Effects of affective ratings and individual differences in English morphological processing

Kaidi Lõo (kloo@ualberta.ca)
Department of Linguistics, University of Alberta
4-32 Assiniboia Hall, Edmonton, AB T6G 2E7, Canada

Abigail Toth [1,2] (a.g.toth@rug.nl)
[1]Department of Linguistics, University of Alberta
4-32 Assiniboia Hall, Edmonton, AB T6G 2E7, Canada
[2]Department of Artificial Intelligence, University of Groningen
Nijenborgh 9, 9747 AG Groningen, The Netherlands

Figen Karaca (karaca@ualberta.ca)
Department of Linguistics, University of Alberta
4-32 Assiniboia Hall, Edmonton, AB T6G 2E7, Canada

Juhani Järvikivi (jarvikiv@ualberta.ca)
Department of Linguistics, University of Alberta
4-55 Assiniboia Hall, Edmonton, AB T6G 2E7, Canada

Abstract

The nature of morphological processing has remained a controversial topic in psycholinguistic research. Some studies (e.g., Rastle, Davis, & New, 2004) have argued that when we read words like *corner* and *talker*, we automatically decompose them into existing morphemes like *talk*, *corn*, and *-er*, regardless of whether it is semantically plausible (e.g., *talker*) or not (e.g., *corner*). Recent studies, however, have challenged this view, by showing early semantic effects of the whole complex word (Järvikivi & Pyykkönen, 2011; Lõo & Järvikivi, 2019; Milin, Feldman, Ramscar, Hendrix, & Baayen, 2017). Using a masked priming paradigm, the present study only found effects of morphological decomposition for true morphological relations (e.g., *talker*) as well as effects of frequency and affective properties of whole words, further challenging automatic decomposition accounts. Finally, we also report that individual differences such as participants' self-reported scholarly reading and openness to new experience, affect processing.

Keywords: morphological processing; masked priming; affective properties; individual differences

Introduction

A large body of psycholinguistic research has focused on the question of how people read words like *cats* or *puppy*. More precisely, the question is whether these words are understood by accessing their morphemic components, for example *cat*, *-s*, *pup*, and *-y* or whether they are processed as any simple word, without recourse to internal structure.

From early on (Taft & Forster, 1975; Manelis & Tharp, 1977) both views have been represented in various forms. Recently, a particularly prominent view has been a variant of the former which states that all morphologically complex words are automatically decomposed in lexical access (Beyersmann et al., 2016; Lázaro, Illera, & Sainz, 2016; Longtin, Segui, & Hallé, 2003; Marslen-Wilson, Bozic, & Randall, 2008; Rastle, Davis, Marslen-Wilson, & Tyler, 2000; Rastle et al., 2004; Rastle & Davis, 2008). Most strikingly, this view takes the decomposition process to operate on the word form alone, without access to any semantic aspects of the word, with the prediction being that all word forms with apparent internal structure should be processed alike.

This approach has found support from masked priming studies (see e.g., Rastle et al., 2004) demonstrating that both pseudo-complex words, where the potential morphemic parts (e.g., *corn* and *-er*) do not make up the meaning of the whole word (e.g., *corner*), as well as transparent complex words with morphemic parts (e.g., *talk* and *-er*) that clearly contribute to the meaning of the whole word (e.g., *talker*), equally facilitate the recognition of their stems (*corn* and *talk*, respectively). Not only that, this research has also shown that words that are not exhaustively divisible into two morphemes, like *turnip* (where *-ip* is not an English affix), do not behave this way, suggesting that automatic decomposition is not only agnostic to semantics but is also driven by online analysis of linguistic structure.

However, not all recent research aligns with this view. Recent studies considering semantic and whole-word properties of the words have started to question this rather simplistic approach to language processing, especially in the case of morphologically complex languages, such as Serbian, Finnish and Estonian (Milin, Filipović Durdević, & Moscoso del Prado Martín, 2009; Moscoso del Prado Martín, Bertram, Häikiö, Schreuder, & Baayen, 2004; Lõo, Järvikivi, & Baayen, 2018; Lõo, Järvikivi, Tomaschek, Tucker, & Baayen, 2018), but also for English (Baayen, Wurm, & Aycock, 2007; Schmidtke, Matsuki, & Kuperman, 2017). For instance, these studies show whole-word frequency effects (Baayen et al., 2007; Schmidtke, Matsuki, & Kuperman, 2017; Lõo et al., 2018), as well as paradigmatic effects (Milin et al., 2009; Moscoso del Prado Martín et al., 2004; Lõo et al., 2018) in the processing of complex words, which does not align well with the automatic decomposition approach.

In priming, Feldman and colleagues have shown for both English (Feldman, O'Connor, & Moscoso del Prado Martin, 2009) and Serbian (Feldman, Kostić, Gvozdenović, O'Connor, & del Prado Martín, 2012) that semantically transparent pairs show stronger priming than opaque pairs. Järvikivi and Pyykkönen (2011) reported that when morphological family size of the prime was accounted for, priming is smaller for pseudo-complex forms compared to real inflected forms in Finnish. Similarly, in a recent English masked priming study by Lõo and Järvikivi (2019) no priming was found for pseudo-complex words when whole-word frequency of the prime was taken into account in the analysis. Milin et al. (2017) included learning-based measures (Baayen, Milin, Filipovic Durdjevic, Hendrix, & Marelli, 2011) and found comparable priming effects for pseudo-derived words (e.g., *corner*) and orthographic controls (e.g., *brothel*) with more experienced readers showing priming to a lesser extent compared to less experienced readers. Along these same lines, Andrews and Lo (2013) reported that participants with relatively high vocabulary scores showed effects of priming in the transparent condition, but no priming in the opaque condition; whereas participants whose orthography knowledge was better than their vocabulary knowledge also showed priming in the opaque condition. Finally, Medeiros and Duñabeitia (2016) conducted a masked priming lexical decision study with Spanish suffixed words and found priming effects for slow readers, but not for fast readers.

In summary, there is accumulating evidence suggesting that both semantics of the complex words (Feldman et al., 2009; Järvikivi & Pyykkönen, 2011; Lõo & Järvikivi, 2019; Milin et al., 2017) and individual differences of the participants affect morphological decomposition (Schmidtke, Van Dyke, & Kuperman, 2017; Falkauskas & Kuperman, 2015; Medeiros & Duñabeitia, 2016; Andrews & Lo, 2013).

In the present study, we will focus on the affective properties (valence, arousal, danger and usefulness ratings) of complex words. Like simplex words, complex words can also be described along different affective dimensions, for example, from very negative (e.g., *murderer*) to very positive (e.g., *puppy*); from very exciting (e.g., *panics*) to very calming (e.g., *sleeping*); from extremely dangerous (e.g., *lionness*) to not dangerous at all (e.g., *echoing*); and from extremely useful to human survival (e.g., *knives*) to not useful at all (e.g., *scorpions*).

Previous research has shown that these affective properties predict lexical processing costs. For instance, positive, calming, useful and dangerous words have been found to elicit the fastest reaction times in word recognition tasks (Kuperman, Estes, Brysbaert, & Warriner, 2014; Wurm, 2007). Kuperman (2013) reported that compound words that had more positive consituents and were also more positive as a whole were processed faster than negative and neutral compounds.

Until now, affective properties of derived and inflected words have not received much investigation, especially, in the masked priming context (see Forster, 1998 for a discussion of this method). According to the automatic morphemic decomposition view, only affective properties of the stem (e.g, *pup*) and not of the whole (inflected or derived) word (e.g., *puppy*) should influence processing costs

The current study also investigates the effects of individual differences on morphological processing by looking at several self-reported language background and personality measures of participants. The personality component will be more exploratory than the language background measures. Lõo, Toth, Karaca, and Järvikivi (2018) found that personality influenced how participants rated different types of complex words. The arousal scale of the complex words was most prominent for participants who scored high on the neuroticism scale of Big Five personality questionnaire (John & Srivastava, 1999). The present study explores whether personality effects also arise in response times of masked priming lexical decision.

In summary, the goal of the current study is two-fold. First, we will study whether automatic decomposition occurs in a large within-item study design study design when lexcal-distributional and affetive properties of the words are included in the analysis. Second, we will examining individual differences on morphological processing, by exploring participants' self-reported language background and personality measures.

Visual Masked Priming Experiment

Participants

57 native speakers of English (43 female, mean age 21 years, range 18-46) with normal or corrected-to-normal vision participated in the experiment for partial course credit.

Materials

Ninety monomorphemic English words were selected as target stimuli from the Massive Auditory Lexical Database (MALD, Tucker et al., 2018). Each target word (e.g., *pup*) was primed within-item in six conditions. The conditions were the following: identity (e.g., *pup*), inflected (e.g., *pups*), derived (e.g., *puppy*), opaque (e.g., *pupal*), stem-embedded (e.g., *pupil*), and unrelated baseline control condition (e.g., *fencing*).

Additionally, 90 nonwords and 90 real words were added to the item set as fillers. Nonword targets (e.g., *sutt*) followed the phonotactics of English and were also selected from the MALD database. Primes for nonword and real word fillers were always real English words, consisting of the same six condition types with the same proportions as for the real word targets.

Design and procedure

The prime-target pairs were counterbalanced across six lists. Each list contained 360 items. 90 experimental prime-target trials, 90 unrelated prime-target filler trials, and 180 word-prime and nonword trials. In the filler trials, prime and target pairs mimicked the six conditions in experimental list. Fillers and nonword trials were the same across lists.

The experiment was carried out using the E-Prime experimental software (Psychology Software Tools Inc.) and a SR-BOX response box. All stimuli were presented in black 32-point font Courier New letters on light gray background at the centre of the computer screen.

Each trial began with a fixation cross (+) appearing in the centre of the screen for 1000 ms, immediately followed by a forward mask (##########) for 500 ms. After that, the prime word appeared in lower case letters in the same location for 50ms. The target word appeared in the same location in upper case letters, and remained on the screen until the participant pressed the “yes” or “no” button on the response box. The participants were instructed to decide as accurately and as fast as possible whether the string of letters was an existing word in English or not. Ten practice trials preceded the experimental trials.

Prior to the main task, participants were asked to fill out a language background questionnaire, where they were asked to reflect on their English language skills and reading habits. For instance, they were asked how often they read scholarly or fictional literature; how they estimate their English vocabulary size, and how fast they consider themselves as readers.

They were also asked to fill out a 60-item HEXACO personality inventory questionnaire (Ashton & Lee, 2009), which provided for each participant a separate score on each of the six personality scales: honesty, emotionality, extroversion, agreeableness, conscientiousness, and openness to experience. The whole procedure (questionnaires and lexical decision task) took approximately 60 minutes to complete.

Analysis and Results

Prior to the analysis, practice trials, nonword trials and fillers were removed from the dataset. Trials with response times more than 1600 ms (1.1% of the data) as well as trials with incorrect responses (6.2% of the data) were removed.

Frequencies for the primes and targets were determined using the Corpus of Contemporary American English (COCA, Davies, 2010). Whole-word frequency (i.e., the token frequency of *pups*, *pups* or *puppy*) was used for the analysis. Frequency was log-transformed prior the analysis to reduce the skewness of the distribution.

Affective ratings of valence, arousal, danger and usefulness for each target and prime were collected during a separate rating experiment (see Lõo et al., 2018). In total, 181 native speakers of English rated the experimental items of the current study on an nine-point Likert-scale either on valence, arousal, usefulness or danger scale (1- sad/not exciting/not useful/not dangerous; 9 - happy/exciting/extremely useful/extremely dangerous). Participants in the rating experiment were different from the participants in the current experiment. A rating score for each target and prime word was calculated by taking the average score for each word across all participants.

The statistical analysis was conducted using Generalized Additive Mixed Models (GAMM, Wood, 2006; the R-package *mgcv*). For visualization, we made use of the R-package *itsadug* (van Rij, Baayen, Wieling, & van Rijn, 2016). We opted for GAMM analysis, because it does not assume linearity between the predictor and response variables.

The response variable of interest was the reaction time of masked priming lexical decision in milliseconds. We opted to use raw reaction times because they followed a normal distribution. However, an analysis with the log-transformed reaction times produced the same results. The main predictors were the condition (identity - M, inflected - I, derived - D, pseudo-complex - PC, stem-embedded - SE, baseline - BL), as well as the log-transformed frequency and affective ratings (valence, arousal, danger and usefulness) of the prime and target words. Additionally, we were interested in the effects of individual differences measures, we investigated whether self-reported language knowledge and reading habits, as well as personality had an effect on reaction times.

The output of the final GAMM-model is presented in Table 1. The parametric part shows that participants were significantly faster in identity (t=-4.90, $p<0.00001$), inflected (t=-4.10, $p<0.00001$) and derived (t=-3.43, p=0.006) conditions, whereas the two other conditions (pseudo-complex and stem-embedded condition) were not significant compared to the baseline condition.

Further, participants' openness to new experience and scholarly reading frequency affected reaction times. Reaction times decreased linearly for the participants who scored higher on the openness to experience scale compared to participants who scored lower on the same scale (t=-2.23, p=0.003). In return, reaction times were slower for participants who read more scholarly articles compared to participants who read fewer scholarly articles (t=3.48, p=0.015). There was neither a significant interaction between the condition and the openness score, nor between the condition and the scholarly reading score. Other self-reported language and personality scores were not significant in the final model.

The first three lines of the non-parametric part of the model output show nonlinear interactions between the prime and the target frequency, between the prime and target arousal score as well as between the prime and target usefulness score. These effects are visualized in Figure 1. The yellow color at the bottom left corner of the left panel shows that the reaction times were the slowest when both the prime and target were low-frequency words. However, the interaction between the target and prime frequency seems to disappear when target and prime frequencies increase. This is indicated by the blue color and wider contour lines at the top right corner of the left panel in Figure 1.

The nonlinear interaction between the prime and target arousal score is represented in the middle panel of Figure 1. Reaction times were the fastest when the target word scored high on the arousal scale and the prime word scored low on the arousal scale as indicated by the blue color at the bottom

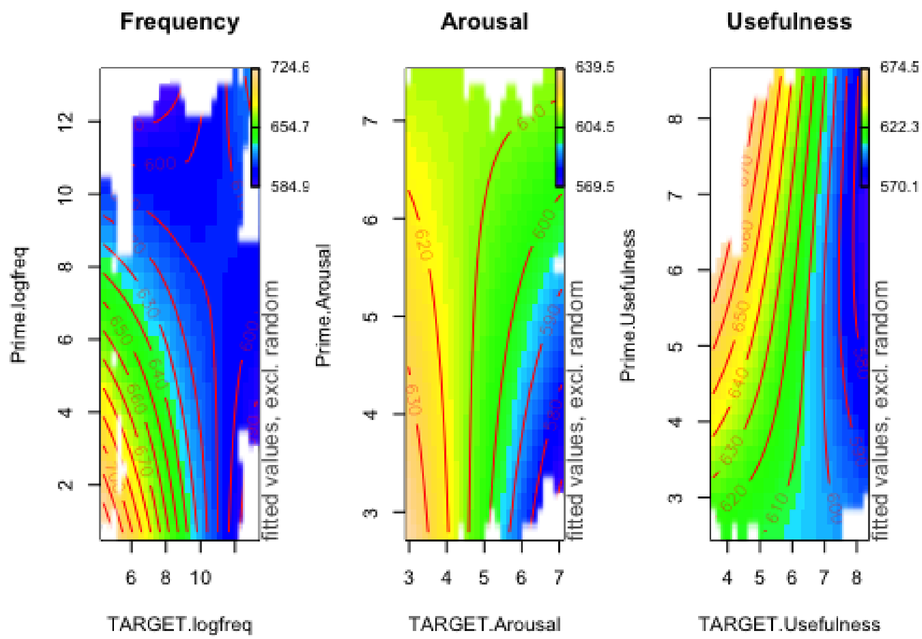

Figure 1: Tensor product smooth for the interaction of prime and target word frequency, arousal and usefulness. Color coding is used to represent model predictions, with yellow indicating slower reaction times, and blue representing faster reaction times

right corner of Figure 1.

Finally, the interaction between the prime and target word usefulness is presented in the right panel of Figure 1. The reaction times were the slowest when the prime was rated as very useful but the target was not, as indicated by the yellow color at the top left corner. Neither the target nor the prime valence and danger scores were significant. Frequency, arousal and usefulness scores did not interact with the condition.

The non-parametric part of the model output also included by-target random intercepts and by-participants random smooths for trial to account for the random variability between the items and participants.

In summary, the GAMM-analysis showed significant priming effects for the identity, derived and inflected conditions, but no priming for the pseudo-complex or stem-embedded conditions. There was a significant interaction between prime and target frequency, arousal and usefulness scores, but this did not interact with the condition. Finally, we found significant effects of participants' openness and scholarly reading, however, these effects did not interact with the condition.

Discussion and Conclusion

The goal of the current study was to investigate English morphological processing using masked priming. Some studies have reported that words like *talker* and *corner* are at least initially processed similarly (Rastle et al., 2004), while others claim that this is not the case, in particular, when various lexical-distributional properties are taken into account (Järvikivi & Pyykkönen, 2011; Lõo & Järvikivi, 2019; Milin et al., 2017), as well as individual differences between participants (Schmidtke, Van Dyke, & Kuperman, 2017; Falkauskas & Kuperman, 2015; Medeiros & Duñabeitia, 2016; Andrews & Lo, 2013).

In line with the latter view, the present study reports priming effects for words with an existing morphological relationship (e.g., *cats*, *puppy*), but no effects of priming for pseudo-complex words (e.g., *corner*). In fact, the processing of pseudo-complex words did not differ at all from either the stem-embedded condition (e.g., *turnip*) or the unrelated baseline condition. This supports findings from another recent English priming study by Lõo and Järvikivi (2019), where there were also no priming effects for pseudo-complex condition, using different materials. Additionally, we showed that the semantics of the complex words plays an important role early on. Like in Lõo and Järvikivi (2019), frequency of the complex word predicted processing costs; however, there were no significant differences between pseudo-complex and truly morphologically complex words in this respect.

Further, we investigated how affective ratings of complex words affect morphological processing. In line with the previous research on compound processing (Kuperman, 2013), we found effects of affective ratings for inflected and derived words. The prime-target ratio of affective ratings influenced the response times in masked priming, further challenging the blind decomposition approach, where the properties of the prime should not have an effect.

Interestingly, out of the four ratings scales (valence, arousal, usefulness and danger), only arousal and useful-

Table 1: Summary of the partial effects in GAMM fitted to masked priming lexical decision reaction times in milliseconds.

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	694.18	86.89	7.99	< 0.0001
conditionD	-31.94	9.31	-3.43	0.0006
conditionI	-35.09	8.56	-4.10	< 0.0001
conditionM	-42.82	8.73	-4.90	< 0.0001
conditionPC	-2.17	8.19	-0.27	0.79
conditionSE	6.25	8.10	0.77	0.44
open.hexaco	-49.86	22.35	-2.23	0.03
Scholarly.Reading	32.81	13.90	2.36	0.02
B. smooth terms	edf	Ref.df	F-value	p-value
te(TARGET.logfreq,Prime.logfreq)	3.80	4.22	3.93	0.003
te(TARGET.Arousal,Prime.Arousal)	3.03	3.05	3.48	0.015
te(TARGET.Usefulness,Prime.Usefulness)	3.68	4.16	3.74	0.004
s(Subject,Trial)	184.49	494.00	4.29	< 0.0001
s(TARGET)	47.37	86.00	1.24	< 0.0001

ness target-prime ratios had an effect. Kuperman (2013) reported valence but not arousal effects in compound processing. However, their study used a standard lexical decision task, whereas the current study used masked priming lexical decision, tapping into earlier processing than the standard lexical decision. Arousal and usefulness ratings may be tapping into the internal state of the individual, thus are more subconscious; whereas, valence ratings may require a more conscious thought, and thus get activated later in time than can be captured by a masked priming study.

In general, the effects of affective properties were not that strong in the current study, and there may be different reasons for this. First, in the current study, the derived and inflected primes (e.g., *puppy*, *pups*) have similar meanings to the target (e.g., *pup*), so the affective polarities may have been very similar (for example, the word *puppy* was as happy, exciting, useful and dangerous as the word *pup*). Second, as the design of the current study did not explicitly control for the emotional affectivess of the stimuli, most of the stimuli were neither very positive nor very negative, neither very useful nor very useless, so there may not have been enough variation between the stimuli.

Finally, the current study also focused on the effects of individual differences in morphological processing. Interestingly, they were again the same for truly complex and pseudo-complex words. From the five personality measures (honesty, emotionality, extroversion, conscientiousness, openness to experience), only participants' openness to new experience had an effect on reaction times. More open participants were faster than less open participants. Participants who are more open to experience in general might be also more open to tasks such as a lexical decision task. From the language background measures (self-reported vocabulary knowledge, reading speed, scholarly reading and fictional reading frequency), only scholarly reading had an effect on reaction times. Participants who read more scholarly literature were slower than participants who read less scholarly literature. This is in line with the research showing that more experience with language slows one down in various language tasks (Ramscar, Hendrix, Shaoul, Milin, & Baayen, 2014).

Also this is important to note, however, that both the topic of affective properties and individual differences in morphological processing are relatively new and thus, our findings require further research.

To conclude, the processing of complex words, even in languages with a relative simple morphology, such as English, seems to be much more complex than just a matter of morphemic decomposition. The current study complements this idea by showing that pseudo-complex and morphologically complex words are indeed processed differently. We also showed that both affective properties and individual differences influence English morphological processing; however, the precise nature of these effects requires further research.

References

Andrews, S., & Lo, S. (2013). Is morphological priming stronger for transparent than opaque words? it depends on individual differences in spelling and vocabulary. *Journal of Memory and Language*, *68*(3), 279–296.

Ashton, M. C., & Lee, K. (2009). The HEXACO–60: A short measure of the major dimensions of personality. *Journal of personality assessment*, *91*(4), 340–345.

Baayen, R. H., Milin, P., Filipovic Durdjevic, D., Hendrix, P., & Marelli, M. (2011). An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Psychological Review*, *118*(3), 438-481.

Baayen, R. H., Wurm, L. H., & Aycock, J. (2007). Lexical dynamics for low-frequency complex words. a regression study across tasks and modalities. *The Mental Lexicon*, *2*, 419–463.

Beyersmann, E., Ziegler, J. C., Castles, A., Coltheart, M., Kezilas, Y., & Grainger, J. (2016). Morpho-orthographic segmentation without semantics. *Psychonomic Bulletin & Review*, *23*(2), 533–539.

Davies, M. (2010). The corpus of contemporary american english as the first reliable monitor corpus of english. *Literary and linguistic computing*, *25*(4), 447–464.

Falkauskas, K., & Kuperman, V. (2015). When experience meets language statistics: Individual variability in processing English compound words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(6), 1607.

Feldman, L. B., Kostić, A., Gvozdenović, V., O'Connor, P. A., & del Prado Martín, F. M. (2012). Semantic similarity influences early morphological priming in serbian: A challenge to form-then-meaning accounts of word recognition. *Psychonomic bulletin & review*, *19*(4), 668–676.

Feldman, L. B., O'Connor, P. A., & Moscoso del Prado Martin, F. (2009). Early morphological processing is morpho-semantic and not simply morpho-orthographic: evidence from the masked priming paradigm. *Psychonomic Bulletin & Review*, *16*(4), 684–691.

Forster, K. (1998). The pros and cons of masked priming. *Journal of Psycholinguistic Research*, *27*(2), 203–233.

Järvikivi, J., & Pyykkönen, P. (2011). Sub-and supralexical information in early phases of lexical access. *Frontiers in Psychology*, *2*.

John, O. P., & Srivastava, S. (1999). The big five trait taxonomy: History, measurement, and theoretical perspectives. *Handbook of personality: Theory and research*, *2*(1999), 102–138.

Kuperman, V. (2013). Accentuate the positive: Semantic access in english compounds. *Frontiers in psychology*, *4*.

Kuperman, V., Estes, Z., Brysbaert, M., & Warriner, A. B. (2014). Emotion and language: valence and arousal affect word recognition. *Journal of Experimental Psychology: General*, *143*(3), 1065.

Lázaro, M., Illera, V., & Sainz, J. (2016). The suffix priming effect: Further evidence for an early morpho-orthographic segmentation process independent of its semantic content. *The Quarterly Journal of Experimental Psychology*, *69*(1), 197–208.

Longtin, C., Segui, J., & Hallé, P. (2003). Morphological priming without morphological relationship. *Language and Cognitive Processes*, *in press*, 0.

Lõo, K., & Järvikivi, J. (2019). Whole-word frequency effects in English masked priming: very little CORN in CORNER and CORNET. *Manuscript submitted for publication*.

Lõo, K., Järvikivi, J., & Baayen, R. H. (2018). Whole-word frequency and inflectional paradigm size facilitate Estonian case-inflected noun processing. *Cognition*, *175*, 20–25.

Lõo, K., Järvikivi, J., Tomaschek, F., Tucker, B. V., & Baayen, R. H. (2018). Production of estonian case-inflected nouns shows whole-word frequency and paradigmatic effects. *Morphology*, *28*, 71–97.

Lõo, K., Toth, A., Karaca, F., & Järvikivi, J. (2018). Effects of word stem and personality in affective ratings for complex words. *Paper presented at the 11th International Conference on the Mental Lexicon in Edmonton, Alberta, September 25-28, 2018*.

Manelis, L., & Tharp, D. A. (1977). The processing of affixed words. *Memory and Cognition*, *5*, 690-695.

Marslen-Wilson, W. D., Bozic, M., & Randall, B. (2008). Early decomposition in visual word recognition: Dissociating morphology, form, and meaning. *Language and Cognitive Processes*, *23*(3), 394–421.

Medeiros, J., & Duñabeitia, J. A. (2016). Not everybody sees the ness in the darkness: Individual differences in masked suffix priming. *Frontiers in psychology*, *7*, 1585.

Milin, P., Feldman, L. B., Ramscar, M., Hendrix, P., & Baayen, R. H. (2017). Discrimination in lexical decision. *PloS one*, *12*(2), e0171935.

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

Moscoso del Prado Martín, F., Bertram, R., Häikiö, T., Schreuder, R., & Baayen, R. H. (2004). Morphological family size in a morphologically rich language: The case of Finnish compared to Dutch and Hebrew. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *30*, 1271–1278.

Ramscar, M., Hendrix, P., Shaoul, C., Milin, P., & Baayen, R. H. (2014). The myth of cognitive decline: Non-linear dynamics of lifelong learning. *Topics in cognitive science*, *6*(1), 5–42.

Rastle, K., & Davis, M. H. (2008). Morphological decomposition based on the analysis of orthography. *Language and Cognitive Processes*, *23*(7-8), 942–971.

Rastle, K., Davis, M. H., Marslen-Wilson, W. D., & Tyler, L. K. (2000). Morphological and semantic effects in visual word recognition: A time-course study. *Language and Cognitive Processes*, *15*(4-5), 507–537.

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*, 1090–1098.

Schmidtke, D., Matsuki, K., & Kuperman, V. (2017). Surviving blind decomposition: A distributional analysis of the time-course of complex word recognition. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *43*(11), 1793-1820.

Schmidtke, D., Van Dyke, J. A., & Kuperman, V. (2017). Individual variability in the semantic processing of English compound words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, in press.

Taft, M., & Forster, K. I. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal*

Behavior, *14*, 638-647.
Tucker, B. V., Brenner, D., Danielson, D. K., Kelley, M. C., Nenadić, F., & Sims, M. (2018). The massive auditory lexical decision (mald) database. *Behavior research methods*, 1–18.
van Rij, J., Baayen, R. H., Wieling, M., & van Rijn, H. (2016). *itsadug: Interpreting time series, autocorrelated data using GAMMs.* (R package version 2.2)
Wood, S. N. (2006). *Generalized additive models*. New York: Chapman & Hall/CRC.
Wurm, L. H. (2007). Danger and usefulness: An alternative framework for understanding rapid evaluation effects in perception? *Psychonomic Bulletin & Review*, *14*, 1218–1225.