

# Morphological processing is gradient not discrete in L1 and L2 English masked priming

Kaidi Lõo, Abigail Toth, Figen Karaca, and Juhani Järvikivi  
University of Tartu | University of Groningen | Radboud University |  
University of Alberta

In recent years, evidence has emerged that readers may have access to the meaning of complex words even in the early stages of processing, suggesting that phenomena previously attributed to morphological decomposition may actually emerge from an interplay between formal and semantic effects. The present study adds to this line of work by deploying a forward masked priming experiment with both L1 (Experiment 1) and L2 (Experiment 2) speakers of English. Following recent research trends, we view morphological processing as a gradient process emerging over time. In order to model this, we used a large within-item stimulus design combined with advanced statistical methods such as generalised mixed models (GAMM) and quantile regression (QGAM). L1 GAMM analyses only showed priming for true morpho-semantic relations (the identity ‘bull’, inflected ‘bulls’ and derived conditions ‘bullish’), with no priming observed in the case of other relations (the pseudo-complex ‘bully’ or the stem-embedded ‘bullet’ conditions). Furthermore, with respect to the time-course of effects, we found significant differences between conditions were present from very early on as revealed by the QGAM analyses. In contrast, L2 speakers showed significant facilitation across all five conditions compared to the baseline condition, including the stem-embedded condition, suggesting early L2 processing is only dependant on the form.

**Keywords:** morphological processing, masked priming, L1 processing, L2 processing, quantile regression

## 1. Introduction

It is not exactly clear how a word’s internal structure influences the way it gets processed when we read. Various theoretical accounts have been proposed in

order to explain how morphology happens in the mind for both native (L1; Baayen, Chuang, Shafaei-Bajestan & Blevins 2019; Baayen, Milin, Filipovic Durdjevic, Hendrix & Marelli 2011; Feldman, O'Connor & Moscoso del Prado Martin 2009; Giraudo & Grainger 2001; Marelli & Baroni 2015; Rastle & Davis 2008; Rastle, Davis & New 2004; Schmidtke, Matsuki & Kuperman 2017; Schreuder & Baayen 1995; Taft 2004; Taft & Forster 1975) and non-native speakers (L2; Clahsen & Felser 2006; Clahsen, Felser, Neubauer, Sato & Silva 2010; Jacob, Heyer & Veríssimo 2018; Lemhöfer, Dijkstra, Schriefers, Baayen, Grainger & Zwitserlood 2008; McDonald 2006). Although research on this topic dates back 50 years (Manelis & Tharp, 1977; Taft & Forster, 1975), recent novel methodological approaches are promising in providing new insights (Amenta, Crepaldi & Marelli, 2020; Günther, Petilli & Marelli, 2020; Hendrix & Sun, 2020; Schmidtke et al., 2017; Ulicheva, Harvey, Aronoff & Rastle, 2020).

The forward masked priming paradigm (Forster & Davis, 1984) has traditionally been the most influential tool in the investigation of the how and when language users make use of morphological structure. In this task, a prime is presented for a very brief period of time (usually between 30–50 ms) in between a forward mask (e.g., #####) and the target itself (i.e., the backward mask). It is assumed that the forward mask prevents participants from becoming aware of the prime word (but see Norris & Kinoshita 2008 and Tzur & Frost 2007 for a discussion) and as such is seen as a way to tap into early, non-conscious, cognitive processes.

The current study applied this paradigm to the morphological processing of L1 and L2 English. Unlike most previous studies, we used six different experimental conditions in a large scale within-item design. We also conducted two analyses for each experiment; the first using generalized additive mixed models and the second using quantile generalized additive models. The quantile generalized additive model was adopted to explore the time-course of effects across the different experimental conditions. Before introducing the details of the current study, we will first review the relevant literature on L1 and L2 morphological processing.

### 1.1 Evidence from L1 processing

Studies using the masked priming paradigm with native (L1) speakers often report that words with real morphological structure, such as *talker* are primed similarly to words with only apparent structure (i.e., a stem + an orthographically possible affix), such as *corner*, whereas words without morphological structure (i.e., a stem + an orthographically impossible affix), such as *turnip* (where *-ip* is not a morpheme in English), do not show priming effects (Beyersmann, Ziegler,

Castles, Coltheart, Kezilas & Grainger, 2016; Kazanina, 2011; Longtin, Segui & Hallé, 2003; Rastle et al., 2004). These findings have led to the hypothesis that processing is initially morpho-orthographic and semantically blind. That is, for all (written) words, morphologically motivated segmentation is attempted automatically based on the form of the (letter) string alone, regardless of whether or not the word in question is truly complex. However, there is another line of research suggesting that the meaning (i.e., the morpho-semantics) of a word may also be accessed early on. In fact, when the effect sizes between words with real versus pseudo morphology have been compared, the former have consistently been found to be larger (e.g., Feldman et al. 2009; Jared, Jouravlev & Joanisse 2017; Järvikivi & Pyykkönen 2011; Morris, Frank, Grainger & Holcomb 2007) and in some cases there was no priming for the pseudo-complex condition (Lõo & Järvikivi, 2019).

Furthermore, if early processing were semantically blind then there should not be differences between different types of real morphological structures, for example between inflected and derived forms. Notably, the majority of studies have not explicitly tested for such differences. This might be surprising given that the semantic relationship between a derived word and its base is typically much more indirect than between an inflected word and its stem, as derivation often changes the meaning of the prime in more idiosyncratic ways. In an unmasked English priming study, Raveh (2002) directly compared inflected and derived words and found a small priming advantage for the inflected condition compared to the derived condition when the prime duration was 50 ms and a significant advantage at 150 and 250 ms prime durations. Furthermore, Feldman (1994) found differences between inflectional and derivational priming in Serbian. These findings suggest that effects, not only between real and pseudo-morphological structure, but also within the real morphological structure may be gradient. However, evidence for differences between different types of real morphologically complex words is more limited than between complex and pseudo-complex words.

Building upon this idea, a recent study by Jared et al. (2017) used masked priming with event-related potentials (ERPs) to compare the processing between semantically transparent (e.g., *foolish*), quasi-transparent (e.g., *bookish*), opaque (e.g., *vanish*) and orthographic (e.g., *bucket*) words, across five lexical and semantic decision experiments. In all the experiments, they observed the largest priming effect for semantically transparent words, followed by quasi-transparent words, and the smallest effect for opaque words, as reflected by the response times. This suggests that morphological processing is more gradient than has been concluded based on studies with limited conditions. Furthermore, the ERP

measurements from the same set of experiments showed semantic transparency effects to be observed as early as 200–250 ms.

Semantic properties, such as compositionality, which itself is a gradient phenomenon (see e.g., Bell & Schäfer, 2016) have also been shown to influence early processing of compounds (e.g., Amenta et al. 2020; Davis, Libben & Segalowitz 2019; Kuperman, Schreuder, Bertram & Baayen 2009; Schmidtke & Kuperman 2019). For example, Davis et al. (2019) showed significantly larger P100 ERPs for compounds with two opaque constituents (e.g., *deadline*) compared to compounds containing transparent constituents (e.g., *grapeseed*). This suggests that the whole compound word plays a role in early word recognition, as early as 100 ms. This goes against what is predicted by a blind decomposition account, given that if constituents are accessed before whole words, it would need to happen within 100 ms, a proposition for which we currently lack evidence.

Early semantic effects in morphological processing is contrary to much of the earlier neuro-psychological morphological processing research, which usually report early morphological decomposition effects and late semantic effects, i.e., semantically blind decomposition (see Leminen, Smolka, Dunabeitia & Pliatsikas, 2019, for an overview). However, recently, this line of research has been further challenged by ERP and eye movement co-registration studies, which show very early semantic effects in eye-movements but interestingly not in the ERP measurements (Schmidtke & Kuperman, 2019). These findings call into question the initial conclusion that late semantic effects in the ERP signal for morphologically complex words necessarily reflect the start of semantic influences. Instead, the findings suggest that such ERP effects perhaps reflect a later semantic integration phase. Thus, the discussion on the timeline of morphological processing seems to be far from resolved.

Finally, recent evidence with nonwords (Beyersmann et al., 2016; Hasenäcker, Beyersmann & Schroeder, 2016; Heathcote, Nation, Castles & Beyersmann, 2018) suggests that even when stems are not followed or preceded by a real affix, such as the nonword *cornil*, they can still show early priming effects. This growing line of research suggests that bottom-up information of the stem form, independent of whether or not there is real morphological structure, can influence word processing early on (see also Morris et al. 2007 for similar evidence in words). Going back to Jared et al. (2017), this suggests that anytime there is no top-down semantic influence, as is the case when the prime is a nonword, or when the top-down semantic information is less disrupting, as in the case of transparent primes, the more likely we are to see facilitation.

In sum, although discrete morphemic decomposition has been often seen as the first building block of the word recognition process, recent evidence on the early influence of a word's semantics points to a possibly more complicated

picture of L1 processing. In particular, it seems that a view whereby morphological structure is the only cue in early processing is problematic. Therefore, a more gradient view of morphology, whereby different types of (sub)lexical information emerge simultaneously, may be a more fruitful approach to this topic. This is also in line with research from semantic similarity judgements, which suggest that morphological complexity lies on a continuum and that there is no clear separation between semantically transparent and semantically opaque words (Gonnerman, Seidenberg & Andersen, 2007). Recent evidence suggests that priming experiments with more conditions and items can also help to create a more natural continuum and automatically bring out more subtle patterns in morphological processing. Furthermore, it is important to find ways to appropriately assess the time course of processing, as differences between categories can arise earlier or later during processing, which cannot always be captured by a traditional mean reaction time analysis.

## 1.2 Evidence from L2 processing

In the case of non-native (L2) morphological processing, the focus has largely been on determining whether L2 speakers use the same processing mechanisms as L1 speakers. In general, recent research indicates that L2 speakers depend more on form properties than L1 speakers. For example, Silva & Clahsen (2008) compared derivation with identity priming and found that the two were equally strong for L1 speakers, whereas for L2 speakers, identity primes showed stronger facilitation compared to derived primes. The same was shown for inflection, such that inflected and identity priming were comparable for L1 speakers, whereas for L2 speakers there was no inflectional priming (but see Voga, Anastassiadis-Symeonidis & Giraudo 2014). More recently, using a within item design, Jacob et al. (2018) found that while L1 speakers showed similar priming for derivation and inflection, L2 speakers only showed priming effects for the derived, but not for the inflected words. However, it should be noted that the derived words were exclusively suffixed (e.g., *Bohrung* ‘drilling’) and the inflected words were prefixed (e.g., *gebohrt* ‘drilled’), which may have influenced the results, such that L2 speakers may be particularly sensitive to form similarity at the word onset.

Further research by Clahsen and colleagues (Heyer & Clahsen, 2015) found that L2 speakers processed words with embedded stems (e.g., *scandal*) and transparent derived words (e.g., *scanner*) similarly, showing facilitation for both, whereas L1 speakers did not show facilitation for embedded stem words. This was also taken to suggest that compared to native speakers, non-native speakers rely more on form-properties. However, other studies have been able to show early morpho-semantic effects also in the case of L2 speakers. For example,

Diependaele, Duñabeitia, Morris & Keuleers (2011) showed larger facilitation for real morphology for both L1 and L2 speakers and less facilitation for pseudo-complex and stem-embedded words with Spanish- and Dutch-English bilinguals. In line with this, a recent masked priming study, with L1 Italian-L2 English speakers, Viviani & Crepaldi (2019) reported statistically larger facilitation for transparent versus opaque words in both L1 and L2 speakers. However, priming in the stem-embedded condition was present only in the case of L2 speakers. Crucially, the authors showed that genuine morphological facilitation only arose as proficiency increased, as both opaque and orthographic priming shrunk as L2 competence increased (see also Coughlin & Tremblay 2015). This suggests that L2 speakers rely increasingly less on the form as their proficiency increases. One possible explanation is that, initially, L2 speakers need to access L2 semantics via their L1, but later they may develop a more direct link between L2 forms to L2 concepts (see e.g., Jiang, 2000).

Recent research by Mulder and colleagues suggests that like the L1 lexicon, the L2 lexicon may be sensitive to semantic information of complex words beyond transparency. Mulder and colleagues (Mulder, Dijkstra & Baayen, 2015; Mulder, Dijkstra, Schreuder & Baayen, 2014) studied Dutch morphological processing with Dutch-English bilinguals and showed that bilinguals are affected by the morphological family size (i.e., the number of derivations and compounds where a word occurs as a constituent) of cognate words in both languages.

Given the discrepant findings regarding morpho-semantic influences in the case of L1, it is unsurprising that the same goes in the case of L2. However, the story is further complicated by differences in L1 and L2 typologies, as well as differences in L2 language proficiency across the different studies. Despite the disagreement about whether morpho-semantics influences L2 morphological processing, there is a consensus that orthographic priming plays a strong role. There is less consensus about the role of orthography in L1 priming because the top-down semantic information is more automatic for L1 speakers than for L2 speakers. Compared to L1 morphological processing, L2 processing has received a lot less attention, especially when it comes to studies using rich stimulus sets.

### 1.3 The current study

Discrepancies in previous research can partly be attributed to differences in statistical methods. However, they are also in-part driven by the assumptions and predictions of the adopted framework through which the results are interpreted (see e.g., Baayen et al. 2011; Feldman et al. 2009; Marelli & Baroni 2015). Another reason for the discrepant findings may be that the experimental conditions (and the exact items within those conditions) are not the same across

studies. More complex item sets naturally lead to more complex patterns of results (see also Feldman & Basnight-Brown 2008 for a discussion on item list composition). Thus, we are currently still unsure how exactly L1 and L2 speakers process morphologically complex words and how processing may differ between the two populations. In what follows, we present a study which tests the different accounts.

The present study deployed the same forward masked priming experiment with both L1 (Experiment 1) and L2 (Experiment 2) speakers of English. We used a full within-item design in order to investigate how various types of morphological structure influence language processing, as well as to better control for the effects of individual target word properties. The critical manipulation was the type of prime used for the processing of a monomorphemic target word (*bull*): identity (ide, *bull*); inflected (inf, *bulls*), derived (der, *bullish*), pseudo-complex (pc, *bully*), stem-embedded (se, *bullet*), and unrelated baseline (*disks*). Thus, compared to the majority of previous research, the present study has more conditions, as well as a within-item design.

Given the previous L1 research, several outcomes are possible. According to the semantically blind decomposition approach (Beyersmann et al., 2016; Kazanina, 2011; Longtin et al., 2003; Rastle et al., 2004), the expected outcome would be similar facilitatory priming effects for the inflected, derived and pseudo-complex conditions compared to the baseline condition, whereas the stem-embedded condition would be expected to show no priming compared to the baseline condition. However, if morpho-semantics has an influence on processing, as predicted by e.g., Feldman et al. 2009; Jared et al. 2017; Järvikivi & Pyykkönen 2011; Morris et al. 2007, the expected outcome would be that the inflected condition should show stronger facilitation than the derived condition, as there is a stronger semantic relationship between the prime and the target. However, compared to the inflected and derived conditions, there should be weaker or no priming at all for the pseudo-complex condition. As for the stem-embedded condition, one also does not expect priming compared to the baseline condition according to this view.

In the case of L2 processing, the main question is whether L2 speakers show similar patterns to L1 speakers. According to Clahsen and colleagues (Heyer & Clahsen, 2015; Silva & Clahsen, 2008), stem-embedded, pseudo-complex, derived and inflected conditions should all show similar facilitation, as form is the primary driving force of processing. However, if semantic relatedness also plays a role in L2 processing (e.g., Diependaele et al. 2011; Viviani & Crepaldi 2019; Voga et al. 2014), differences between the inflected and derived conditions, as well as between the real morphological (inflected and derived) and pseudo-morphological conditions might arise.

It should be noted that an identity prime will serve as a control condition across our analyses. As the identity prime is identical to the target in both form and meaning, we expect the largest priming effect of this condition for both L1 and L2 speakers.

In order to investigate the time-course of when the different conditions exert their influence on processing, we performed a distributional analysis on the reaction times using quantile non-parametric additive models (Fasiolo, Wood, Zaffran, Nedellec & Goude, 2020). Quantile regression is a statistical method for predicting the deciles of a response variable, whereas standard regression methods consider only the mean of the response. With respect to the current study, this technique can help us reveal when exactly the differences between the experimental conditions emerge. Quantile regression allows us to investigate the extent to which overlap in form and/or meaning influences the earliest processing, by looking into specific deciles capturing only very fast or very slow reaction times. This technique has also been previously used by other morphological processing studies. For example, Lõo, Järvikivi, Tomaschek, Tucker & Baayen (2018) used this technique to study word naming of inflected noun processing in Estonian and Baayen & Smolka (2020) used it to study German verb processing with overt priming. Other related methods of distributional reaction time analyses also highlight the importance of looking at the time-course in morphological processing (Andrews & Lo, 2013; Hasenäcker et al., 2016; Schmidtke et al., 2017).

Accordingly, differences driven by morpho-semantics may emerge in the current quantile regression analysis between apparent and real morphology. According to the morpho-semantic account, the differences between the transparent conditions (inflected and derived) and pseudo-complex conditions should be present in the earliest deciles, as well as between the derived and inflected conditions. However, according to the semantically blind decomposition account, there should be no differences between morphologically decomposable words, especially not in the early deciles. As most previous research suggests that L2 speakers are more driven by form-properties than L1 speakers, we also expect less morpho-semantic influences in the quantile regression; they may however, arise in the later deciles.

Next, we will introduce the details of Experiment 1 with native speakers of English.



## 2. Experiment 1: L1 English speakers

### 2.1 Materials and methods

#### 2.1.1 Participants

Ninety native speakers of English (68 female, mean age 20 years, range 17–46) with normal or corrected-to-normal vision participated in the experiment. All participants were undergraduate students from the University of Alberta and received partial course credit for their participation. The study received ethics approval from the University of Alberta Research Ethics Board. Informed written consent was obtained prior to participation.

#### 2.1.2 Materials

Ninety monomorphemic English words were selected as target stimuli from the Massive Auditory Lexical Decision Database (MALD, Tucker, Brenner, Danielson, Kelley, Nenadić & Sims 2019). Each target word (e.g., *bull*) was primed within-item across six conditions: identity (ide, e.g., *bull*), inflected (inf, e.g., *bulls*), derived (der, e.g., *bullish*), pseudo-complex (pc, e.g., *bully*), stem-embedded (se, e.g., *bullet*), and unrelated baseline (e.g., *disks*). As baseline words, an equal number of unrelated monomorphemic, inflected, derived, pseudo-complex and stem-embedded words were selected from the MALD. Regarding the affix type, primes were either (pseudo-)suffixed (e.g., *talker*, *corner* 381 stimuli) or pseudo-(prefixed) (e.g., *dispose*, *misadd*, *exact*, 54 stimuli) or did not have an affix (105 identity primes e.g., *bull* and some of the baseline primes e.g., *gem*). An additional 90 real word and 180 nonword targets were selected from MALD (Tucker et al., 2019) and added to the item set as fillers. All nonword targets (e.g., *sutt*) followed the phonotactics of English. Primes for both real word and nonword fillers were always real English words and fell into one of the six conditions mentioned above, matching the experimental trials. The prime-target pairs were counterbalanced across six lists. Each list contained 360 items: 90 experimental prime-target trials, 90 real word and 180 nonword filler trials. All filler trials were the same across the six lists.

With respect to stimuli selection, we were unable to control for the lexical properties of the primes between the conditions given the within-item design and multiple conditions. However, these properties were taken into account during the statistical analyses. The average length (in characters) of the primes for each condition were as follows: identity 3.5, inflected 5.1, derived 6.1, pseudo-complex 5.5, stem-embedded 5.5, baseline 5.5. The average log frequency of the primes for each condition were as follows: identity 9.0, inflected 7.4, derived 4.8, pseudo-

complex 7.1, stem-embedded 6.7, baseline 7.6. The complete stimulus set is available on Open Science Framework (<https://osf.io/b5vst/>).

### 2.1.3 Procedure and design

The experiment was carried out with E-Prime experimental software using an SRBOX response box (Psychology Software Tools Inc.). All stimuli were presented in black 32-point Courier New font on a light gray background at the centre of a 21-inch LCD monitor.

Participants were tested individually and sat approximately 60 cm in front of the computer screen. All participants were randomly assigned to one of the six experimental lists and instructed to decide whether the string of letters on the screen was a real English word, as quickly and accurately as possible.

Each trial began with a fixation cross (+) appearing in the centre of the screen for 1000 ms. A forward mask (#####) was then presented for 500 ms, followed by the prime word appearing in lower-case letters in the same location for 50 ms. The prime was then immediately replaced by the target word, which appeared in upper-case letters and remained on the screen until the participant either pressed “yes” or “no” on the response box. Ten practice trials preceded the experimental trials.

## 2.2 Analysis

Prior to data analysis, practice trials and fillers were removed from the data set. Three participants and one item (the target word *CAD*) were removed from the data set as their error rates were 50% or higher. Finally, trials with incorrect responses (9.4% of the data) were removed from the data set. There were no significant differences in error rates between conditions.

Statistical analyses were conducted in two parts. First, we ran a standard regression analysis using generalized additive mixed-effects models (GAMM, Wood 2017; the R-package *mgcv*), which provided predictions based on mean reaction times. Second, we conducted a quantile analysis of the reaction times using quantile generalized additive mixed modelling (QGAM, Fasiolo et al. 2020, the R-package *qgam*).

GAMM is an extension to mixed-effects regression methods (Baayen, Davidson & Bates, 2008) that allows for non-linear modeling of predictor terms. This is achieved via smooth functions (Wood, 2017), which allow the regression line to ‘wiggle’ when warranted by the data. In addition to random intercept and random slope adjustments, GAMM also models random smooths, which adjust the trend of a predictor in a nonlinear way (naturally incorporating random intercepts and slopes). Prior to the GAMM analysis, trials with response times above

or below 2.5 standard deviations of the grand mean were identified as outliers and removed (1.8% of the data).

The final model was obtained using a backward model fitting approach while keeping the random effects fixed (see Baayen, Vasishth, Bates & Kliegl 2017). In this procedure, all the variables that did not significantly contribute to the model fit were removed from the model one by one. The contribution of each individual predictor was also checked by model comparisons using the fREML and AIC scores from the *itsadug*-package (van Rij, Baayen, Wieling & van Rijn, 2016).

Besides the prime condition, which was our main variable of interest, we also considered affix type, length, frequency, orthography-semantics consistency, morphological family size, orthographic neighbourhood size and the semantic similarity between the prime and target word, as possible covariates. Frequencies of the primes and targets were determined using the Corpus of Contemporary American English (Davies 2010). Morphological family size was taken from the CELEX English database (Baayen, Piepenbrock & Gulikers, 1995). The orthography-semantics consistency measures were taken from the English orthography-semantics consistency database (Marelli & Amenta, 2018). The orthographic neighbourhood size is based on the Morpholex database (Sanchez-Gutierrez, Mailhot, Deacon & Wilson, 2018). The semantic similarities between the prime and the target were taken from the Snaut database (Mandera, Keuleers & Brysbaert, 2017).

### 2.3 GAMM results

Results from the standard GAMM analysis fitted to the log-transformed reaction times with the unrelated prime condition as the reference level are summarized in Table 1. The parametric terms represent the main effects of the categorical variables and linear continuous variables. We found that the effect of condition was gradient: the identity condition (*ide*) showed the largest facilitation compared to the baseline condition, followed by the inflected (*inf*) and derived (*der*) conditions. Notably, there were no significant differences between the baseline condition and the pseudo-complex (*pc*) and stem-embedded (*se*) conditions.

We also compared the conditions with each other by changing the reference level of condition to each of the five other levels (*ide*, *inf*, *der*, *pc* and *se*). Pairwise comparisons showed that the priming for the inflected condition was significantly larger than for the derived condition and that the priming for the derived condition was significantly larger than for the pseudo complex and the stem-embedded conditions. The only insignificant pairwise differences were between the identity and inflected conditions and between the pseudo-complex and stem-embedded

conditions. Further, there was a linear effect of word length: longer target words were processed more slowly than shorter target words.

The smooth terms of Table 1 represent both main effects of non-linear continuous variables, as well as random effects. The first smooth term shows that there was a nonlinear effect of target word frequency. This is indicated by the edf-score being greater than 1 (where the the edf reflects the degree of non-linearity of a regression line). It is important to note that the smooth estimates do not indicate the direction of the effect, which instead is determined via visualization. This revealed that more frequent words were recognized faster. The second smooth term represents the by-subject random smooth adjustments for trial (i.e., the order in which the stimuli was presented, ranging from 1 to 360). The third smooth term represents the random intercept adjustments for target items.

Affix type (prefix, suffix, no affix), prime frequency, prime length, semantic similarity, orthography-semantic consistency, as well prime and target morphological family size and orthographic neighbourhood size measures did not significantly contribute to the model fit. All interactions between the predictors were insignificant, as were random intercepts for prime.

**Table 1.** Summary of the partial effects of GAMM fitted to log-transformed masked priming lexical decision reaction times in milliseconds in Experiment 1

A. parametric coefficients	Estimate	Std. Error	T-value	P-value
(Intercept)	6.3724	0.0342	186.2930	< 0.0001
conditionDer	-0.0274	0.0084	-3.2405	0.0012
conditionInf	-0.0544	0.0085	-6.4376	< 0.0001
conditionIde	-0.0694	0.0085	-8.2048	< 0.0001
conditionPc	-0.0058	0.0085	-0.6799	0.4966
conditionSe	0.0008	0.0085	0.0973	0.9225
target length	0.0301	0.0083	3.6294	0.0003
B. smooth terms	Edf	Ref.df	F-value	P-value
smooth for log target frequency	2.5544	2.7013	13.2688	< 0.0001
by-subject random smooths for trial	290.2493	774.0000	5.8276	< 0.0001
random intercepts for target	66.9470	86.0000	3.6992	< 0.0001

## 2.4 QGAM results

In order to obtain model predictions beyond mean reaction times, we conducted a quantile analysis of the reaction times using quantile generalized additive mixed

modelling (QGAM, Fasiolo et al. 2020, the R-package *qgam*), which provided predictions across different reaction time deciles (0.1, 0.2, ..., 0.9). This approach complements the standard GAMM analysis as we can investigate whether certain variables are predictive for short versus long responses, as well as investigate how effects evolve over time. This can provide insights into the gradual change within the time-course between the different conditions. Given that the data is split into deciles, the model estimates are based on a smaller amount of data than in the previous analysis. Therefore, in order to be more conservative in our interpretation of the results, we opted for a p-value significance threshold of 0.01 for the QGAM analysis.

The fixed effects of the final QGAM model were the same as the final GAMM model, with condition as our predictor of interest and target frequency and target length as co-variates. However, given that the current version of the *qgam*-package cannot handle complex random structure, we opted for a simpler random effects structure, including random intercepts for subjects and for target words.

Figure 1 depicts the results. The bar graphs from top left to bottom right present the effects of condition for the nine deciles. The intercept values for each decile model are shown below the bars (e.g., 541.26 for the 0.1st decile), with the p-values within the bars indicating the difference compared to the unrelated baseline condition. The black p-values below the bars represent pairwise differences with the other conditions. As indicated by the changing bar sizes across the panels, the effect of condition changed across the reaction time distribution. In the 0.1st decile, there was a significant facilitation for all conditions compared to the baseline. This effect was gradient, mirroring the magnitude of semantic relatedness between the prime and the target: the identity prime (*ide*) had the strongest priming, followed by inflected (*inf*), derived (*der*), pseudo-complex (*pc*) and stem-embedded (*se*) primes, each showing gradually less facilitation.

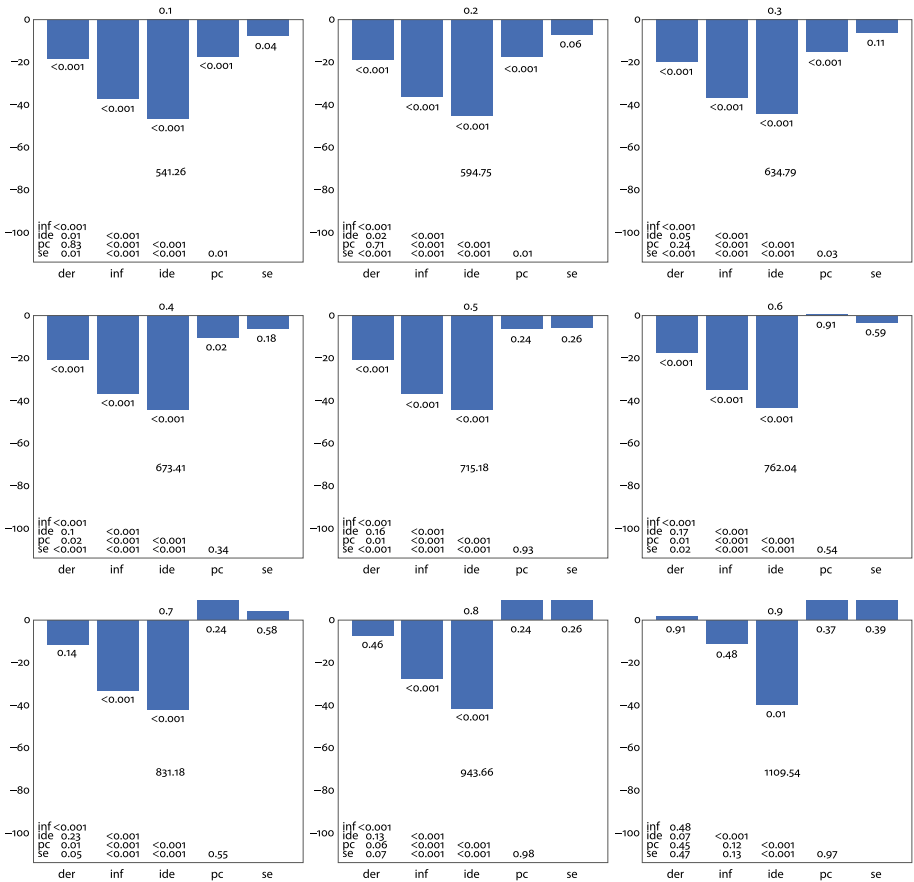
This pattern gradually changed across the deciles. The stem-embedded condition was insignificant from the 0.1st decile onwards and the pseudo-complex condition from the 0.4th decile onwards. The morpho-semantically related conditions remained significant until later deciles, with the derived condition becoming insignificant from the 0.7th decile and the inflected condition from the 0.9th decile. The identity condition remained significant through the last decile. Thus, even if the semantic difference between the inflected and identity conditions is not large, the results suggest that it exists (see Booij 1996).

As we can see from the pairwise condition comparisons in the first decile, the only insignificant comparisons were between the pseudo-complex and derived conditions as well as between the identity and derived conditions and between the pseudo-complex and derived conditions, with other comparisons being signifi-

cant. Pairwise differences between conditions remained significant until the later deciles.

The effect of target frequency and length increased across the deciles, whereas affix type had no effect.

Experiment 2 was the same as Experiment 1 but it was conducted with advanced non-native speakers of English.



**Figure 1.** Effects of Condition (with unrelated baseline as the reference level) in a QGAM fitted to primed lexical decision latencies for 0.1st to 0.9th decile in Experiment 1. Black p-values within the bars show the difference to the baseline, black p-values below the bars pairwise difference between other conditions. Ide-identity; inf-inflected; der-derived; pc-pseudo-complex, se-stem-embedded

### 3. Experiment 2: Advanced L2 English speakers

#### 3.1 Materials and methods

##### 3.1.1 Participants

Seventy-two advanced L2 English speakers (49 female, mean age 20 years, range 18–38) with normal or corrected-to-normal vision participated in the experiment. All participants were students at the University of Alberta with at least a B2 level of English proficiency. Participants came from various L1 language backgrounds, including German, French, Spanish, Arabic, and the most common being Cantonese and Mandarin (44 participants). Written consent was obtained prior to participation and all participants received partial course credit for their participation.

##### 3.1.2 Materials

The materials were the same as in Experiment 1.

##### 3.1.3 Procedure

The procedure and the design were the same as in Experiment 1.

#### 3.2 Analysis

Prior to the analysis, practice trials and fillers were removed from the data set. One participant and one item (the target word *BOG*) were removed from the data set as their error rates were 50% or higher. Trials with incorrect responses (10.6% of the data) were removed. We again performed two analyses, a standard GAMM analysis and a QGAM analysis with the same modeling procedure and variables. Prior to the GAMM analysis trials with response times above or below 2.5 standard deviations of the grand mean were also removed (1.8% of the data).

#### 3.3 GAMM results

The results of the final GAMM model for L2 speakers are shown in Table 2. The GAMM analysis revealed that primes in both semantically related (*ide*, *inf*, and *der*) and semantically unrelated conditions (*pc* and *se*) showed significant facilitation compared to the baseline condition. The largest priming was again found for the identity condition (*ide*) and the smallest for the stem-embedded (*se*) condition. When comparing other critical conditions with each other by changing the reference level, we found that all pairwise differences were insignificant. To summarize, the identity, inflected, derived, pseudo-complex and stem-embedded

conditions showed a larger priming effect than the baseline condition, but there were no differences between the identity and inflected conditions, between the inflected and derived conditions, between the derived and pseudo-complex conditions or between the pseudo-complex and stem-embedded conditions.

Furthermore, unlike for the L1 speakers, there was a significant effect of affix type: prefixed words were processed more slowly than suffixed words, with no difference between suffixes and non-affixed primes. Prime frequency, prime length, orthography-semantics consistency, orthographic neighbourhood size, semantic similarity as well prime and target morphological family size were all insignificant. The final model included condition, target length and affix type as fixed effects, as well as a nonlinear smooth for target frequency, by-subject random smooths for trial and random intercepts for target items. All interactions between the predictors were insignificant, as were random intercepts for prime.

**Table 2.** Summary of the partial effects in GAMM fitted to log-transformed masked priming lexical decision reaction times in milliseconds in Experiment 2

A. parametric coefficients	Estimate	Std. Error	T-value	P-value
(Intercept)	6.4519	0.0495	130.3991	< 0.0001
conditionDer	-0.0492	0.0137	-3.5961	0.0003
conditionInf	-0.0661	0.0136	-4.8493	< 0.0001
conditionIde	-0.0708	0.0259	-2.7337	0.0063
conditionPc	-0.0569	0.0139	-4.0828	< 0.0001
conditionSe	-0.0451	0.0141	-3.2012	0.0014
affixPre	0.0507	0.0147	3.4533	0.0006
no affix	-0.0019	0.0268	-0.0721	0.9425
target length	0.0585	0.0106	5.5446	< 0.0001
B. smooth terms	Edf	Ref.df	F-value	P-value
smooth for log target freq	2.1889	2.3940	44.6460	< 0.0001
by-subject random smooths for trial	209.2505	666.0000	8.1116	< 0.0001
random intercepts for target	59.2282	86.0000	2.2631	< 0.0001

### 3.4 QGAM results

Next, we conducted a quantile analysis of the reaction times using quantile generalized additive mixed modelling (QGAM, Fasiolo et al. 2020, the R-package *qgam*) to capture the time-course of L2 processing. The fixed effects of the final QGAM model were the same as of the final GAMM model, with condition as our



predictor of interest and target frequency and target length as well as affix type as co-variates. However, given that the current version of the *qgam*-package cannot handle complex random structure, we opted again for a simpler random effects structure, including random intercepts for subjects and for target words.

Figure 2 depicts the results of the QGAM. The general pattern in the 0.1st decile resembles that of Experiment 1. There was a significant facilitation for all conditions compared to the baseline, each showing gradually less facilitation in order of the semantic relatedness between the prime and the target: the largest priming was for the identity condition and the smallest priming for the stem-embedded condition. Surprisingly, the significant priming effect for the identity condition disappeared from 0.6th decile onwards, but remained present for the other conditions. As for pairwise comparisons, almost all were insignificant, with the only significant difference being between the stem-embedded and inflected conditions, remaining significant until the 0.6th decile.

Finally, the effect of target length and target frequency increased across deciles, such that later deciles were more affected by frequency and length than earlier deciles. The effect of affix type also changed over time, with prefixed words becoming increasingly more inhibitory across deciles. The data and the code for the complete statistical analyses of Experiment 1 and 2 are available on Open Science Framework (<https://osf.io/b5vst/>).

In summary, our results suggest several differences between L1 and L2 speakers. First, when looking at Figures 1 and 2, it appears that the effect sizes for most conditions are larger for L2 speakers compared to L1 speakers. Second, unlike for L1 speakers, priming effects for the pseudo-complex and stem-embedded conditions remained significant until later deciles.

#### 4. Discussion

The present study investigated morphological processing with both native (L1) and advanced nonnative (L2) speakers of English, using the masked priming lexical decision paradigm. We used a full within-item design consisting of six different conditions: identity (*bull*); inflected (*bulls*), derived (*bullish*), pseudo-complex (*bully*), stem-embedded (*bullet*), and unrelated baseline prime (*disks*).

In Experiment 1, L1 speakers only showed facilitation for true morpho-semantic relations: neither pseudo complex (e.g., *bully*) nor stem-embedded (*bullet*) words produced significant priming in a standard regression analysis. As for the true morpho-semantic conditions, inflected words (*bulls*) showed more priming than derived ones (*bullish*).

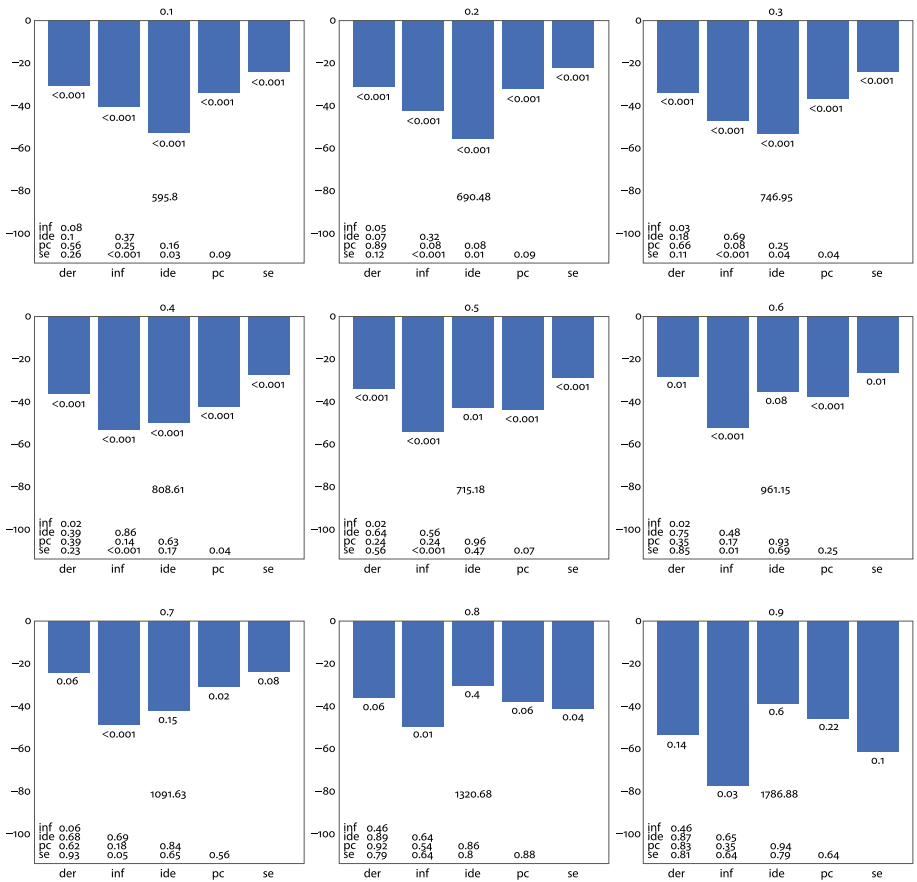
Quantile regression analysis extended the findings of the standard analysis in two important ways. First, it revealed differences between conditions varying with the semantic relatedness between the prime and target from the first decile onward, such that the identity condition showed the most priming and the pseudo-complex condition the least. Second, it also showed an almost one-to-one correspondence between the degree of semantic relatedness and the decile at which the condition in question ceased to be significant. For example, the identity condition remained significant until the last decile, whereas the stem-embedded condition was not reaching the significance threshold even in the first decile.

Thus, both of our L1 analyses are in line with research suggesting that from early on, morphological priming is gradient and sensitive to semantics (e.g., Diependaele, Sandra & Grainger 2009; Feldman et al. 2009; Jared et al. 2017; Järvikivi & Pyykkönen 2011; Morris et al. 2007), in contrast to the predictions made by a semantically blind approach (e.g., Rastle et al., 2004). At the same time, it is important to point out that our study did not manipulate or control for the semantic relatedness or transparency between the prime and the target, rather this seems the most natural explanation given how the effects differ between conditions, as well as how those differences emerge over time.

The observation that the processing of complex words is driven by word level properties, is supported by other recent L1 morphological processing research. For example, it has been shown that distributional properties of the surface word form, including whole-word frequency, paradigm size (Baayen, Wurm & Aycok, 2007; Lõo, Järvikivi & Baayen, 2018; Lõo et al., 2018), morphological family size (De Jong, Schreuder & Baayen, 2003; Moscoso del Prado Martín, Bertram, Häikiö, Schreuder & Baayen, 2004; Schreuder & Baayen, 1997), affective properties, such as psychological arousal and valence (positivity) (Kuperman, 2013; Schmidtke et al., 2017) and orthography-semantics consistency measures (Amenta et al., 2020), influence lexical processing very early on, occurring at the same time or even earlier than properties of component morphemes.

Coming from a blind decomposition perspective (Rastle et al., 2004), the gradient effects presented in Experiment 1 can in principle be explained by a processing mechanism by which stems and affixes of a word are separated and put back together. However, in that case, the timing of morphological decomposition and recombination would need to occur before the fastest response times in the earliest deciles, which already show a difference between pseudo and real morphology. This seems unlikely given what we know about the onset of semantic effects in word processing in general (Hauk, Davis, Ford, Pulvermüller & Marslen-Wilson 2006; Reingold, Reichle, Glaholt & Sheridan 2012).

In Experiment 2, L2 speakers showed significant facilitation across all five conditions compared to the baseline, including the stem-embedded condition



**Figure 2.** Effects of condition (with unrelated baseline as the reference level) in a QGAM fitted to primed lexical decision latencies for 0.1st to 0.9th decile in Experiment 2. Black p-values within the bars show the difference to the baseline, black p-values below the bars pairwise difference between other conditions. Ide-identity; inf-inflected, der-derived; pc-pseudo-complex; se-stem-embedded

(bullet). However, there were no significant differences between the experimental conditions themselves. The quantile regression analysis was in line with the standard GAMM analysis, showing that unlike for L1 speakers the effect of all conditions, including the pseudo-complex and stem-embedded conditions, remained significant until the later deciles. This suggests that L2 processing is primarily driven by bottom-up information. It is possible that L2 processing is also driven by top-down information, but L2 speakers were unable to access the meaning of the words fast enough in rapid tasks, such as masked priming lexical decision. Delayed access to the meaning of L2 words has been also often been reported

in the literature on the multilingual lexicon (Dijkstra & van Heuven, 2018). In our data, possible semantic influence is suggested by the fact that the stem-embedded (*bullet*) condition resulted in significantly less facilitation than the inflected condition (*bulls*) across the first six deciles. However, as the identity condition became insignificant from the 0.6th decile, the QGAM results should be considered with some caution. One possibility is that the identity condition ceased to be significant due to a task effect. Given that the majority of primes are complex, L2 speakers may start to expect complex primes, in particular with longer responses. Another possibility is data sparsity for certain conditions in certain deciles.

In general, L2 speakers seem to be both slower and less sensitive to the semantics of the word and effects are mostly driven by the form overlap (Heyer & Clahsen, 2015; Silva & Clahsen, 2008). However, it is also important to point out that the L2 participants came from various language backgrounds and we only used self-reported proficiency scores to determine L2 language proficiency. Conclusions based on a more homogeneous group of L2-speakers may differ. Given the potential issues raised above, the evidence for L2 processing remains more inconclusive than the evidence for the L1 processing.

Our results suggest that L1 morphological processing, like many other aspects of language comprehension (see e.g., Tanenhaus, Spivey-Knowlton, Eberhard & Sedivy 1995; Traxler 2014), is first and foremost a quest for meaning, in which multiple – often conflicting – sources of information emerge (see also Kuperman et al. 2009; Libben 2006). Even a short presentation of a prime seems to trigger the emergence of bottom-up and top-down information about the word, and these effects change rapidly over the time course of processing. The current study shows that the inspection of the time-course, as well as the use of more fine-grained stimulus conditions, can add new insights into how and when certain types of information emerge in morphological processing.

These findings do not dismiss the importance of morphological structure as a cue in language processing. Rather, they suggest that morphological effects can emerge from the shared form and meaning within the word. A gradient view to morphological processing does not assume stems and affixes as separate entities in a symbolic-like computational system, but as a network of interconnected form-meaning mappings. This view aligns with linguistic theories such as the Construction Grammar (e.g., Diessel, 2019) but also with computational models of morphological processing within the connectionist framework (e.g., Plaut & Gonnerman, 2000). Accordingly, we suggest that morphologically complex words are subject to early facilitation from orthography (i.e., bottom-up information), which becomes interfered with by increasing semantic influence. For example, even a word like *corner* can facilitate the processing of *corn* early on based on

the orthographic overlap, but as semantic information is accessed, the facilitation starts to diminish.

This study adds to recent research approaches, using novel methodological and conceptual approaches, all of which point to a picture that is more complex than what a pure feed-forward decomposition approach to morphological processing would predict (Amenta et al., 2020; Günther et al., 2020; Hendrix & Sun, 2020; Schmidtke et al., 2017; Ulicheva et al., 2020). Finally, many recent accounts in theoretical morphology have already accepted words as basic building blocks in their own right (Anderson, 1992; Aronoff, 1994; Beard, 1995; Blevins, 2016; Booij, 2010; Marzi, Blevins, Booij & Pirrelli, 2020).

A recent review acknowledges (Crepaldi, Amenta & Marelli, 2019) that the discreteness at the morphemic level that most current models of morphological processing assume is not actually observed in data: “[...] part of the reason behind this state of affairs is the still dominant view of morphemes as discrete units that undergo combinatorial rules; a discrete, black-and-white, compositional morphology. [...] data quite often depict blurred boundaries, seem to lie on a continuum, and are probabilistic in nature.” This is also exactly what we observe in our current study. We suggest that the processing of complex words is gradient not discrete when we compare the different conditions with each other, but also when we observe the changes in priming effects over time. How exactly this can be translated into morphological processing models requires future research.

## Funding

This research was supported by Estonian Research Council grant PSG743 and Mobilitas Plus postdoctoral researcher grant (MOBJD408) given to the first author and by a Social Sciences and Humanities Research Council (<http://www.sshrc-crsh.gc.ca/>) Partnership Grant (Words in the World, 895-2016-1008) given to the last author.

## Acknowledgements

We would like to thank Melanie Bell, Daniel Schmidtke and one anonymous reviewer for their helpful comments on an earlier version of this paper.

## References

- Amenta, S., Crepaldi, D., & Marelli, M. (2020). Consistency measures individuate dissociating semantic modulations in priming paradigms: A new look on semantics in the processing of (complex) words. *Quarterly Journal of Experimental Psychology*, 1747021820927663. <https://doi.org/10.1177/1747021820927663>

- Anderson, S. R. (1992). *A-morphous morphology*. Cambridge: Cambridge University Press.  
<https://doi.org/10.1017/CBO9780511586262>
- 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. <https://doi.org/10.1016/j.jml.2012.12.001>
- Aronoff, M. (1994). *Morphology by Itself: Stems and Inflectional Classes*. Cambridge, Mass.: The MIT Press.
- Baayen, R. H., Chuang, Y.-Y., Shafaei-Bajestan, E., & Blevins, J. P. (2019). The discriminative lexicon: A unified computational model for the lexicon and lexical processing in comprehension and production grounded not in (de) composition but in linear discriminative learning. *Complexity*, 2019. <https://doi.org/10.1155/2019/4895891>
- Baayen, R. H., Davidson, D. J., & Bates, D. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59, 390–412. <https://doi.org/10.1016/j.jml.2007.12.005>
- 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. <https://doi.org/10.1037/a0023851>
- Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database (CD-ROM)*. University of Pennsylvania, Philadelphia, PA: Linguistic Data Consortium.
- Baayen, R. H. & Smolka, E. (2020). Modeling morphological priming in German with naive discriminative learning. *Frontiers in Communication*, 5, 17. <https://doi.org/10.3389/fcomm.2020.00017>
- Baayen, R. H., Vasishth, S., Bates, D., & Kliegl, R. (2017). The cave of shadows. addressing the human factor with generalized additive mixed models. *Journal of Memory and Language*, 56, 206–234. <https://doi.org/10.1016/j.jml.2016.11.006>
- 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. <https://doi.org/10.1075/ml.2.3.06baa>
- Beard, R. (1995). *Lexeme-morpheme base morphology: A general theory of inflection and word formation*. Albany, NY.: State University of New York Press.
- Bell, M. J. & Schäfer, M. (2016). Modelling semantic transparency. *Morphology*, 26 (2), 157–199. <https://doi.org/10.1007/s11525-016-9286-3>
- 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. <https://doi.org/10.3758/s13423-015-0927-z>
- Blevins, J. P. (2016). *Word and paradigm morphology*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199593545.001.0001>
- Booij, G. (2010). Construction morphology. *Language and linguistics compass*, 4(7), 543–555. <https://doi.org/10.1111/j.1749-818X.2010.00213.x>
- Booij, G. E. (1996). Inherent versus contextual inflection and the split morphology hypothesis. In G. E. Booij & J. v. Marle (Eds.), *Yearbook of Morphology 1995* (pp. 1–16). Dordrecht: Kluwer Academic Publishers. [https://doi.org/10.1007/978-94-017-3716-6\\_1](https://doi.org/10.1007/978-94-017-3716-6_1)
- Clahsen, H. & Felser, C. (2006). Grammatical processing in language learners. *Applied psycholinguistics*, 27(1), 3. <https://doi.org/10.1017/S0142716406060024>

- Clahsen, H., Felser, C., Neubauer, K., Sato, M., & Silva, R. (2010). Morphological structure in native and nonnative language processing. *Language Learning*, 60(1), 21–43. <https://doi.org/10.1111/j.1467-9922.2009.00550.x>
- Coughlin, C. E. & Tremblay, A. (2015). Morphological decomposition in native and non-native french speakers. *Bilingualism: Language and Cognition*, 18(3), 524–542. <https://doi.org/10.1017/S1366728914000200>
- Crepaldi, D., Amenta, S., & Marelli, M. (2019). For a probabilistic and multidisciplinary approach to the investigation of morphological processing. *Cortex*, 116. <https://doi.org/10.1016/j.cortex.2019.05.010>
- 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. <https://doi.org/10.1093/lilc/fqq018>
- Davis, C. P., Libben, G., & Segalowitz, S. J. (2019). Compounding matters: Event-related potential evidence for early semantic access to compound words. *Cognition*, 184, 44–52. <https://doi.org/10.1016/j.cognition.2018.12.006>
- 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. <https://doi.org/10.1515/9783110910186.65>
- Diependaele, K., Duñabeitia, J. A., Morris, J., & Keuleers, E. (2011). Fast morphological effects in first and second language word recognition. *Journal of Memory and Language*, 64 (4), 344–358. <https://doi.org/10.1016/j.jml.2011.01.003>
- Diependaele, K., Sandra, D., & Grainger, J. (2009). Semantic transparency and masked morphological priming: The case of prefixed words. *Memory & Cognition*, 37(6), 895–908. <https://doi.org/10.3758/MC.37.6.895>
- Diessel, H. (2019). *The grammar network*. Cambridge University Press. <https://doi.org/10.1017/9781108671040>
- Dijkstra, T. & van Heuven, W. J. (2018). Visual word recognition in multilinguals. *The Oxford handbook of psycholinguistics*, 118–143.
- Fasiolo, M., Wood, S. N., Zaffran, M., Nedellec, R., & Goude, Y. (2020). Fast calibrated additive quantile regression. *Journal of the American Statistical Association*, 1–11.
- Feldman, L. B. (1994). Beyond orthography and phonology: Differences between inflections and derivations. *Journal of Memory and Language*, 33(4), 442–470. <https://doi.org/10.1006/jmla.1994.1021>
- Feldman, L. B. & Basnight-Brown, D. M. (2008). List context fosters semantic processing: Parallels between semantic and morphological facilitation when primes are forward masked. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34 (3), 680.
- 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. <https://doi.org/10.3758/PBR.16.4.684>
- Forster, K. & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10 (4), 680–698. <https://doi.org/10.1037/0278-7393.10.4.680>
- Giraudo, H. & Grainger, J. (2001). Priming complex words: Evidence for supralelexical representation of morphology. *Psychonomic Bulletin and Review*, 8, 127–131. <https://doi.org/10.3758/BF03196148>

- Gonnerman, L. M., Seidenberg, M. S., & Andersen, E. S. (2007). Graded semantic and phonological similarity effects in priming: Evidence for a distributed connectionist approach to morphology. *Journal of Experimental Psychology: General*, 136 (2), 323. <https://doi.org/10.1037/0096-3445.136.2.323>
- Günther, F., Petilli, M. A., & Marelli, M. (2020). Semantic transparency is not invisibility: A computational model of perceptually-grounded conceptual combination in word processing. *Journal of Memory and Language*, 112, 104104. <https://doi.org/10.1016/j.jml.2020.104104>
- Hasenäcker, J., Beyersmann, E., & Schroeder, S. (2016). Masked morphological priming in German-speaking adults and children: Evidence from response time distributions. *Frontiers in Psychology*, 7, 929. <https://doi.org/10.3389/fpsyg.2016.00929>
- Hauk, O., Davis, M., Ford, M., Pulvermüller, F., & Marslen-Wilson, W. (2006). The time course of visual word recognition as revealed by linear regression analysis of ERP data. *NeuroImage*, 30, 1383–1400. <https://doi.org/10.1016/j.neuroimage.2005.11.048>
- Heathcote, L., Nation, K., Castles, A., & Beyersmann, E. (2018). Do 'blacheap' and 'subcheap' both prime 'cheap'? an investigation of morphemic status and position in early visual word processing. *Quarterly Journal of Experimental Psychology*, 71 (8), 1645–1654. <https://doi.org/10.1080/17470218.2017.1362704>
- Hendrix, P. & Sun, C. C. (2020). The role of information theory for compound words in Mandarin Chinese and English. *Cognition*, 205, 104389. <https://doi.org/10.1016/j.cognition.2020.104389>
- Heyer, V. & Clahsen, H. (2015). Late bilinguals see a scan in scanner and in scandal: dissecting formal overlap from morphological priming in the processing of derived words. *Bilingualism: Language and Cognition*, 18(3), 543–550. <https://doi.org/10.1017/S1366728914000662>
- Jacob, G., Heyer, V., & Veríssimo, J. (2018). Aiming at the same target: A masked priming study directly comparing derivation and inflection in the second language. *International Journal of Bilingualism*, 22(6), 619–637. <https://doi.org/10.1177/1367006916688333>
- Jared, D., Jouravlev, O., & Joanisse, M. F. (2017). The effect of semantic transparency on the processing of morphologically derived words: Evidence from decision latencies and event-related potentials. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43(3), 422.
- Järvikivi, J. & Pyykkönen, P. (2011). Sub- and supralephical information in early phases of lexical access. *Frontiers in Psychology*, 2. <https://doi.org/10.3389/fpsyg.2011.00282>
- Jiang, N. (2000). Lexical representation and development in a second language. *Applied linguistics*, 21 (1), 47–77. <https://doi.org/10.1093/applin/21.1.47>
- Kazanina, N. (2011). Decomposition of prefixed words in Russian. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(6), 1371.
- Kuperman, V. (2013). Accentuate the positive: Semantic access in English compounds. *Frontiers in Psychology*, 4. <https://doi.org/10.3389/fpsyg.2013.00203>
- Kuperman, V., Schreuder, R., Bertram, R., & Baayen, R. H. (2009). Reading of multimorphemic Dutch compounds: Towards a multiple route model of lexical processing. *Journal of Experimental Psychology: HPP*, 35, 876–895.
- Lemhöfer, K., Dijkstra, T., Schriefers, H., Baayen, R. H., Grainger, J., & Zwitserlood, P. (2008). Native language influences on word recognition in a second language: A megastudy. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34 (1), 12.



- Leminen, A., Smolka, E., Dunabeitia, J.A., & Pliatsikas, C. (2019). Morphological processing in the brain: The good (inflection), the bad (derivation) and the ugly (compounding). *Cortex*, 116, 4–44. <https://doi.org/10.1016/j.cortex.2018.08.016>
- Libben, G. (2006). Why study compound processing? an overview of the issues. *The representation and processing of compound words*, 1–22.
- Longtin, C., Segui, J., & Hallé, P. (2003). Morphological priming without morphological relationship. *Language and Cognitive Processes*, 18(3), 313–334. <https://doi.org/10.1080/01690960244000036>
- Lõo, K. & Järvikivi, J. (2019). Whole-word frequency effects in English masked priming: very little CORN in CORNER and CORNET. *Proceedings of The 11th International Conference on the Mental Lexicon*, 1, e072.
- 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. <https://doi.org/10.1016/j.cognition.2018.02.002>
- 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. <https://doi.org/10.1007/s11525-017-9318-7>
- Mandera, P., Keuleers, E., & Brysbaert, M. (2017). Explaining human performance in psycholinguistic tasks with models of semantic similarity based on prediction and counting: A review and empirical validation. *Journal of Memory and Language*, 92, 57–78. <https://doi.org/10.1016/j.jml.2016.04.001>
- Manelis, L. & Tharp, D.A. (1977). The processing of affixed words. *Memory and Cognition*, 5, 690–695. <https://doi.org/10.3758/BF03197417>
- Marelli, M. & Amenta, S. (2018). A database of orthography-semantics consistency (osc) estimates for 15,017 English words. *Behavior Research Methods*, 50 (4), 1482–1495. <https://doi.org/10.3758/s13428-018-1017-8>
- Marelli, M. & Baroni, M. (2015). Affixation in semantic space: Modeling morpheme meanings with compositional distributional semantics. *Psychological Review*, 122(3), 485. <https://doi.org/10.1037/a0039267>
- Marzi, C., Blevins, J.P., Booij, G., & Pirrelli, V. (2020). Inflection at the morphology-syntax interface. *Word Knowledge and Word Usage*, 228. <https://doi.org/10.1515/9783110440577-007>
- McDonald, J.L. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language*, 55(3), 381–401. <https://doi.org/10.1016/j.jml.2006.06.006>
- Morris, J., Frank, T., Grainger, J., & Holcomb, P.J. (2007). Semantic transparency and masked morphological priming: An ERP investigation. *Psychophysiology*, 44 (4), 506–521. <https://doi.org/10.1111/j.1469-8986.2007.00538.x>
- 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.
- Mulder, K., Dijkstra, T., & Baayen, R.H. (2015). Cross-language activation of morphological relatives in cognates: The role of orthographic overlap and task-related processing. *Frontiers in Human Neuroscience*, 9, 16. <https://doi.org/10.3389/fnhum.2015.00016>
- Mulder, K., Dijkstra, T., Schreuder, R., & Baayen, R.H. (2014). Effects of primary and secondary morphological family size in monolingual and bilingual word processing. *Journal of Memory and Language*, 72, 59–84. <https://doi.org/10.1016/j.jml.2013.12.004>

- Norris, D. & Kinoshita, S. (2008). Perception as evidence accumulation and bayesian inference: Insights from masked priming. *Journal of Experimental Psychology*, 137(3), 434–455. <https://doi.org/10.1037/a0012799>
- 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. <https://doi.org/10.1080/01690960050119661>
- Rastle, K. & Davis, M.H. (2008). Morphological decomposition based on the analysis of orthography. *Language and Cognitive Processes*, 23(7–8), 942–971. <https://doi.org/10.1080/01690960802069730>
- 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. <https://doi.org/10.3758/BF03196742>
- Raveh, M. (2002). The contribution of frequency and semantic similarity to morphological processing. *Brain and Language*. <https://doi.org/10.1006/brln.2001.2527>
- Reingold, E.M., Reichle, E.D., Glaholt, M.G., & Sheridan, H. (2012). Direct lexical control of eye movements in reading: Evidence from a survival analysis of fixation durations. *Cognitive psychology*, 65(2), 177–206. <https://doi.org/10.1016/j.cogpsych.2012.03.001>
- Sánchez-Gutiérrez, C.H., Mailhot, H., Deacon, S.H., & Wilson, M.A. (2018). MorphoLex: A derivational morphological database for 70,000 English words. *Behavior Research Methods*, 50(4), 1568–1580. <https://doi.org/10.3758/s13428-017-0981-8>
- Schmidtke, D. & Kuperman, V. (2019). A paradox of apparent brainless behavior: The time-course of compound word recognition. *Cortex*, 116, 250–267. <https://doi.org/10.1016/j.cortex.2018.07.003>
- 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. <https://doi.org/10.1037/xlm0000411>
- Schreuder, R. & Baayen, R.H. (1995). Modeling morphological processing. In L.B. Feldman (Ed.), *Morphological Aspects of Language Processing* (pp. 131–154). Hillsdale, New Jersey: Lawrence Erlbaum.
- Schreuder, R. & Baayen, R.H. (1997). How complex simplex words can be. *Journal of Memory and Language*, 37, 118–139. <https://doi.org/10.1006/jmla.1997.2510>
- Silva, R. & Clahsen, H. (2008). Morphologically complex words in L1 and L2 processing: Evidence from masked priming experiments in English. *Bilingualism: Language and Cognition*, 11(2), 245–260. <https://doi.org/10.1017/S1366728908003404>
- Taft, M. (2004). Morphological decomposition and the reverse base frequency effect. *The Quarterly Journal of Experimental Psychology*, 57A, 745–765. <https://doi.org/10.1080/02724980343000477>
- Taft, M. & Forster, K.I. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior*, 14, 638–647. [https://doi.org/10.1016/S0022-5371\(75\)80051-X](https://doi.org/10.1016/S0022-5371(75)80051-X)
- Tanenhaus, M., Spivey-Knowlton, M., Eberhard, K., & Sedivy, J. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, 268, 1632–1634. <https://doi.org/10.1126/science.7777863>
- Traxler, M.J. (2014). Trends in syntactic parsing: Anticipation, bayesian estimation, and good-enough parsing. *Trends in Cognitive Sciences*, 18(11), 605–611. <https://doi.org/10.1016/j.tics.2014.08.001>

- Tucker, B. V., Brenner, D., Danielson, D. K., Kelley, M. C., Nenadić, F., & Sims, M. (2019). The massive auditory lexical decision (MALD) database. *Behavior Research Methods*, 51 (3), 1187–1204. <https://doi.org/10.3758/s13428-018-1056-1>
- Tzur, B. & Frost, R. (2007). SOA does not reveal the absolute time course of cognitive processing in fast priming experiments. *Journal of Memory and Language*, 56 (3), 321–335. <https://doi.org/10.1016/j.jml.2006.11.007>
- Ulicheva, A., Harvey, H., Aronoff, M., & Rastle, K. (2020). Skilled readers' sensitivity to meaningful regularities in English writing. *Cognition*, 195, 103810. <https://doi.org/10.1016/j.cognition.2018.09.013>
- 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.
- Viviani, E. & Crepaldi, D. (2019). Masked morphological priming tracks the development of a fully mature lexical system in L2. <https://doi.org/10.31219/osf.io/xdsng>
- Voga, M., Anastassiadis-Symeonidis, A., & Giraudo, H. (2014). Does morphology play a role in L2 processing?: Two masked priming experiments with Greek speakers of ESL. *Linguisticae Investigationes*, 37(2), 338–352. <https://doi.org/10.1075/li.37.2.10vog>
- Wood, S. N. (2017). *Generalized additive models: an introduction with R*. CRC press. <https://doi.org/10.1201/9781315370279>

## Address for correspondence

Kaidi Lõo  
Institute of Estonian and General Linguistics  
University of Tartu  
Jakobi 2-405  
50090 Tartu  
Estonia  
[kaidi.loo@ut.ee](mailto:kaidi.loo@ut.ee)

## Co-author information

Abigail Toth  
Department of Artificial Intelligence –  
Cognitive Modeling Group  
University of Groningen  
[a.g.toth@rug.nl](mailto:a.g.toth@rug.nl)

Figen Karaca  
Radboud University  
[f.karaca@let.ru.nl](mailto:f.karaca@let.ru.nl)  
Juhani Järvikivi  
Department of Linguistics  
University of Alberta  
[jarvikivi@ualberta.ca](mailto:jarvikivi@ualberta.ca)

## Publication history

Date received: 28 April 2021

Date accepted: 15 December 2021

Published online: 18 March 2022