



Humanwissenschaftliche Fakultät

Tina Marusch | Lena A. Jäger | Leander Neiß | Frank Burchert |
Lyndsey Nickels

Overt language production of German past participles

investigating (ir-)regularity

Suggested citation referring to the original publication:
Language, Cognition and Neuroscience (2018)
DOI <http://dx.doi.org/10.1080/23273798.2018.1527936>
ISSN (print) 2327-3798
ISSN (online) 2327-3801

Postprint archived at the Institutional Repository of the Potsdam University in:
Postprints der Universität Potsdam
Humanwissenschaftliche Reihe ; 492
ISSN 1866-8364
<http://nbn-resolving.de/urn:nbn:de:kobv:517-opus4-420621>

Overt language production of German past participles: investigating (ir-)regularity

Tina Marusch ^{a,b,d}, Lena A. Jäger ^{a,e}, Leander Neiß^c, Frank Burchert ^a and Lyndsey Nickels ^{b,d}

^aDepartment of Linguistics, University of Potsdam, Potsdam, Germany; ^bARC Centre of Excellence in Cognition and its Disorders, Department of Cognitive Science, Macquarie University, Sydney, Australia; ^cHasso Plattner Institut, Potsdam, Germany; ^dInternational Doctorate for Experimental Approaches to Language and Brain (IDEALAB), Potsdam, Germany; ^eDepartment of Computer Science, Potsdam, Germany and the Weizenbaum Institute for the Networked Society, Berlin, Germany

ABSTRACT

We report two experiments and Bayesian modelling of the data collected. In both experiments, participants performed a long-lag primed picture naming task. Black-and-white line drawings were used as targets, which were overtly named by the participants. Their naming latencies were measured. In both experiments, primes consisted of past participle verbs (*er tanzt/er hat getanzt* "he dances/he has danced") and the relationship between primes and targets was either morphological or unrelated. Experiment 1 additionally had phonologically and semantically related prime-target pairs as well as present tense primes. Both in Experiment 1 and 2, participants showed significantly faster naming latencies for morphologically related targets relative to the unrelated verb primes. In Experiment 1, no priming effects were observed in phonologically and semantically related control conditions. In addition, the production latencies were not influenced by verb type.

ARTICLE HISTORY

Received 21 July 2017
Accepted 17 September 2018

KEYWORDS

Overt language production; long-lag priming; regularity; Bayesian analysis; German past participles

1. Introduction

The way in which regular and irregular verbs are represented and processed has been studied intensively over the past decades using various experimental methods. Priming is one such method, however, the majority of priming studies have focused on language comprehension and studies on the processing of inflectional verb morphology in language production are rare. Moreover, regular and irregular verbs are often treated as dichotomous verb categories in the literature. Yet, in many languages, including English and German, irregular verbs show a variety of idiosyncratic patterns which provide test cases for current models of inflectional morphology.


The approach taken in the current study aimed to investigate the mechanisms underlying (ir-)regularity in language production using a primed picture naming paradigm testing German past participle and present tense forms. In the section below, we will first provide an overview of German inflectional morphology. Subsequently, relevant theories of morphological processing of regular and irregular verbs in language comprehension are introduced, followed by a review of

experimental research focused on the effects of (ir-)regularity on language production.

1.1. Verbal inflectional morphology in German

Psycholinguistic experiments typically contrast English regular and irregular verb morphology in order to understand the processing of these different verb classes. However, regular and irregular are not necessarily straightforward verb categories. This is true for both English and German as they are closely related languages. Both languages use suffixation to form the simple past (that is, *-te* in German and *-ed* in English, which is phonologically realised in English as /t/, /d/ or /ɪd/ depending on the stem final phoneme). German linguist Jacob Grimm labelled such verbs *weak verbs* because they need help from a suffix to form past tense (Elsen, 2011, p. 179) and these verbs can be considered regular. *Strong verbs*, on the other hand, form their past tense by changing the vowel in the verb stem undergoing a process known as *ablaut* and *gradation* (e.g. *know – knew*). Strong verbs can be considered to be a subgroup of irregular verbs (Elsen, 2011,

CONTACT Tina Marusch  marusch@uni-potsdam.de

 Supplemental data for this article can be accessed at <https://doi.org/10.1080/23273798.2018.1527936>

© 2018 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

p. 179). However, German and English irregular past participle forms show very many different patterns, which may include stem changes (e.g. English: *to sing* – *sung*), *-n* suffixation (e.g. English: *blow* – *blown*, German: *halten* “to hold” – *gehalten* “held”) or both suffixation and stem changes (English: *break* – *broken*, German: *singen* “to sing” – *gesungen* “sung”; Quirk et al., 1973). There is also an intermediate group of verbs which show stem changes in the past tense and past participle but also require the weak affix *-t* (e.g. German: *brennen* “to burn”, *brannte* “burnt”, *gebrannt* “has burnt”, English: *sleep*, *slept*, *slept*). These are called *mixed verbs* (e.g. Marcus et al., 1995).

As this study focuses on German, we will now highlight key aspects of German verb inflection. In regular verbs, German uses the dental suffix *-te* to form the simple past. In contrast, irregular verbs form the simple past by changing the stem vowel (e.g. *lesen* “to read” – *las* “read”).

For the past participle, the dental suffix *-t* is used for regular verbs (e.g. *fragen* “to ask” – *hat gefragt* “has asked”), while for irregular past participles, the suffix *-en* is attached to the stem (e.g. *halten* “to hold” – *hat gehalten* “has held”). In addition, the prefix *ge-* is added to form the past participle of all verbs which are stressed on the first syllable (Wiese, 1996).

Irregular past participles may or may not show a stem change and differ in whether the past participle stem is identical to the simple past stem. This property is reflected in Paul’s (2007) commonly used terminology in which verbs are classified as AAA, ABA, ABB and ABC verbs. The capital letters reflect the number of different stems in past tense and past participle relative to the infinitive. For example, the regular verb *tanzen* “to dance” is an AAA verb. It has the stem *tanz* in the infinitive form (*tanzen*). This stem is “A”, the same stem also appears in the past tense *tanzte* “danced” (A) and *getanzt* “has danced” (A). In contrast, *beissen*, is an ABB irregular verb. It has the stem *beis* (A) in the infinitive, *biss* “bit” in the past tense (B) and *biss* (B) in the past participle *biss* “has bitten”. (ABC: *springen* “to jump” (A), *sprang* “jumped” (B), *gesprungen* “has jumped” (C)) (see Table 1).

Table 1. German verb class definitions used in this study.

Verb class	Paul (2007)	Infinitive	Past participle		
			Affix type	Stem change	Example
Regular	AAA	<i>tanzen</i>	–t	no	<i>er hat getanzt</i> “he has danced”
Irregular 1	ABA	<i>lesen</i>	–n	no	<i>sie hat geles-en</i> “she has read”
Irregular 2	ABB/ ABC	<i>gießen</i>	–n	yes	<i>er hat gegoss-en</i> “he has watered”

Irregular verbs (ABA, ABB, ABC) are therefore not a homogeneous group. Even though irregular verbs historically followed a predictable pattern (Mailhammer, 2007), in present day German, their formation cannot be captured by a rule as is the case for regular verbs. For example, second language learners of German might assume that verbs which are phonologically similar to other verbs would follow the same ablaut pattern (e.g. *trinken* “to drink” – *trank* “drunk” – *getrunken* “drank”). While generalising this pattern works for some verbs (e.g. *stinken* “to stink” – *stank* “stank” – *gestunken* “stunk”), other verbs follow the regular declension (e.g. *blinken* “to flash” – *blinkte* “flashed” – *geblinkt* “flashed”) and phonological similarity is by no means a reliable indicator of the correct form.

For German, irregularities in past participles can arise by affixation and/or through stem changes. The four verb types that arise from the combination of these features vary in their degree of irregularity and unpredictability. Regular participles are the most predictable because they may be derived from a grammatical rule, while irregular 2 past participles, which contain a stem change and take the affix *-n* (*gießen* “to water” – *gegossen* “has watered”) are the least predictable and most irregular. However, it is not obvious whether the combination stem change and affix *-t* (mixed verbs: *brennen* “to burn” – *gebrannt* “has burnt”) or the combination absence of stem change and affix *-n* (irregular 1: *lesen* “to read” – *gelesen* “has read”) is more “irregular”.

These morphological patterns offer a source of evidence in the debate regarding how past tense is represented and processed because different accounts make differing predictions regarding the processing of these verb classes. For the purpose of testing these predictions, we maintain the traditional distinction between regular and irregular verbs and subdivide irregular verbs into two groups: irregular 1 and irregular 2 verbs (Smolka et al., 2007, see Table 1), which differ in the presence of stem changes in past participle.

1.2. Morphological processing of (ir-)regularity in language comprehension

A number of explanatory approaches have been put forward to account for the representation and processing of regular and irregular verbs. The *Words and Rules Approach* also known as the *Dual Mechanism Model* (Pinker & Prince, 1994; Clahsen, 1999; Pinker & Ullman, 2002), assumes that regular and irregular verbs are processed in fundamentally different ways. In the Dual Mechanism Model, regular verbs are described by symbolic rules and, hence, word forms of a verb’s paradigm are predictable from the verb stem. In German and

English, regular verb forms are decomposed into stem and affix during language comprehension and stem and affix are combined during language production (Pinker & Prince, 1994).

Among the first studies to explore the processing of regularity in German was the study by Sonnenstuhl et al. (1999) which used a cross-modal priming design to investigate past participles. Participants heard spoken primes and were asked to perform lexical decision to the immediately-presented visual targets. First person present tense targets of regular and irregular verbs were preceded either by themselves (identity condition: prime *tanze* – target *tanze* “I dance”), by their past participle (morphological condition: prime *getanzt* “has danced” – target *tanze* “I dance”) or an unrelated form (unrelated condition: prime *wünsche* “I wish” – target *tanze* “I dance”). Irregular verbs were controlled for stem changes: Only those verbs were used which do not change the stem in the past participle (irregular 1 verbs in our study).

There was “full” priming observed for regular first person present tense targets which had been preceded by regular past participles (prime *getanzt* “has danced” – target *tanze* “I dance”). However, for irregular first person present tense targets which had been preceded by irregular past participles (prime *gelesen* “has run” – target *lese* “I read”), participle priming was less than identity priming by irregular verbs, and thus, only partial. Nevertheless, as noted by Smolka et al. (2007), these effects may have been due to differences in surface frequency across the verb classes tested.

Notably, this experiment used a dichotomous contrast between regular and irregular verbs, even though irregular past participles with and without stem change were not simply conflated to one category as in other studies (e.g. Weyerts et al., 1996). Stimuli were restricted to one subtype of irregular past participles (irregular 1 past participles, e.g. *lesen* “to lesen” – *hat gelesen* “has read”). However, if only two verb types are tested, differences will always appear to be dichotomous (Smolka et al. 2013).

Although, the Dual Mechanism Model does not explicitly state how mixed verbs and sub-types of irregular verbs are represented and processed, it assumes that “the unpredictable must be stored” (Pinker & Prince, 1994, p. 342). Shades of unpredictability are not differentiated. It follows that different types of unpredictable forms, whether mixed verbs or sub-types of irregular verbs, are processed similarly, that is, they are handled by associative memory. Yet, this is an assumption that has yet to be tested empirically.

Connectionist accounts, on the other hand, deny the necessity of symbolic rules and exceptions (Rumelhart & McClelland, 1986) and do not incorporate separate

cognitive mechanisms. Instead, it is suggested that one connectionist network can handle the processing of both regular and irregular forms. In a connectionist network, the speaker’s phonological and semantic knowledge of language is applied to form past tense (Joanisse & Seidenberg, 1999). According to this account, regular and irregular forms rely on phonological and semantic knowledge to a differential degree. For regular verbs in German and English, the mapping from infinitive stem to simple past forms is argued to be primarily dependent on phonological processes (McClelland & Patterson, 2002). Although semantic knowledge is relevant for both regular and irregular verbs, as irregular verbs are limited in phonological consistency between infinitive forms and their past tense, they rely on semantic knowledge much more than regular verbs (Joanisse & Seidenberg, 2005).

A study using event-related potential (ERPs) by Justus et al. (2008) illustrated how important it is to consider various patterns of irregularity. Participants performed an auditory lexical decision task to targets which immediately followed their auditory primes. Experimental stimuli consisted of regular verbs and irregular verbs as well as two additional control conditions. Irregular verbs contained both mixed verbs (e.g. *burn* – *burnt*) and irregular verbs (*sing* – *sang*). The authors used the N400 component for their investigation. The N400 is characterised by a negative going wave about 350 msec after presentation of a meaningful stimulus, e.g. a word or a picture (Kutas & Federmeier, 2011). It is related to the semantic plausibility of the stimulus in the existing context: The more predictable a stimulus, the smaller is the N400 amplitude. Moreover, the N400 has been found to reflect morphological decomposition (McKinnon et al., 2003). Effects of facilitatory priming in ERPs can be seen in an attenuated N400 amplitude (McKinnon et al., 2003; Koester & Schiller, 2008).

Justus et al. (2008) compared ERP waveforms of regular and irregular verbs. They observed significant N400 reductions in the primed condition relative to the unprimed condition both in response to regular and irregular verbs. Regular and irregular verbs were able to prime their present tense verbs. For irregular verbs, the difference between the primed and the unprimed conditions lasted longer compared to regular verbs, i.e. was detectable in the 500–700 msec time window. However, when they factored in the distinction between mixed and irregular verbs, Justus et al. (2008) found that the difference between primed and unprimed condition was much larger for irregular verbs than for mixed and regular verbs. Crucially, for mixed and regular verbs, the difference between primed and unprimed condition was not significantly different.

The authors argued that their results were most compatible with connectionist network models such as that of Joanisse & Seidenberg (1999). According to such model, irregular verbs rely on semantic resources to a larger extent resulting in a larger priming effect.

Initially, most evidence in the past-tense debate came from studies with English native speakers. However, subsequently, data from languages such as German have been taken into account (Clahsen, 1999). In contrast to English simple past tense, German past participle forms involve inflectional affixes for both regular and irregular forms. Thus, regularity is not tied to the presence or absence of affixes and, therefore, has been considered a better test case (Clahsen, 1999; Penke & Westermann, 2006; Cholin et al., 2010).

Smolka et al. (2007) also present data from German native speakers. They propose a model for the processing of verbal inflection which is neither connectionist nor dual route. According to this model, stems which have a similar meaning cluster together in the mental lexicon (e.g. *sing* – *sang* – *sung*). With each stem represented once. For example, a stem such as *kauf* can mean “a purchase” or be part of a past participle form *gekauft* (“has bought”). Upon encountering *gekauft* auditorily or visually, the past participle will be segmented into its constituents *ge-*, *kauf* and *-t*. *Kauf* will activate the meaning of “to buy” at the conceptual level while *ge-* and *-t* will activate the meaning of past. *Kauf* meaning “a purchase”, on the other hand, will also activate the meaning of “to buy” in the mental lexicon. Whether a combination of affixes is acceptable or not is determined by their frequency of occurrence in the language: The strength of connections between stem and affix is a function of the frequency of occurrence of that combination.

To test whether regular and irregular 2 verbs are processed differently in language comprehension, Smolka et al. (2007) used German past participles, illegal combination past participles (e.g. *ge-worf-t* similar to “threw-ed”) and verbal or non-verbal past participles (e.g. *ge-wurf-t*, a combination of the noun “throw” plus *-ed*) as primes in a lexical decision task. They found that regular and irregular 2 past participles facilitated the response to a morphologically related infinitive target verb to the same extent. Moreover, the fact that even illegal combination past participles and non-word stimuli were able to prime morphologically related verb targets supported their conclusion that all past participles irrespective of their verb class are accessed through their stems and are processed by similar mechanisms.

Subsequently, Smolka et al. (2013) investigated whether the degrees of regularity that are seen in the

German verbs are processed in a continuous or categorical way using visual-visual priming and ERPs. Targets were presented as 1st person singular forms and belonged to one of three verb types (regular, irregular 1 verbs *lesen* “to read” – *gelesen* “has read”, irregular 2 verbs *gießen* “to water” – *gegossen* “has watered”). Primes appeared in one of five conditions (prime identical to target: *tanze* “dance”, morphological prime: *getanzt* “has danced”, semantic associate 1st person: *steppe* “tap dance”, semantic associate participle: *gestept* “has tap danced”, unrelated: *wünsche* “wish”). They observed a difference between identity and past participle priming condition in frontal and parietal areas for regular past participles. For irregular 2, differences between identity and past participle conditions were seen in parietal areas. Irregular 1 past participles produced priming effects which were intermediate in amplitude and topographical distribution compared to regular and irregular 2 past participles. These data supported the model proposed in Smolka et al. (2007) and Smolka et al. (2013) suggested, that these graded results result from differences in stem connectivity between regular, irregular 1 and irregular 2 verbs. Stem connectivity is the frequency of occurrence of a stem within a verb’s paradigm. Regular stems, are high in stem connectivity as they occur throughout the paradigm. Stems, such as *sung* in *gesungen* are very low in stem connectivity because the past participle stem only occurs once in the entire paradigm. Importantly, this study highlights the importance of including subtypes of irregular verbs and going beyond the conventional regular-irregular dichotomy.

This brief review illustrates that there is no consensus as to how verbs with different ways of forming past participles are processed. It also shows that a dichotomous division of verbs into regular and irregular, is a matter of experimental convenience and not necessarily psychological reality. Adhering to such a categorisation may give us unreliable evidence and lead to a distorted picture of the processing of different patterns found in natural language (Justus et al., 2008).

1.3. (Ir-)regularity in spoken production

Most evidence in the debate on the processing of verb (ir-)regularity originates from research on language comprehension. Spoken production is inherently more difficult to study as it is hard to control spoken output in the same way as it is possible to control stimuli presented in language comprehension studies (Koester & Schiller, 2008; Harley, 2010). However, language production is an essential part of our ability to use language in everyday life. After all, there is no language comprehension without someone to produce language in the

first place. Thus, it is vital to understand processes of language production, that is, how (ir-)regularity is processed when we speak.

ERPs are not yet commonly used to study language production. However, an ERP study by Budd et al. (2014) used a silent production paradigm to test children's and adult's production of English regular and irregular present and past-tense forms. The participants were shown the infinitive form of a verb and instructed to silently produce 3rd person present tense when a picture of a dog appeared on the screen or the past-tense form when they saw a picture of a dinosaur. Subsequently, an additional cue prompted the participants to produce their response overtly, and only targets with correct overt responses were included in the analysis. The ERP response was time-locked to the cue for silent production (dog or dinosaur). For the adult's production of regular past-tense forms, Budd et al. found a negativity 300–450 msec post-onset relative to the irregular past-tense forms. Yet, no such effect was found for present tense forms.

According to the authors, these effects support the Dual Mechanism Model. However, such differences are explained by connectionist approaches through formal differences between the verb types instead of separate mechanisms (Justus et al., 2008). Budd et al. (2014) argue that the delayed silent production paradigm is “closer to real language production” than picture naming (Budd et al., 2014, p. 2). However, as Ganushchak et al. (2011, p. 1) point out, “covert speech may (...) involve different processes than overt speech production”.

It is clear that more experimental evidence is necessary to answer questions about mechanisms underlying the overt production of regular and irregular verbs as well as their subtypes. This paper reports two experiments. The first experiment investigates the effectiveness of the long-lag priming paradigm for studying the production of verb morphology using naming latency. The second experiment aimed to replicate the first but with greater power. Thus, the second experiment included more items, and items were presented twice throughout the experiment. Finally, in the data collected in these experiments is modelled using a Bayesian analysis.

2. Experiment 1

The goal of Experiment 1 was to evaluate long-lag priming as a method to investigate the mechanisms that underlie language production of German past participles. In the critical condition, prime and target were morphologically related (e.g. prime: *er hat gesungen*

“he has sung”, target: *singen* “to sing”). However, this means that prime and target were necessarily also related with regards to meaning and form. Thus, semantic and phonological relatedness may have contributed to any morphological priming effect. However, it has been shown that if the temporal distance between prime and target is increased, neither semantic nor phonological priming effects survive this lag (Zwitsers et al., 2002; Dohmes et al., 2004; Koester & Schiller, 2008, 2011).

This experiment employed a primed picture naming task (Koester & Schiller, 2008), in which German native speakers overtly named black-and-white line drawings using the infinitive form of the verb to describe the action displayed. Importantly, picture naming “engages the full production process, from conceptualization to articulation” (Tabak et al., 2010) and is, thus, considered a suitable method to study language production.

As noted above, past participles vary in morphological structure in that a stem may take either a productive or non-productive affix and may or may not show stem changes. For this reason, primes consisted of past participles or present tense forms belonging to different verb classes (see Table 1).

2.1. Methodology

2.1.1. Participants

37 native speakers of German participated in the study (11 male). Participants were an average of 25.6 years of age ($SD = 5.4$; age range = 19–44) and had normal or corrected to normal vision. They were given either monetary compensation or course credit for their participation.

2.1.2. Materials

2.1.2.1. Targets. The verbs denoting target pictures fell into three categories, that is, regular verbs and two different types of irregular verbs, i.e. irregular 1 and irregular 2 verbs (see Table 1, and Appendix A Table A1 for a full list of targets). Verbs from each verb class were matched for lemma frequency (Heister et al., 2011), number of phonemes, number of syllables and number of letters. A total of 48 verbs were included in the experiment (see Table 2 for matching properties of target verbs).

Black-and-white line drawings for all 48 simple verbs were created by an artist (see Figure 1 for an example). The drawings were tested for name agreement using a web-based questionnaire with native speakers of German. Participants were instructed to look at each picture and to fill in the verb that described the action displayed in the picture. Participants used the

Table 2. Properties of targets and primes of Experiment 1.

Relatedness	Prime	Target	Target			Prime				
			Log Lemma Freq. (SD)	Letters (SD)	Mean no. of Phonemes (SD)	Syllables (SD)	Log Lemma Freq. (SD)	Letters (SD)	No. of Phonemes (SD)	Syllables (SD)
Morph.	<i>getanzt/tanzt</i> "has danced/dances"	<i>tanzen</i>	1.74	5.65	4.52	1.06	1.7 (0.6)	5.65 (1.1)	4.52 (0.76)	1.06 (0.24)
Semantic	<i>gesteppi/stept</i> "has tap danced/ tap dances"						0.58 (0.58)	7.3 (1.8)	6.33 (1.5)	1.9 (0.7)
Phon.	<i>getarnt/tarnt</i> "has disguised/disguises"	"to dance"	(0.6)	(1.1)	(0.76)	(0.24)	0.88 (0.7)	6.8 (1.5)	5.8 (1.36)	1.7 (0.62)
Unrel.	<i>gebeichtet/beichtet</i> "has confessed/confesses"						0.9 (0.71)	7.14 (1.8)	6.3 (1.54)	2 (0.72)

**Figure 1.** Example of a black-and-white line drawing used as target picture.

comment section that was provided to write down alternative verbs they considered. Their feedback was used to revise the drawings and tested again. Three web-based experiments with a total of 104 participants were performed.

2.1.2.2. Primes. Prime-target pairs were related morphologically, semantically, phonologically or unrelated and primes were presented in present tense or past participle (see Table 2).

If available, minimal pairs were selected as phonologically related primes (e.g. *rauschen* "to swoosh" - *rauchen* "to smoke"; Zwitserlood et al., 2000). If no minimal pair was available, words that overlapped in onset and vowel (e.g. *schreiben* "to write" - *schreinern* "to do wood-work") or that rhymed (e.g. *schwimmen* "to swim" - *stimmen* "to tune") were selected instead. Semantically related primes were related by hyponymy (e.g. *schwimmen* "to swim" *kraulen* "to swim the crawl") or by belonging to the same semantic field (*pfeifen* "to whistle" - *flöten* "to play the flute").

Semantic associations between semantically related, phonologically related and unrelated primes and their targets were tested in a web-based association test. Participants were asked to rate each of the 144 prime-target pairs on a scale of one (unrelated) to seven (almost identical). For semantically related primes, even near synonyms of the target verb were not considered to be suitable primes because they may have been potential names for the target. In order to be able to identify near synonyms, participants were asked to rate a verb "seven" only if they thought the meaning was almost identical. The instructions emphasised that participants should rate the prime-target pair quickly one after

another without thinking for too long. For mean semantic relatedness by condition see Appendix A Table A2.

2.1.2.3. Fillers. Fillers consisted of unrelated written verbs in present tense, past participle, infinitive form and black-and-white line drawings of unrelated actions. For each prime-target pair a pool of nine unrelated filler items and one unrelated filler drawing was created.

Fillers had no phonological overlap with the respective target verb and were not semantically related. Occasionally, filler verbs occurred twice throughout the experiment for different verb-target pairs. In such cases, a different form of the verb was used (e.g. infinitive, present tense or past participle). Since filler items were selected in a way that they were not related to the respective target verb, there is no reason to assume that repeated exposure to a filler item affected naming latency to a target picture.

2.1.3. Design

Morphologically related and unrelated prime-target pairs were separated by five to eight filler trials (see Figure 2) (Zwitserslood et al., 2002). Morphological priming was evaluated in reference to the baseline condition, i.e. the unrelated prime-target condition. To test whether semantic and phonological similarity between prime and target had any effect on the response latencies, semantically and phonologically related prime-target pairs were included in the experiment.

2.1.4. Apparatus

Participants were tested individually in a darkened room sitting 60–70 cm from the computer screen. Participants wore a microphone. Stimuli were presented in the middle of the screen in white lower case letters on a grey background. The experimental software was written in Python 2.7.3, was run on a Dell XPS 13 laptop running Ubuntu 12.04. and presented on a 1920 × 1080 pixel computer screen.

2.1.5. Procedure

After the participant had given informed consent, they completed a questionnaire providing personal background information. The experiment was preceded by a familiarisation phase during which the participant was familiarised with the drawings they were going to see in the experiment. After receiving instructions, participants saw the line drawing with the target verb written underneath.

Participants were instructed to look at the picture, read aloud the target verb and to press the button to move to the next picture. Familiarization was self-paced and participants were asked to complete the task at their own pace but in a timely manner. Depending on the participant, familiarisation took about four to five minutes. After completing familiarisation, participants received written instructions for the primed picture naming task. Participants were told to use simple verbs in the infinitive to name the pictures which appeared on the screen as they had in the familiarisation phase and to read aloud any text exactly as it appeared on the screen. Participants were instructed to respond as soon as a word or picture appeared. In order to get used to the task, participants completed 20 practice trials and had the opportunity to ask questions afterwards. During the experiment, the timing was identical for trials showing a prime, target or filler: an asterisk appeared for 250 msec followed by a blank screen for 250 msec. Then the stimulus item (a target picture, a prime or a filler) appeared and remained on the screen for 2000 msec during which time the participant was required to respond. Koester & Schiller (2008) allowed between 1400 to 1700 msec for a response. Since overt action naming is known to be slower than object naming (Szekely et al., 2005), we allowed 2000 msec for the response. Each participant saw 48 prime-target pairs in one of the four prime-target conditions (morphologically, semantically, phonologically related and unrelated) with five to eight filler trials between prime and target. Five to eight fillers were randomly selected from the respective pool of filler items available

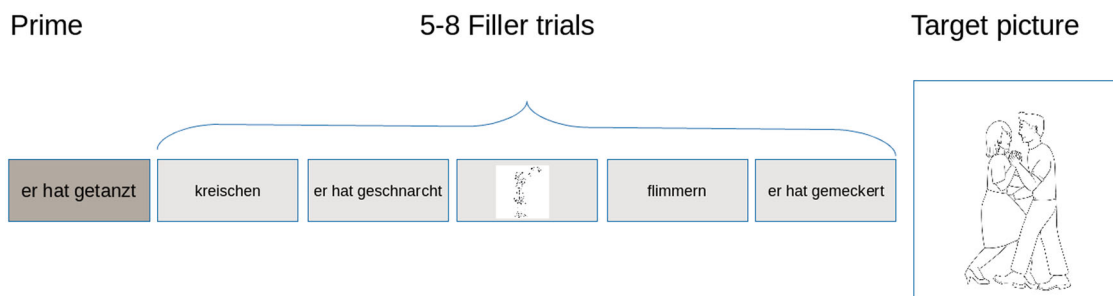



Figure 2. An experimental trial consisted of a prime word, five to eight filler trials (words or line drawings) and a target picture.

Table 3. Mean naming latencies, accuracy and naming latency difference for different levels of relatedness relative to the unrelated condition.

Relatedness	Prime	Target <i>tanzen</i> ("to dance")	Mean Lat. (SE)	Accuracy in %	Δ Lat.
Morphological	<i>tanzt</i> ("dances")		1145 msec (10)	94.82	18 msec
Semantic	<i>stept</i> ("tap dances")		1179 msec (10)	92.34	-16 msec
Phonological	<i>tarn</i> ("disguises")		1173 msec (11)	92.34	-10 msec
Unrelated	<i>beichtet</i> ("confesses")		1163 msec (10)	91.89	-

Note: SE = Standarderror; Δ Lat. = Primingeffect.

for each prime-target pair for each participant and on each experimental trial. A participant saw each target picture only once. The 48 experimental trials were distributed over three blocks with two breaks in between. No feedback was given during the experiment. As stimulus presentation was implemented by means of a Python script, this script was also used to randomise prime-target pairs for each participant. The entire primed picture naming task took 25 minutes.

2.1.6. Data analysis

Participant's responses were checked for accuracy. False starts, time out of a response, overlaps and incorrect responses were excluded from the analysis (9.35% of all trials). If the participant started to utter the beginning of a word that was different from the target verb but corrected him/herself this was classified as a false start. Overlaps occurred if responses to preceding filler trials were too slow and overlapped with the recording of the target trial. A response was classified as timeout if it took longer than 2000 msec for the participant to respond and none or only the onset of a word was captured in the recording. An incorrect response was any response which deviated from the target form. Praat (Version 5.3.16; Boersma & Weenink, 2009) was used to measure response latencies from onset of the target picture until the voice onset. The person measuring the naming latencies was blind to the experimental condition of each trial. Naming latency distributions are often positively skewed and, thus, violate the assumption of normality which is a prerequisite of frequentist linear mixed models (Kliegl et al., 2010). Hence, the Box-Cox function (Box & Cox, 1964) which is built into the R package MASS (Venables & Ripley, 2002) was used to identify the type of transformation that would result in the data fitting the normal distribution best. The Box-Cox procedure and the inspection of the quantile-quantile plots suggested that the Inverse-Gaussian transformation was the most suitable method. Transforming the naming latencies preserves ordinal relation of means and does not change the direction of the effects or the significance of main effects (Kliegl et al., 2010).

To assess the research questions, a frequentist linear mixed model (LMMs) was fitted using R (R Core Team, 2013) and the lme4-package (Bates et al., 2014). A LMM

allows the specification of "subjects" and "items" as random factors in a single LMM and as such one model replaces two separate F1- and F2 ANOVAs (Kliegl et al., 2011). In addition, LMMs "suffer less severe loss of statistical power if an experimental design loses balance due to missing data" (Kliegl et al., 2011). This was important in the current experimental design as data loss was expected due to incorrect responses in picture naming.

2.2. Results

Our research question asked whether Relatedness (morphological, semantic, phonological and unrelated) affected naming latencies of target pictures in a long-lag priming paradigm. That is, did morphologically, phonologically, semantically or unrelated primes facilitate or inhibit RTs to target pictures? Mean naming latencies, standard error SE, mean accuracy and naming latency difference can be seen in Table 3. The overall accuracy was 92.85 %.

We fitted a frequentist linear mixed model which had inverse transformed negative RT as the dependent variable. Relatedness was modelled with contrast coding using sum contrasts. Morphologically, semantically and phonologically related prime-target pairs were each compared to unrelated prime-target pairs (see Appendix B Table B1 for the contrast matrix). Time reference frame and verb type were not included in the analysis. The random effect structure included random intercepts for participants and items as well as random slopes for participants. Since the number of fillers intervening between prime and target varied (five to eight trials), "number of fillers" was included as a covariate into the model along with name agreement and logarithmic lemma frequency of the target.

Target pictures preceded by a morphologically related prime were named significantly faster than those preceded by unrelated primes ($\beta = 0.02, SE = 0.01, t = 2.61$). However, no significant priming effects were found for phonologically related prime-target pairs ($\beta = -0.004, SE = 0.01, t = -0.49$) or semantically related prime-target pairs ($\beta = -0.01, SE = 0.01, t = -1.26$). If a target picture had higher name agreement scores in the web based

Table 4. Estimates and coefficients of linear mixed model 1 (Experiment 1).

Contrast	Estimate	SE	t-value
(Intercept)	-0.65	0.05	-12.35
Unrelated – Morphologically related	0.02	0.01	2.61
Unrelated – Phonologically related	-0.004	0.01	-0.49
Unrelated – Semantically related	-0.01	0.01	-1.26
Name agreement	-0.3	0.06	-5.33
Fillers	0.01	0.002	1.98

test, participants responded faster ($\beta = -0.3$, $SE = 0.06$, $t = -5.33$) and the smaller the number of fillers between prime and target the faster the response. However, this effect was only marginally significant ($\beta = 0.01$, $SE = 0.002$, $t = 1.98$).¹ Log lemma frequency had no significant effect on naming latencies. Therefore, it was dropped as a covariate; this did not change the effects. The numbers reported here are, thus, based on a model with name agreement and number of fillers as covariates. A summary of the models coefficients can be found in Table 4.

2.3. Discussion

In this experiment we addressed whether morphological priming effects could be found in language production when the distance between prime and target was five to eight trials. Semantic and phonological prime-target pairs were included to test whether semantic and/or phonological similarity between prime and target in the morphological condition may have contributed to morphological priming effects.

Pictures were named faster when they had been preceded by morphologically related primes, which was expected. By reading aloud the morphologically related prime word, (e.g. *er hat getanzt* “he has danced”, (constituent) morphemes are activated. So when the target picture, e.g. *tanzen* “to dance”, is to be named, the (constituent) morphemes are hypothesised to still be active and facilitate naming. Our results replicate the results by Zwitserlood et al. (2000) and Koester & Schiller (2008) who found morphological priming effects in priming studies of compounds with similar lags between prime and target.

In contrast to morphological priming, semantic and phonological priming effects have been found not to last over the temporal distance created by the lag of five to eight trials between prime and target (Zwitserlood et al., 2000; Koester & Schiller, 2008). Our results, likewise, showed no semantic or phonological priming effects. However, the absence of such an effect does not provide evidence in favour of the null hypothesis. To test the plausibility of the null hypothesis, we, therefore, also performed a Bayesian analysis (see Section 4.1).

3. Experiment 2

The main goal of Experiment 2 was to test the effect of regularity on naming latencies. Since no semantic or phonological priming was found in Experiment 1 we assume that the long-lag priming paradigm is effective and eliminates semantic and phonological contributions to morphological priming. Hence, we included a morphologically related and unrelated condition and past participle primes.

As described in the introduction, the Dual Mechanism Model assumes that two distinct mechanisms are responsible for the processing of regular and the two types of irregular verbs (Clahsen, 1999) whereas alternative approaches such as connectionist networks or the model by Smolka et al. (2007) do without separate mechanisms.

If differences in priming patterns for regular and the two types of irregular verbs were observed, this would provide evidence for distinct mechanisms underlying their processing (Sonnenstuhl et al., 1999). Similar priming patterns, on the other hand, would speak in favour of a single system. The Dual Mechanism Model (Clahsen, 1999; Pinker & Ullman, 2002) predicts facilitatory priming effects for regular targets which have been preceded by morphologically related primes: reading aloud a morphologically related prime will activate constituent morphemes (stem + affix) and, speed naming the target picture in the infinitive. Irregular 1 and irregular 2 verb forms, on the other hand, are hypothesised to be stored in lexical entries separate from their stem form. Reading aloud an irregular 1 or irregular 2 prime would activate its infinitive stem only indirectly and, therefore, should result in reduced or no facilitation for the target verb in the morphological condition (Sonnenstuhl et al., 1999; Clahsen, 1999), irrespective of whether there is a stem change.

Although the model of Smolka et al. (2007) has been proposed as a model for visual word recognition, it can also be used to predict priming patterns in language production (Smolka et al., 2007, 2013). In the model, both regular and irregular verbs are accessed through their stem and, hence, a morphological priming effect is predicted for regular, irregular 1 and irregular 2 verbs. Smolka et al. (2013) point out that regular, irregular 1 and irregular 2 verbs differ in stem connectivity, again, the frequency of occurrence of a stem within verb’s paradigm. Yet, the graded patterns reported by Smolka et al. showed up only in the ERPs but not in RTs. We argue, however, that, if stem connectivity is a decisive factor, it’s effects should be observable in naming latencies as well. As noted previously, in connectionist models, a single mechanism processes regular and irregular

Table 5. Predicted priming pattern for regular, irregular 1 and irregular 2 past participles.

Approach	Priming effect
Dual Mechanism Model	Regular > (Irregular 1 = Irregular 2) <ul style="list-style-type: none"> ◦ facilitatory priming for regular targets ◦ reduced or no priming for irregular verbs ◦ no difference between irregular 1 and 2 past participles
Connectionist models	Regular = Irregular 1 = Irregular 2 <ul style="list-style-type: none"> ◦ no morphological priming ◦ the same for all verb types
Smolka et al. (2007)	Regular = Irregular 1 = Irregular 2 <ul style="list-style-type: none"> ◦ morphological priming ◦ the same for all verb types ◦ a graded pattern is indicative of the influence of stem connectivity

forms (Joanisse & Seidenberg, 1999). However, connectionists claim that irregular verbs rely more on the semantic knowledge of the speaker and, hence, profit more from a semantic relationship between prime and target than regular verbs. The processing of regular verbs is hypothesised to depend more on the phonological knowledge of the speaker. Therefore, priming effects arise either due to phonological similarity (i.e. for regular and irregular 1 primes and targets: *getanzt* – *tanzen*, *gelesen* – *lesen*) or semantic links (i.e. for irregular 2 primes and targets: *gesungen* – *singen*). No separate morphological level of processing is assumed under a connectionist view. Yet, the long-lag paradigm which we adopted is supposed to eliminate semantic and phonological priming effects (Zwitsers et al., 2002; Dohmes et al., 2004; Koester & Schiller, 2008). Consequently, connectionist accounts predict no effects for any of the verb types. The predictions are the same as in Experiment 1 and are summarised in Table 5.

3.1. Methodology

3.1.1. Participants

49 participants aged between 20 and 39 took part in the study (mean age = 26.6; 26 men and 23 women). Data from three participants were excluded because they did not follow the instructions. Participants were right handed as determined by the Edinburgh Lateralality Index (Oldfield, 1971). They had no history of neurological or mental illness and either normal or corrected-to-normal vision. Participants were reimbursed for their participation.

3.1.2. Materials

3.1.2.1. Targets. The same target pictures were used as in Experiment 1. However, to increase the overall number of items, 23 additional black-and-white line

drawings were created resulting in 71 items (regular: $n=25$, irregular 1: $n=21$, irregular 2: $n=25$). The 71 targets were presented twice throughout the experiment. All were matched for log lemma frequency using dlexDB (Heister et al., 2011). Again, target verbs were balanced as far as possible across sets for word length (number of letters, number of phonemes and number of syllables; see Table 6). All 71 items were tested for name agreement with 75 German native speakers in a web-based questionnaire. They were asked to look at each drawing and to name it by typing the corresponding verb. Participants were asked to enter alternative verbs which they considered in cases they were not able to decide.

3.1.2.2 Primes. Unrelated primes were chosen such that they belonged to the same verb type (i.e. irregular 1 prime – irregular 1 target, regular prime – regular target). In a few cases ($n=14$), irregular targets could not be paired up with an appropriate unrelated prime of the same verb type. In such cases, irregular (n -) participles were chosen from another irregular verb type (i.e. unrelated irregular 2 prime – irregular 1 target).

3.1.2.3 Fillers. Filler items, which were used to create the lag between prime and target, were either verbs (2nd person sg. present tense, past participles, infinitive verbs) or black-and-white line drawings. The same criteria as in Experiment 1 were used to create the filler items.

3.1.3. Design

The design of Experiment 2 was similar to that of Experiment 1: primed picture naming was used in a within subject design. Primes were presented five to eight filler trials before the target (Koester & Schiller, 2008, see also Experiment 1). Regularity of the target verb was a between-items factor with three levels (regular, irregular 1, irregular 2). In Experiment 2, Relatedness was a within-items factor with only two levels (morphologically related/unrelated). Naming latencies were measured. The unrelated condition served as a baseline condition to evaluate morphological priming in the naming latencies.

3.1.4. Apparatus

A Python script (Python 2.7.3) was used to present visual stimuli and to record participants' overt responses. A Sennheiser PC31 II headset which the participant wore around the neck was used for the recording.

Table 6. Properties of targets and primes of Experiment 2.

Verb Class	Relatedness	Prime	Target	Target				Prime			
				Mean no. of		No. of		Log Lemma		No. of	
				Letters (SD)	Phonemes (SD)	Syllables (SD)	Frequency (SD)	Letters (SD)	Phonemes (SD)	Syllables (SD)	Frequency (SD)
Regular	Morphological	<i>getanzt</i>	<i>tanzen</i> "to dance"	6.36 (0.86)	5.52 (0.71)	2 (0)	1.44 (0.53)	7.52 (0.96)	6.56 (0.96)	2.16 (0.37)	
	Unrelated	<i>gebeichtet</i>					1.4 (0.6)	7.48 (1.12)	6.8 (1.08)	2.24 (0.44)	
Irregular 1	Morphological	<i>gebraten</i>	<i>braten</i> "to fry"	6.1 (0.83)	5.33 (0.58)	2 (0)	1.96 (0.71)	8.1 (0.83)	7.43 (0.6)	3 (0)	
	Unrelated	<i>gekommen</i>					1.81 (0.61)	8.76 (1.26)	7.43 (0.51)	3 (0)	
Irregular 2	Morphological	<i>gesungen</i>	<i>singen</i> "to sing"	6.76 (1.27)	5.36 (0.57)	2 (0)	1.73 (0.65)	8.56 (1.26)	7.44 (0.65)	3 (0)	
	Unrelated	<i>geschmolzen</i>					1.36 (0.73)	8.68 (1.07)	7.52 (0.65)	3 (0)	

3.1.5. Procedure

After giving informed consent, participants completed a familiarisation phase to become familiar with the picture stimuli used in the experiment. Participants saw a black-and-white line drawing which had the target verb written underneath. Both experimental drawings and 55 filler drawings were presented. The participants were instructed to look at each drawing and then read aloud the verb which was written below. Familiarization was self-paced, although participants were instructed to go through the drawings in a timely manner.

After completing the familiarisation, there was a 20 to 25 minute break, before the participants received the written instructions. The participants were instructed to respond as soon as a word or drawing appeared on the screen by overtly naming the drawing using a simple verb or by reading aloud any words exactly as they appeared. 20 practice trials which were identical to the trials of the experiment preceded the experiment and the participant had a chance to ask questions.

The timing was the same for trials showing a prime, target picture or a filler item. A fixation cross appeared for 250 ms to centre the gaze of the beginning of each trial. It was followed by a blank screen for 250 ms. Then the stimulus item (word(s) or a drawing) appeared for 2000 ms during which the participant was to respond. Written stimuli were presented in the middle of the screen in white Serif lower case on a grey background. Pictures were presented on a grey screen in constant size (900 × 900).

The 71 prime-target pairs were shown in two prime-target conditions (morphologically related/unrelated) with five to eight filler trials separating prime and target. The python script that was used for stimulus presentation was also used to ensure that the order of presentation of conditions was counterbalanced both for each participant and across participants. This was done such that, the presentation of morphologically related and unrelated conditions was pseudorandom for each participant. Whether a morphological or unrelated prime-target pair was presented first was also randomized across participants. Moreover, prime-target pairs were presented in a different randomised order for each participant. 142 experimental trials were distributed over seven blocks with six breaks in between. A target never appeared twice in the same block. No feedback was given during the experiment. Each prime-target pair had a pool of 10 unrelated filler items (pictures or words). Five to eight fillers were chosen randomly from that pool for each subject. The entire primed picture naming task took about 55 minutes (excluding breaks, which depended on each individual).

3.1.6. Data analyses

Scoring of participants responses was identical to Experiment 1. After checking the responses for their accuracy, 9.11% of the data points were classified as incorrect and did not enter further data analysis.

In addition to scoring for accuracy of the targets, primes and filler pictures were checked for their accuracy as well. Trials were rejected as incorrect if a prime was not read aloud properly or if filler pictures were named incorrectly and resulted in a related form that could have served as prime. Thus, an additional 2.97 % were removed. The Box-Cox procedure as well as the inspection of the quantile-quantile plots revealed that the Inverse Gaussian transformation would be the most appropriate. Transforming naming latency data does not change the ordinal relation of means and does not alter the effects or significance of main effects (Kliegl et al., 2010).

3.2. Results

3.2.1. Main effect of priming

One goal of this experiment was to replicate the effect of Relatedness found in Experiment 1. The data can be seen in Table 7 and Figure 3. The overall accuracy was 87.9%.

To test whether Relatedness had an effect on the naming latencies, we fitted a frequentist linear mixed model with inverse transformed negative RT as dependent variable. Analogous to the model in Experiment 1, we modelled Relatedness using sum contrasts, that is, unrelated prime-target pairs were compared to morphologically related prime-target pairs (for contrast matrix see Appendix B Table B2). Random intercepts for participants and items and random slopes for participants were included as random effects. Name agreement, number of fillers, lemma frequency and presentation (first/second) were included as additional predictors.

Again, a significant main effect of Relatedness was found ($\beta = 0.04, SE = 0.01, t = 7.6$). The model's coefficients can be found in Table 8. As in the previous experiment, name agreement scores were a significant predictor of naming latencies. That is, the higher name agreement scores were, the faster a picture was named ($\beta = -0.4, SE = 0.06, t = -6.4$). Unsurprisingly, naming latencies were significantly faster on the second

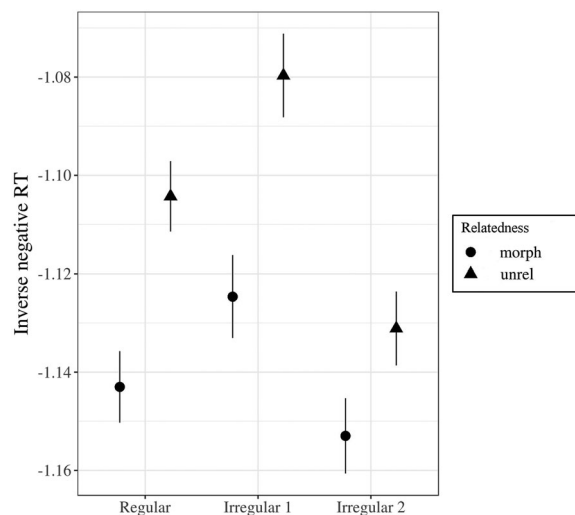


Figure 3. Mean RTs for primed and unprimed prime-target pairs by verb class for Experiment 2.

presentation ($\beta = -0.12, SE = 0.01, t = -25.2$). Neither number of fillers nor lemma frequency showed a significant effect. As these predictors did not improve the fit of the model they were dropped. This did not influence the effects. Hence, the coefficients are based on the model with presentation and name agreement as predictors.


3.2.2. Main effect of verb type

In the second analysis, we were interested in priming patterns by verb type. Table 9 provides descriptive measures.

We fitted a linear model with inverse negative naming latencies as the dependent variable. The factors Relatedness (morphological/unrelated) and verb type (regular, irregular 1 and irregular 2 verbs) were independent variables. We compared irregular 2 to irregular 1 verbs and then irregular 2 and irregular 1 verbs taken together to regular verbs (Helmert contrast).

The relevant comparison involving verb type was the interaction between the main effect of Relatedness and irregular 2 versus irregular 1 verbs. Moreover, the interaction between Relatedness and irregular 2 versus irregular 1 taken together and regular verbs was also included (see Appendix B Table B3 for the contrast matrix). Name agreement, presentation, lemma

Table 7. Mean target naming latencies, accuracy and naming latency difference for morphologically related and unrelated trials for Experiment 2.

Relatedness	Prime	Target <i>tanzen</i> ("to dance")	Mean Lat. (SE)	Accuracy in %	Δ Lat.
Morphological	<i>tanzt</i> ("dances")		921 msec (4)	87.72	30 msec
Unrelated	<i>beichtet</i> ("confesses")		951 msec (4)	87.14	–

Note: SE = Standarderror; Lat. = Primingeffect.

Table 8. Estimates and coefficients of linear mixed model 2 (Experiment 2).

Contrast	Estimate	SE	t-value
(Intercept)	-0.6	0.06	-11
Unrelated – Morphologically related	0.04	0.01	7.6
Name agreement	-0.4	0.06	-6.4
Presentation	-0.12	0.01	-25.2

frequency of the target and the number of fillers intervening between prime and target were included as predictors. Except for the main effect of Relatedness ($\beta = 0.04, SE = 0.01, t = 7.67$), none of the other comparisons or interactions were found to be significant (see Table 10).

3.3. Discussion

In contrast to Experiment 1, Experiment 2 tested only morphologically related and unrelated prime-target pairs. Drawings, which were preceded by morphologically related words, were named significantly faster than drawings which were preceded by unrelated words. Thus, we are able to reject the null hypothesis that the means in the unrelated and morphologically related condition do not differ.

This result was expected because it replicates both the reports in the literature as well as the findings of Experiment 1 (Zwitsers et al., 2000; Koester & Schiller, 2008). Therefore, we conclude that the long-lag priming paradigm works well and that morphological priming effects can survive a lag of five to eight trials.

Although we found a robust effect of morphological priming, we did not find any significant difference




Table 10. Estimates and coefficients of linear mixed model 3 (Experiment 2).

Contrast	Estimate	SE	t-value
(Intercept)	-0.6	0.06	-10.54
Relatedness	0.04	0.01	7.67
Irregular 2 – Irregular 1	0.002	0.03	0.05
Regular – (Irregular 2 & Irregular 1)	0.03	0.03	1.05
Relatedness × (Irregular 2 – Irregular 1)	-0.01	0.01	-1.12
Relatedness × (Regular – (Irregular 2 & Irregular 1))	-0.004	0.01	-0.72
Name agreement	-0.41	0.07	-6.34
Presentation	-0.12	0.01	-25.2

between the verb types. This outcome cannot be explained by the Dual Mechanism Model (Clahsen, 1999) nor by Connectionist models (Joanisse & Seidenberg, 1999). They presuppose similar processing mechanisms, and hence, it should be possible to prime all three verb types to the same degree. However, connectionist models attribute this to semantic and phonological priming effects, which we can rule out due to the lag between prime and target (Zwitsers et al., 2000). The model proposed in Smolka et al. (2007) can account for these results.

However, before possible explanations are discussed, we present Bayesian linear mixed models. Bayesian statistics allows us to model how likely a scientific hypothesis is given the experimental data. This is especially important for those results of Experiment 1 and 2, which did not show a significant difference between conditions. The Bayesian analysis will allow us to determine the confidence with which we can say that our data support the null hypothesis that there is no difference between the conditions.

Table 9. Mean target naming latencies, accuracy and naming latency difference for primed and unprimed prime-target pairs by verb class for Experiment 2.

Verb type	Condition		Target	MeanLat.(SE)	Accuracy in %	Δ Lat.Unrel.–Morph.
	Prime	Relatedness				
Regular	Morph.	tanzt/getanzt ("dances/has danced")	tanzen ("to dance") 	915 msec (7)	87.83	34 msec
	Unrel.	beichtet/gebeichtet ("confesses/has confessed")		949 msec (7)	86.52	–
Irregular 1	Morph.	liest/gelesen ("reads/has read")	lesen ("to read") 	937 msec (8)	86.23	43 msec
	Unrel.	bgelt/gebgelt ("irons/has ironed")		979 msec (9)	85.71	–
Irregular 2	Morph.	gießt/gegossen ("waters/watered")	gießen ("to water") 	913 msec (7)	88.87	16 msec
	Unrel.	schmuggelt/geschmuggelt ("smuggles/has smuggled")		929 msec (7)	88.96	–

Note: Morph. = Morphologically related; Unrel. = Unrelated; Lat. = Priming effect.

4. Bayesian analysis of naming latencies

In frequentist hypothesis testing, the question that is asked is whether the null hypothesis can be rejected or not; the p -value is the probability of getting an effect as extreme or even more extreme than the observed results under the assumption that the null hypothesis is true. Thus, in the frequentist paradigm, we can only obtain evidence *against* the null hypothesis. Neither can we find evidence in favour of the null nor can we infer anything about the actual effect size. The Bayesian paradigm, by contrast, offers a way to evaluate how convincing a hypothesis is given the experimental data at hand. From the likelihood of the data and our prior belief about plausible effect sizes, a probability distribution can be computed over possible effect sizes given the observed data by using Bayes' Theorem. This probability distribution is referred to as *posterior distribution* of the experimental effect. From this posterior distribution, we can directly calculate how likely it is that the effect size lies in any given interval. Conversely, we can also compute a so-called *95% credible interval*, which provides the interval within which we can be 95% certain that the real effect size lies (note that the frequentist confidence interval does not provide any information about plausible effect sizes, see Morey et al., 2015). Our prior knowledge or belief about the effect size is also expressed in a probability distribution over possible effect sizes, referred to as *prior distribution* of the experimental effect. This prior can reflect a belief about plausible effect size based, for example, on earlier research. Alternatively, in case one does not have prior knowledge about the expected effect size, one can use an *uninformative* prior, which in general is a widely spread out distribution reflecting uncertainty about the effect size. The posterior distribution is computed by applying Bayes' Theorem: the posterior distribution is proportional to the prior distribution multiplied with the likelihood of the observed data (i.e. the probability of the observed data dependent on the model parameter, or here, the effect size in question).

There is not a single standard way to perform hypothesis testing within the Bayesian paradigm. One common approach is the usage of the *Bayes factor* (BF). The Bayes factor compares the support for one model (e.g. the effect size according to the alternative hypothesis) over another model (e.g. the null hypothesis). Technically, the Bayes factor is the ratio of the probability of the observed data given the first model (i.e. the *likelihood* of the first model) and the probability of the observed data given the second model. Importantly, the Bayes factor only quantifies how much one of the two models is superior over the other; it does not provide

Table 11. Overview of Bayesian and corresponding frequentist models.

Bayesian model	Data	Comparisons	Corresponding frequentist model
1	Experiment 1	Unrel. – Morph. Unrel. – Sem. Unrel. – Phon. Name agreement Number of Fillers	Frequentist linear mixed model 1
2	Experiment 2	Relatedness (Unrel.–Morph.) Irreg. 2 – Irreg. 1 Regular – (Irreg. 2 & Irreg. 1) Relatedness × (Irreg. 2 – Irreg. 1) Relatedness × (Regular – (Irreg. 2 & Irreg. 1)) Name agreement	Frequentist linear mixed model 3

any information about the correctness of a model per se. If one is more interested in what a plausible model (here, effect size) would be, this information can be directly obtained by inspecting the posterior distribution. For a tutorial on Bayesian data analysis in psycholinguistics, see Nicenboim & Vasishth (2016) and Sorensen et al. (2016).

We modelled negative reciprocal naming latencies (multiplied by the factor 1000) as the dependent variable by fitting two Bayesian linear mixed models using R (R Core Team, 2013) together with the probabilistic programming language Stan, Version 2.14.1. (Stan Development Team, 2016).

First, we modelled the naming latencies observed in Experiment 1 with Relatedness, name agreement and number of fillers as predictors. The aim of this analysis was to establish the validity of the experimental design used in the experiments presented in this paper. As discussed above, the assumption behind our experimental design was that (i) morphological priming does exist despite the relatively long lag between prime and target and (ii) semantic and phonological priming does not survive the lag between prime and target.

Second, we fitted a model to the data of Experiment 2 to test whether morphological priming differed between verb types. To achieve this aim, fitted a model to the naming latencies of Experiment 2 (see Table 11 for an overview of the comparisons).

4.1. Model 1: the effect of relatedness

In the frequentist analysis of Experiment 1 (Model 1, see Section 2.2), we found a statistically significant facilitatory morphological priming effect. Semantic and phonological prime-target pairs, in contrast, did not show a significant difference between the

unprimed and the primed trials. In other words, while the null hypothesis could be rejected for the difference between morphologically related and unrelated trials, we were not able to reject the null hypothesis regarding the difference between semantically related and unrelated prime-target pairs and for the difference between phonologically related and unrelated prime-target pairs. Importantly, within the frequentist paradigm, we cannot interpret the failure to reject the null hypothesis as evidence for the absence of a semantic and phonological priming effect. A Bayesian analysis, by contrast, allows us to directly test how plausible the null hypothesis (or any other effect size) is given our data.

In the first Bayesian linear mixed model, we included the same predictors as in model 1 of the frequentist analysis (see Appendix B Table B1 for a summary of the contrast coding): we compared morphologically related (coded as -1) with unrelated (coded as $+1$) conditions, semantically related (coded as -1) with unrelated (coded as $+1$) conditions and phonologically related (coded as -1) with unrelated (coded as $+1$) conditions. Name agreement and the number of fillers intervening between prime and target were included as additional predictors. As in the frequentist analysis, random intercepts were fit for both participants and items, and random slopes were fit for participants only.

As there is not much literature available on primed picture naming of verbs, we chose to use only weakly informative priors. Remember that naming latencies were transformed to negative reciprocals multiplied by 1000. As prior distribution for the intercept, we used a normal distribution centred around -1 , which corresponds to 1000 ms on the scale, with variance 1: $\hat{\beta}_{intercept} \sim N(-1, 1)$. We used -1 as the mean of the prior as 1000 msec is a rough guess about mean naming latencies. Importantly, this number will not have much of an influence on the estimation of the posterior distribution as the prior's variance of 1 (corresponding to 1000 msec on the scale) is very large. For the three priming effects (morphological priming effect $\hat{\beta}_{morph}$, semantic priming effect $\hat{\beta}_{sem}$ and phonological priming effect $\hat{\beta}_{phon}$) as well as for the effects of name agreement and number of intervening fillers, we used the same prior distribution, namely a normal distribution centred around 0 with variance 0.1, i.e. $\hat{\beta} \sim N(0, 0.1)$. As prior for the variance-covariance matrix of the predictors, we also used $\hat{\beta} \sim N(0, 0.1)$ as prior distribution.

An overview of the results is provided in Table 12. The mean of the posterior distribution of the morphological priming effect $\hat{\beta}_{morph}$ is 0.02 and the probability of the morphological priming effect being larger than zero P ($\hat{\beta}_{morph} > 0$) is 0.99.

Table 12. Mean of the posterior distribution, posterior probability of the effect being greater than 0, probability of the effect being zero or negligibly small, 95% credible interval and Bayes factor for the effect of Relatedness in Experiment 1.

Comparison	mean($\hat{\beta}$)	P($\hat{\beta} > 0$)	95% CrI	BF
Unrelated – Morphologically related	0.02	0.99	[0.004, 0.03]	BF ₀₁ = 0.64
Unrelated – Phonologically related	-0.004	0.27	[-0.02, 0.01]	BF ₀₂ = 12.7
Unrelated – Semantically related	-0.01	0.08	[-0.02, 0.004]	BF ₀₃ = 6.02
Name Agreement	-0.22	0	[-0.32, -0.12]	BF ₀₄ = 0.002
No. fillers	0.01	0.98	[0.00, 0.01]	BF ₀₅ = 5.1

The mean of the posterior distribution of the semantic priming effect $\hat{\beta}_{sem}$ is -0.01 . The credible interval is $[-0.02, 0.004]$. The null hypothesis was that there is no effect of semantic priming. The credible interval is more or less centred around zero, therefore we can conclude that either the null hypothesis that postulates the absence of a semantic priming effect is true or that the effect is very small, i.e. close to zero.

We calculated the Bayes factor to directly compare the plausibility of the null hypothesis (targets which are preceded by semantically related primes are produced faster than targets preceded by an unrelated prime) and the alternative hypothesis (targets which are preceded by semantically related primes are produced faster than targets preceded by an unrelated prime) against each other. Indeed, the BF₀₃ = 6.02 indicates that the data is 6.02 times more likely to be generated under the null hypothesis than under the alternative hypothesis. Thus, the data can be interpreted as evidence in favour of the null hypothesis.

The mean of the posterior distribution of the phonological priming effect $\hat{\beta}_{phon}$ is -0.004 . Zero lies within the credible interval of $[-0.02, 0.01]$. The null hypothesis that there is no priming effect is either true or the effect is very small, i.e. close to zero. Indeed, BF₀₂ = 12.7, meaning that the null hypothesis is 12.7 times more likely than the alternative hypothesis. Thus, the data speak in favour of the null hypothesis, i.e. the absence of phonological priming, being true.

The mean of the posterior distribution of the effect of name agreement on naming latencies $\hat{\beta}_{nameAgr}$ is -0.22 and the probability of the effect of name agreement P ($\hat{\beta}_{nameAgr} > 0$) is 0. By implication, this means that P ($\hat{\beta}_{nameAgr} < 0$) is 1. Thus, it is extremely likely that we observe a negative effect. These results are consistent with the frequentist analysis where we found statistically significant evidence for a negative effect of name agreement.

The mean of the posterior distribution of the effect of the number of fillers $\hat{\beta}_{fillers}$ is 0.01 and the probability of the effect of the number of fillers being larger than zero $P(\hat{\beta}_{fillers} > 0)$ is 0.98.

4.2. Models 2: the effect of relatedness, verb type and their interaction

Model 2 analyses the morphological priming effect by Verb Type. The frequentist analysis led us to reject the null hypothesis for the comparison of morphologically related and unrelated trials (see Table 4), in other words, there was significant evidence for a morphological priming effect. However, none of the other comparisons reached statistical significance.

We ran another Bayesian linear mixed model with the data of Experiment 2 (BLMM 2). The factors and comparisons corresponded to the frequentist linear mixed model we ran on the data of Experiment 2 (for LMM 5 see Section 3.2.2): the main effect of Relatedness (i.e. morphological priming effect), and the comparisons between irregular 2 and irregular 1 conditions, between the regular conditions and the irregular 2 and irregular 1 conditions together, and the interaction of the latter two comparisons with the main effect of Relatedness; see Appendix B Table B2 for an overview of the applied contrast coding. As in the corresponding frequentist analysis, name agreement was included as predictor. In Experiment 2, items were presented twice to increase the number of trials. Thus, “presentation” was included as predictor indicating whether an item was presented first or second. Random intercepts were fitted for participants and items. We used the same weakly informative priors as in the first Bayesian model, that is $N(-1, 1)$ as the prior for the intercept and $N(0, 0.1)$ as prior for each of the experimental comparisons and predictors.

The mean of the posterior distribution, the posterior probability of the effect being larger than 0, the 95% credible interval, and the Bayes factor comparing the likelihood of the null hypothesis to the likelihood of the alternative hypothesis are summarised in Table 13.

In line with the frequentist analysis of Experiment 2, there is very strong evidence for a morphological priming effect in Experiment 2 (BLMM 2). The null hypothesis claims that pictures which have been preceded by a morphological prime are not named faster than those pictures which are preceded by an unrelated prime. As can be seen in Table 13, the probability that the true effect of morphological priming is greater than 0 is larger than 0.99. We can thus safely reject the null hypothesis and conclude that there is a facilitatory morphological priming effect.

The interactions of Relatedness with the two Verb Type comparisons did not reach significance. To evaluate how plausible the null hypothesis is, we calculated the Bayes factor. The Bayes factor tells us to what extent the null hypothesis is to be preferred over the alternative hypothesis. Table 13 also shows that for the interaction between Relatedness and the comparison of regular versus irregular 1 and 2 verbs $BF_{09} = 8.96$, meaning that the null hypothesis is 8.96 times more likely than the alternative. For the interaction between Relatedness and the comparison of irregular 1 versus irregular 2 verbs, $BF_{10} = 12.3$, showing that the null hypothesis, i.e. the absence of an interaction, is more plausible than the alternative. In sum, the frequentist analysis showed no significant results, i.e. it was statistically inconclusive. The Bayesian analysis, by contrast, can be interpreted as strong evidence favouring the null hypothesis, namely the absence of an interaction between Verb Type and morphological priming.

The mean of the posterior distribution of the effect of name agreement was -0.33 . The probability of the effect of name agreement $P(\hat{\beta}_{nameAgr} > 0)$ was 0. Again, this means, that $P(\hat{\beta}_{nameAgr} < 0)$ is 1 and that a negative effect of name agreement was very likely. This finding, too, was in line with the statistically significant effect of name agreement in the frequentist analysis, in which high name agreement resulted in shorter naming latencies.

The frequentist analyses of Experiment 2 included the predictor “presentation”. Hence, this predictor was also included in BLMM 2. The mean of the posterior distribution of the effect of presentation was -0.12 . The

Table 13. Mean of the posterior distribution, posterior probability of the effect being greater than 0, probability of the effect being zero or negligibly small, 95% credible interval and Bayes factor for each of the comparisons in Experiment 2.

Data	Comparison	mean($\hat{\beta}$)	$P(\hat{\beta} > 0)$	95% CrI	BF
Experiment 2	Relatedness	0.05	1	[0.04, 0.07]	$BF_{06} = 0.00$
	Irregular 2 – Irregular 1	-0.01	0.39	[-0.06, 0.05]	$BF_{07} = 3.46$
	Regular – (Irregular 2 & Irregular 1)	0.01	0.69	[-0.03, 0.06]	$BF_{08} = 3.67$
	Relatedness \times (Reg. – (Irregular 2 & Irregular 1))	-0.01	0.12	[-0.02, 0.004]	$BF_{09} = 8.96$
	Relatedness \times (Irregular 2 – Irregular 1)	-0.004	0.3	[-0.02, 0.01]	$BF_{10} = 12.26$
	Name Agreement	-0.33	0	[-0.44, -0.2]	$BF_{11} = 0.00$
	Presentation	-0.12	0	[-0.12, -0.11]	$BF_{12} = 0.00$

probability of the effect of presentation $P(\hat{\beta}_{presentation} > 0)$ was also 0. Thus, a negative effect of presentation (i.e. the more presentations the faster the naming latencies) was very likely and coincides with the statistically significant effect of presentation observed in the frequentist analysis.

To summarise, the Bayesian analysis is largely consistent with the frequentist analysis: it showed clear evidence for a morphological priming effect. In addition, the Bayesian analysis provided strong evidence that the morphological priming effect is not modulated by Verb Type. Moreover, the Bayesian analysis of Experiment 1 revealed that neither semantic nor phonological priming outlasted the long temporal lag between prime and target present in this experimental design.

5. General discussion

This study investigated morphological processing of three different verb types (regular verbs, e.g. *tanzen* – *getanzt*, irregular 1 verbs, e.g. *lesen* – *gelesen* and irregular 2 verbs, e.g. *gießen* – *gegossen*) in overt language production by means of a primed picture naming task. It is one of the few studies to investigate morphological processing in overt language production. Since there were few previous studies, one of our goals was to evaluate long-lag primed picture naming as a suitable method to investigate verb morphology. A second goal was to study how regular verbs and two sub-types of irregular verbs are processed and represented in the cognitive systems of healthy adults.

5.1. Morphological priming in a long-lag primed picture naming

The primed picture naming design of both Experiment 1 and 2 resembled the experimental design by Koester & Schiller (2008; 2011) in which prime words and noun picture targets were separated by five to eight intervening filler items. For morphologically related prime-target pairs, a semantic and phonological relationship is inherent. The lag between prime and target in Koester Schiller's experiment (2008) was designed to eliminate any semantic and phonological priming effects because these effects have been shown not to survive long lags (Zwitsers et al., 2000, see also Koester & Schiller, 2011). Thus, adapting their experimental design to our experiments provided the opportunity to minimise semantic and phonological contributions to morphological priming effects.

However, verbs are not elicited as easily as nouns because it is more difficult to depict actions compared to objects. Therefore, we tested whether this design

resulted in morphological priming and whether semantic and phonological priming effects survived the lag created by items intervening between prime and verb targets.

In both Experiments 1 and 2 we found a morphological priming effect. That is, participants named target pictures significantly faster if they were preceded by a morphologically related prime (e.g. prime: *sie hat getanzt* "she has danced" target: *tanzen* "to dance") relative to target pictures which were preceded by an unrelated prime (e.g. prime: *sie hat gebeichtet* "she has confessed" target: *tanzen* "to dance"). Therefore, supported by the Bayesian analysis, we were able to reject the null hypothesis that the mean naming latencies in unrelated and morphologically related conditions did not differ.

These results replicated the findings by Koester & Schiller (2008) for noun compounds and showed that primed picture naming in a long-lag design can be successfully used to elicit priming effects in verbs. Moreover, naming latencies did not differ for either the semantically or phonologically related condition relative to the unrelated condition and Bayesian linear mixed model analysis provided evidence in support of the null hypothesis of no difference.

Hence, we argue that the morphological priming effect in Experiment 1 and 2 was morphological in nature and independent of semantic and phonological contributions. Therefore, primed picture naming can be used successfully to investigate morphological processes in verb production.

5.2. Morphological priming across verb types in language production

The main goal of this study was the investigation of the production of regular and two sub-types of irregular verbs in order to test predictions from theoretical accounts of verb production. In Experiment 2, the production latencies of target verbs showed the same degree of priming irrespective of their verb type. This outcome was supported by the results of a Bayesian analysis.

Yet, such an outcome is not expected under the Dual Mechanism Model (Pinker & Prince, 1994; Clahsen, 1999) which predicts that irregular 1 and irregular 2 verbs show the same priming patterns and that these dissociate from the pattern shown by regular verbs. Moreover, only regular verbs are predicted to show a priming effect while past participle irregular 1 and 2 primes should not speed up naming the target, or at least, not to the same degree, in contrast to our findings in Experiment 2.

Although we did not find evidence that the three types of verbs are processed differently, we do not suggest that these different types of verbs ought to be lumped together in future experiments: it needs to be demonstrated that these verbs also show no difference in the language production of different populations (e.g. neuropsychological populations) and tasks. Thus, controlling for subtypes of irregular verbs remains necessary to advance the debate on processing and representation of regularity.

Connectionist models (Joanisse & Seidenberg, 1999) argue that differences in processing between regular and irregular past participles relate not to their morphological differences but rather to the phonological and semantic knowledge a speaker has about these verbs. Hence, differences in priming are ascribed to different semantic and phonological priming. For example, Justus et al. (2008) found stronger priming for irregular 2 prime-target pairs than regular prime-target pairs and attributed this to the fact that irregular verbs profit more from semantic priming. However, in the long lag paradigm used here, neither phonological nor semantic priming occurs and hence the connectionist model would predict no difference in processing speed for all three verb types. This was indeed the pattern we found, however, critically, we nevertheless found (equal) morphological priming for all verb types. This cannot be explained by the connectionist account as it stands: as noted in the introduction, the connectionist model has no dedicated morphological processing and hence does not predict any priming in our current priming paradigm.

The priming patterns of our second production experiment was similar to the priming patterns which were found by Smolka et al. (2007) for lexical decision: Just as in their study, we found a morphological priming effect which did not differ by verb type. Accordingly, the results of our experiments are consistent with the predictions made by the model proposed in Smolka et al. (2007). This model maintains that all verb types, regardless of their regularity, are accessed through their stem. Although it is a model for visual word recognition, as suggested in the Introduction, it can be adapted to account for language production as well. That is, when reading the past participle (e.g. *er hat getanzt*/"he has danced" or *er hat gesungen*/"he has sung"), the orthographic input is segmented into its constituent morphemes (e.g. *ge-tanz-t*; *ge-sung-en*). *ge-*, *-t* and *-en* activate the concept *past* while *sung* and *tanz* activate the concept of *sing* and *dance*, respectively. Upon encountering the black-and-white drawing on the screen, the appropriate concept and the corresponding constituent morphemes are selected. Since the constituent morphemes have been

activated already when they were read aloud earlier, the naming latencies for the infinitive target verbs are speeded up in the morphological condition. This is assumed to be true both for regular verbs (prime: *getanzt* – target: *tanzen*) but also for irregular 1 (prime: *gelesen* – target: *lesen*) and irregular 2 verbs (prime: *gesungen* – target: *singen*). Thus, targets were predicted to be able to be primed regardless of their verb type.

Although, Smolka et al. (2007) assume the same processing mechanism for processing different types of verbs, they do recognise differences among those verbs such as stem connectivity (Smolka et al., 2013). Stem connectivity refers to the number of different stems which are present throughout a verbs paradigm and the number of connections each stem has. For example, regular verbs have only one stem throughout the entire paradigm and this one stem has many connections to different bound morphemes. Irregular 1 verbs have 2-3 different stems but each stem enters only few connections to bound morphemes. Irregular 2 verbs have the highest number of different stems in the verbal paradigm but the least number of connections, (e.g. stem *sung* – *gesungen*).

These differences in stem connectivity were thought to explain the graded effects Smolka et al. (2013) saw in ERP patterns to lexical decision but that were not, however, apparent in their RTs. However, if stem connectivity is a decisive factor in the representation and processing of these different types of verbs, effects of stem connectivity should affect RTs as well. That is, a large number of stem-affix combinations of a stem would inhibit selection of the appropriate stem - affix combination and should slow down the production of these forms (i.e. regular verbs). Thus, we expected inhibitory or no priming for regular verbs and facilitatory priming for irregular 1 and 2 verbs. In summary, while our results support the model by Smolka et al. (2007), we did not find any evidence in favour of the influence of stem connectivity.

6. Conclusion

Our study is the first study to investigate the processing of regular and two types of irregular verbs in overt language production using a long-lag priming paradigm (Koester & Schiller, 2008). While we found morphological priming in both experiments, semantic and phonological control conditions in Experiment 1 showed no significant effect. Importantly, morphological priming was the same for the three verb types in Experiment 2.

We substantiated our frequentist analysis of Experiment 1 and 2 with a Bayesian analysis. It allowed us to provide statistical evidence that the null hypotheses

were plausible, i.e. that no phonological or semantic priming effects were present and that morphological priming effects, which were present for all three verb types, did not differ by Verb Type. We conclude that these results can neither be explained by the Dual Mechanism Account (Clahsen, 1999) nor connectionist network (e.g. Joanisse & Seidenberg, 1999). We also did not find any evidence in our data supporting the role of stem connectivity (Smolka et al., 2013). Rather, our findings are most compatible with the model proposed by Smolka et al. (2007) which claims that all verb types are accessed through their stems by a single mechanism.

Note

1. Note that a negative coefficient for covariates such as name agreement also means a faster response. A faster response implies a shorter RT. Thus, a negative coefficient, e.g. in the case of naming agreement, mean that high name agreement leads to shorter RTS, i.e. a faster response.

Acknowledgments

This work was partially funded by the German Science Foundation under grant SFB1287, and by the German Federal Ministry of Research and Education under grant 16DII116-DII. Lyndsey Nickels was funded by an Australian Research Council Future Fellowship (FT120100102).

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Tina Marusch  <http://orcid.org/0000-0001-6191-9449>
 Lena Jäger  <http://orcid.org/0000-0001-9018-9713>
 Frank Burchert  <http://orcid.org/0000-0002-5470-1588>
 Lyndsey Nickels  <http://orcid.org/0000-0002-0311-3524>

References

- Bates, D. M., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4 [Computer software manual]. R package version 1.1-5. Retrieved from <http://CRAN.R-project.org/package=lme4>
- Boersma, P., & Weenink, D. (2009). *Praat: doing phonetics by computer (version 5.1.13)*. <http://www.praat.org>
- Box, G. E. P., & Cox, D. (1964). An analysis of transformations (with discussion). *Journal of the Royal Statistical Society B*, 26, 211–252.
- Budd, M. J., Paulmann, S., Barry, C., & Clahsen, H. (2014). Brain potentials during language production in children and adults: An ERP study of the English past tense. *Brain and Language*, 127, 345–355.
- Cholin, J., Rapp, B., & Miozzo, M. (2010). When do combinatorial mechanisms apply in the production of inflected words? *Cognitive Neuropsychology*, 27(4), 334–359.
- Clahsen, H. (1999). Lexical entries and rules of language: A multidisciplinary study of German inflection. *Behavioral and Brain Sciences*, 22, 991–1060.
- Dohmes, P., Zwitserlood, P., & Bölte, J. (2004). The impact of semantic transparency of morphologically complex words on picture naming. *Brain and Language*, 90, 203–212. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0093934X03004334>
- Elsen, H. (2011). *Grundzüge der Morphologie des Deutschen [Essentials of German morphology]*. Berlin: De Gruyter.
- Ganushchak, L., Christoffels, I., & Schiller, N. O. (2011). The use of electroencephalography (EEG) in language production research: A review. *Frontiers in Psychology*, 2, 208.
- Harley, T. A. (2010). *The psychology of language: From data to theory*. New York: Psychology Press.
- Heister, J., Würzner, K. M., Bubenzer, J., Pohl, E., Hanneforth, T., Geyken, A., & Kliegl, R. (2011). dlexDB - Eine lexikalische Datenbank für die psychologische und linguistische forschung [dlexDB - A lexical database for psychological and linguistic research]. *Psychologische Rundschau*, 62(1), 10–20. Retrieved from <http://psycontent.metapress.com/content/L12056U222554076>
- Joanisse, M. F., & Seidenberg, M. S. (1999). Impairments in verb morphology after brain injury: A connectionist model. *Proceedings of the National Academy of Sciences of the United States of America*, 96, 7592–7597.
- Joanisse, M. F., & Seidenberg, M. S. (2005). Imaging the past: Neural activation in frontal and temporal regions during regular and irregular past-tense processing. *Cognitive, Affective, and Behavioral Neuroscience*, 5(3), 282–296.
- Justus, T., Larsen, J., de Mornay Davies, P., & Swick, D. (2008). Interpreting dissociations between regular and irregular past-tense morphology: Evidence from event-related potentials. *Cognitive, Affective, and Behavioral Neuroscience*, 8(2), 178–194.
- Kliegl, R., Masson, M. E. J., & Richter, E. M. (2010). A linear mixed model analysis of masked repetition priming. *Visual Cognition*, 18(5), 655–681.
- Kliegl, R., Wei, P., Dambacher, M., Yan, M., & Zhou, X. (2011). Experimental effects and individual differences in linear mixed models: Estimating the relation of spatial, object, and attraction effects in visual attention. *Frontiers in Psychology*, 238(1), 1–12. Retrieved from http://www.frontiersin.org/quantitative_psychology_and_measurement/10.3389/fpsyg.2010.00238/abstract
- Koester, D., & Schiller, N. O. (2008). Morphological priming in overt language production: Electrophysiological evidence production. *NeuroImage*, 42, 1622–1630.
- Koester, D., & N. O. Schiller (2011). The functional neuroanatomy of morphology in language production. *NeuroImage*, 55(2), 732–741. Retrieved from <http://www.sciencedirect.com/science/article/pii/S1053811910015223>
- Kutas, M., & Federmeier, K. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, 62, 621–647.
- Mailhammer, R. (2007). Islands of resilience: The history of the German strong verbs from a systemic point of view. *Morphology*, 17(1), 77–108.

- Marcus, G. F., Brinkmann, U., Clahsen, H., Wiese, R., & Pinker, S. (1995). German inflection: The exception that proves the rule. *Cognitive Psychology*, 29, 189–256.
- McClelland, J. L., & Patterson, K. (2002). Rules or connections in past-tense inflections: what does the evidence rule out? *Trends in Cognitive Sciences*, 11(6), 465–472. Retrieved from <http://www.sciencedirect.com/science/article/pii/S1364661302019939>
- McKinnon, R., Allen, M., & Osterhout, L. (2003). Morphological decomposition involving non-productive morphemes: ERP evidence. *Neuroreport*, 14(6), 883–886.
- Morey, R. D., Hoekstra, R., Rouder, J. N., Lee, M. D., & Wagenmakers, E. J. (2015). The fallacy of placing confidence in confidence intervals. *Psychonomic Bulletin & Review*, 23(1), 103–123.
- Nicenboim, B., & Vasishth, S. (2016). Statistical methods for linguistic research: Foundational ideas - part ii. ArXiv e-prints.
- Oldfield, R. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9, 97–113.
- Paul, H. (2007). *Mittelhochdeutsche Grammatik [middle high German grammar]*. Halle: Niemeyer.
- Penke, M., & Westermann, G. (2006). Broca's area and inflectional morphology: Evidence from Broca's aphasia and computer modeling. *Cortex*, 42(4), 563–576.
- Pinker, S., & Prince, A. (1994). Regular and irregular morphology and the psychological status of rules of grammar. In S. D. Lima, R. Corrigan, and G. Iverson (Eds.), *The reality of linguistic rules* (pp. 321–351). Amsterdam: John Benjamins.
- Pinker, S., & Ullman, M. T. (2002). The past and future of the past tense. *Trends in Cognitive Sciences*, 6, 456–463.
- Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1973). *A grammar of contemporary English*. Harlow: Longman Group.
- R Core Team (2013). R: A language and environment for statistical computing [Computer software manual]. Vienna. Retrieved from <http://www.R-project.org/>
- Stan Development Team (2016). *rstanarm: Bayesian applied regression modeling via Stan*. R package version 2.13.1. Retrieved from <http://mc-stan.org/>
- Rumelhart, D. E., & McClelland, J. L. (1986). Parallel distributed processing: explorations in the microstructure of cognition. In (pp. 216–272). Cambridge: MIT Press.
- Smolka, E., Khader, P. H., Wiese, R., Zwitserlood, P., & Rösler, F. (2013). Electrophysiological evidence for the continuous processing of linguistic categories of regular and irregular verb inflection in German. *Journal of Cognitive Neuroscience*, 25(8), 1284–1304. doi:10.1162/jocn_a_00384
- Smolka, E., Zwitserlood, P., & Rösler, F. (2007). Stem access in regular and irregular inflections: Evidence from German principles. *Journal of Memory and Language*, 57, 325–347.
- Sonnenstuhl, I., Eisenbeiss, S., & Clahsen, H. (1999). Morphological priming in the German mental lexicon. *Cognition*, 72(3), 203–236. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0010027799000335>
- Sorensen, T., Hohenstein, S., & Vasishth, S. (2016). Bayesian linear mixed models using stan: A tutorial for psychologists, linguists, and cognitive scientists. *Quantitative Methods for Psychology*, 12(3), 175–200.
- Szekely, A., Devescovi, A., Federmeier, K., Herron, D., Iyer, G., & Bates, E. (2005). Timed action naming and object naming. *Cortex*, 41, 7–25.
- Tabak, W., Schreuder, R., & Baayen, R. H. (2010). Producing inflected verbs: A picture naming study. *The Mental Lexicon*, 5, 22–46. doi:10.1075/ml.5.1.02tab.
- Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S* (4th ed.), New York: Springer. ISBN 0-387-95457-0. Retrieved from <http://www.stats.ox.ac.uk/pub/MASS4>
- Weyerts, H., T. F. Münte, Smid, H. G., & Heinze, H. J. (1996). Mental representations of morphologically complex words: An event-related potential study with adult humans. *Neuroscience Letters*, 206(2–3), 125–128. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0304394096124423>
- Wiese, R. (1996). *The phonology of German*. Oxford: Clarendon Press.
- Zwitserlood, P., Bólte, J., & Dohmes, P. (2000). Morphological effects on speech production: Evidence from picture naming. *Language and Cognitive Processes*, 15(4–5), 563–591. Retrieved from <http://www.tandfonline.com/doi/abs/10.1080/01690960050119706>
- Zwitserlood, P., Bólte, J., & Dohmes, P. (2002). Where and how morphologically complex words interplay with naming pictures. *Brain and Language*, 81(1–3), 358–367.