphological Aspects of Language Processing* (pp. 131–154). Hillsdale, NJ: Lawrence Erlbaum.

Stickgold, R., & Walker, M. P. (2005). Memory consolidation and reconsolidation: what is the role of sleep? *Trends in Neurosciences, 28*(8), 408–415.

Taft, M. (2004). Morphological decomposition and the reverse base frequency effect. *The Quarterly Journal of Experimental Psychology, 57A*, 745–765.

Tamminen, J., & Gaskell, M. G. (2008). Newly learned spoken words show long-term lexical competition effects. *The Quarterly Journal of Experimental Psychology, 61*(3), 361–371.

Walker, M. P. (2005). A refined model of sleep and the time course of memory formation. *Behavioral and Brain Science, 28*, 51–104.

Zipf, G. K. (1935). *The Psycho-Biology of Language*. Boston: Houghton Mifflin.

Appendix. Experimental materials of Experiments 1 and 2:

annoteerbaar; arceerbaar; balsembaar; bekeurbaar; bekogelbaar; bekrasbaar; bespuwbaar; bestijgbaar; bestormbaar; bestraatbaar; bestuifbaar; betoverbaar; bevuilbaar; bewapenbaar; blondeerbaar; braadbaar; föhnbaar; frankeerbaar; gedenkbaar; omspanbaar; ontkalkbaar; ontkleurbaar; ontruimbaar; ontrukbaar; onttroonbaar; ontwapenbaar; overreedbaar; paneerbaar; pocheerbaar; scandeerbaar; shockeerbaar; tutoyeerbaar; verfraaibaar; verknoeibaar; verkreukbaar; vernauwbaar; verpootbaar; verpotbaar; verspilbaar; verversbaar

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

afkluiving; afzouting; bebotering; bedamming; bedrinking; bedrupping; begiering; behakking; bekonkeling; bekruising; bekwijling; benageling; beolieing; bepareling; bepoeiering; beriddering; bespijkering; betralieing; bewalming; bijeenlegging; omduikeling; omkeiling; ontadeling; ontharsing; ontmugging; onttuiging; ontvloeiing; ontzadeling; oppieping; toezinging; verboersing; vergokking; verkijking; verkletsing; verkuiling; vernaaiing; verrijding; versmeding; verspelding; wegstoting

Author's address

Laura de Vaan
Center for Language Studies
Radboud University Nijmegen
P.O. Box 310
6500 AH Nijmegen
The Netherlands

laura.devaan@mpi.nl