

# Effects of a Morpheme-Based Spelling Intervention Challenging Previous Results

Viktoria Jöbstl<sup>\*a</sup>, Reinhard Kargl<sup>b</sup>, Anna E. Prattes<sup>c</sup>, Elisabeth Beyersmann<sup>d</sup>, Karin Landerl<sup>e</sup>

Received : 13 February 2021  
Revised : 9 May 2021  
Accepted : 7 June 2021  
DOI : 10.26822/iejee.2021.219

<sup>\*a</sup> **Corresponding Author:** Viktória Jöbstl,  
Institute of Psychology, University of Graz, Austria.  
E-mail: viktoria.joebstl@uni-graz.at  
ORCID: <https://orcid.org/0000-0002-1343-3871>

<sup>b</sup> Reinhard Kargl, Institute for Reading and Spelling, Graz, Austria.  
E-mail: legasthenie@aon.at

<sup>c</sup> Anna E. Prattes, Institute of Psychology, University of Graz, Austria.  
E-mail: anna.prattes@gmail.com

<sup>d</sup> Elisabeth Beyersmann, Department of Cognitive Science and Macquarie University Centre for Reading, Macquarie University, Sydney.  
E-mail: lisi.beyersmann@mq.edu.au  
ORCID: <https://orcid.org/0000-0001-9653-6106>

<sup>e</sup> Karin Landerl, Institute of Psychology, University of Graz, Austria and Department of Cognitive Science and Macquarie University Centre for Reading, Macquarie University, Sydney, Australia.  
E-mail: karin.landerl@uni-graz.at  
ORCID: <https://orcid.org/0000-0003-4074-0233>

## Abstract

Morpheme-based literacy training programs are widely used in German primary schools. This study investigated whether (1) morphological training is effective early in development (Grade 2) and (2) literacy gains can be attributed to advanced morphological processing. Fifty-two German-speaking second-graders participated in an eight-week morpheme-based training program, while an age-matched control group ( $n = 41$ ) attended regular language classes. All children completed spelling, reading and morphological awareness tasks and participated in a masked primed lexical decision experiment, once before training, and then again following training. We observed training effects for spelling and reading morphologically complex words, with bigger increases in the intervention than the control group. In addition, the masked priming results revealed that lexical decision times decreased more strongly in the intervention than the control group, but there was no clear training impact on the pattern of morphological priming. However, the performance on standardized reading tests and a morphological awareness task did not differ across participant groups. Thus, while written language processing improved, it is unclear whether these effects can be attributed to morphological processing or rather general gains in orthographic knowledge.

## Keywords:

Morpheme-Based Spelling Intervention; Spelling; Reading; Visual Word Processing; Morphological Priming

## Introduction

This study aimed to investigate to what extent young German-speaking children can benefit from a morpheme-based spelling intervention. We also wanted to examine whether any gains in spelling (and reading) proficiency could be directly related to improvements in morphological abilities or are simply due to an increase in orthographic knowledge. Morpheme-based instruction and intervention were aimed at teaching children how to identify the semantic building blocks within morphologically complex words (morphemes). For example, the compound "cowboy" consists of two stem morphemes ("cow" and



Copyright ©  
[www.iejee.com](http://www.iejee.com)  
ISSN: 1307-9298

© 2021 Published by KURA Education & Publishing.  
This is an open access article under the CC BY-NC-ND license. (<https://creativecommons.org/licenses/by/4.0/>)

“boy”). The verb “played” consists of the stem morpheme (“play”) and a grammatical morpheme marking past tense (“-ed”). There is increasing interest in morpheme-based instruction and intervention in children’s reading acquisition (e.g., Bowers & Bowers, 2018; Bowers & Kirby, 2010), and a number of arguments in favor of such approaches within the Indo-European language context have been proposed.

### *Transparency of Morphological Information in Print*

While there is large variability among orthographic systems in their transparency and reliability of letter-sound correspondences (De Simone et al., 2021; Schmalz et al., 2016), Indo-European orthographies typically represent morphological structure in a consistent way (e.g., Ulicheva et al., 2018). The orthographic principle of morphological constancy has been mostly discussed for English, which is notoriously inconsistent and irregular at the letter-sound level. Arguably, many inconsistencies at the phonological level are a by-product of the fact that orthography makes linguistically “deeper” morpho-semantic relations transparent, which are not directly evident on the surface level of phonology (Chomsky & Halle, 1968). For example, although the vowel in the two words “heal” and “health” is pronounced differently, their spelling reflects their morpho-semantic relation, and the silent letter “w” in “two” relates it to “twice”, “twin”, and “twenty”. Morpheme-based interventions are assumed to help children deal with the phonological inconsistencies and irregularities of English orthography (Apel et al., 2013; Bowers & Bowers, 2017, 2018).

In many other orthographies, like Spanish, Greek, or German, morphological information is not mandatory in order to decode word pronunciations during reading, as letter-sound relations are mostly consistent and transparent to the reader. Still, these orthographies adhere to the principle of morpheme constancy, meaning that morphemic units are spelt consistently across word forms. In morphologically rich languages this is particularly important as the same stem morpheme can appear in a large number of different word forms (e.g., German “fahr” – Engl.: drive; “fahren”, “fährst”, “fährt”, “gefahren”, “vorfahren”, “Fahrer”, “Gefahr”, “Gefährt”, “Einfahrt”, “Auffahrt”, “Zufahrt” and in a multitude of compounds like “Fahrstuhl”, “Fahrschule”, “Radfahrer”, etc.). Identical spelling of the stem morpheme helps the reader to access word semantics and identical spelling of the numerous inflectional and derivational morphemes helps to quickly identify the grammatical form of the particular word. Thus, recognizing the building blocks of morphologically complex words can support word reading, especially for long words like German “hinterhergegangen” (Engl.: went after) or the compound “Schulsportwettbewerb” (Engl.: school sports competition).

Morphological knowledge is particularly important for spelling (e.g., Ulicheva et al., 2018). Even in “shallow” orthographies with consistent grapheme-phoneme correspondences, phoneme-grapheme correspondences can be highly inconsistent. In German, for instance, the frequent prefix “ver-” is consistently spelt with “v” and not with “f”, although both letters represent the same labiodental fricative and “f” is overall more frequently used for this sound. Orthographic marking of vowel length is particularly inconsistent in German (see Landerl, 2017 for details). For each vowel, there are two or three different spelling alternatives (e.g., /a:/ in “Wal” – Engl.: whale, “Wahl” – Engl.: election, and “Aal” – Engl.: eel) and short vowels are often marked by a subsequent double consonant (“kommen” – Engl.: come) that is retained when followed by another consonant only if this consonant is the onset of a new morpheme (“kommt” – Engl.: he/she/it comes vs. monomorphemic “Hund” – Engl.: dog). Memorizing the spelling of the most frequent word stems and grammatical morphemes can thus help children to correctly spell a large number of words (Bowers & Kirby, 2010; Scheerer-Neumann, 1979).

### *Morphological Skills and Written Language Processing*

Children use morphology productively early in their language development (Mussar et al., 2020). In preschool, they are well able to adapt unfamiliar forms morphologically, as for example demonstrated by Berko’s (1958) classic wug-test (e.g., “This is a wug. Now there is another one. There are two of them. There are two \_\_\_\_.”) and numerous studies since (e.g., Apel et al., 2013; Berninger et al., 2010; Casalis & Louis-Alexandre, 2000; Rispens et al., 2008). The ability to segment and recombine morphemic units is referred to as *morphological awareness* (Apel, 2014; Carlisle, 2000). Variance in morphological awareness explains unique variance in reading (e.g., Deacon & Kirby, 2004; Kirby et al., 2012), and spelling skills (e.g., Kargl & Landerl, 2018; Rispens et al., 2008) above and beyond phonological awareness or vocabulary skills (e.g., Kirby et al., 2012; Levesque et al., 2017).

Despite the evidence for the early development of children’s morphological awareness, the automatization and integration of morphological knowledge into more implicit reading processes does not appear to take place until much later in children’s reading development (e.g., Beyersmann et al., 2012; Dawson et al., 2018; Schiff et al., 2012). Evidence for rapid, automatic morphological processing comes primarily from masked morphological priming, a task that has been extensively used in skilled readers (initially introduced by Longtin et al., 2003 and Rastle et al., 2004) and has by now also been applied in a growing number of studies with developing readers (e.g., Beyersmann et al., 2012; Beyersmann, Grainger, et al., 2015; Beyersmann et al., 2019; Beyersmann et al., 2021; Quémart et al., 2011). In a typical masked

priming paradigm, a mask (e.g., a string of hash keys) is presented for 500 ms, followed by the presentation of the prime for about 50 ms, and the target word, for which participants perform a lexical decision response. The key comparison in this body of research has been between prime-target pairs sharing a semantically transparent morphological relationship (e.g., “farmer – farm”), a semantically opaque morphological relationship (e.g., “corner – corn”) and a purely orthographic relationship (e.g., “cashew – cash”). The widely replicated pattern of results from skilled readers show significant priming in the “farmer – farm” and “corner – corn” conditions, but not in the “cashew – cash” condition, suggesting that adults rapidly decompose morphologically complex words, independently of semantics (for reviews, see Amenta & Crepaldi, 2012; Rastle & Davis, 2008). As opposed to adults however, younger readers appear to rely on more semantically driven morphological processing techniques (“farmer – farm”; Beyersmann et al., 2012; Schiff et al., 2012; but see Quémart et al., 2011).

More recently, there has been a new push in the literature towards the examination of morphologically complex nonwords in developing readers (e.g., Beyersmann, Grainger, et al., 2015; Beyersmann et al., 2020; Beyersmann et al., 2021; Hasenäcker et al., 2016, 2020; Mousikou et al., 2020). The key comparison in this task is between four different prime conditions: a suffixed word condition (e.g., “flexible – flex”), a suffixed nonword condition (e.g., “flexify – flex”), a non-suffixed nonword condition (e.g., “flexint – flex”), and an unrelated control condition (e.g., “faulty – flex”). This paradigm has the advantage that the same targets can be used across all four conditions, cancelling out uncontrolled differences between target words. Moreover, the use of complex nonwords allows to remove any apparent semantic relationship between the prime (e.g., “flexify”) and the target (e.g., “flex”), thus making it possible to investigate the automaticity of morphological processing in a context in which it is impossible to fall back onto processing of the whole letter string.

The overall pattern of results shows that lexical decision times for monomorphemic target words are faster when a preceding prime was a suffixed word, a suffixed nonword, or a non-suffixed nonword, compared to the unrelated control (adults: e.g., Beyersmann, Casalis, et al., 2015; Beyersmann, Cavalli, et al., 2016; Heathcote et al., 2018; Morris et al., 2011; children: Beyersmann, Grainger, et al., 2015; Beyersmann, et al., 2021; Hasenäcker et al., 2016, 2020). This suggests that stems embedded in complex nonwords are rapidly activated independently of whether or not they are accompanied by a suffix or a non-suffix. Additionally, there seems to be an advantage, when the prime and the target share a semantic relationship (suffixed word primes; Beyersmann, Grainger, et al., 2015).

### *Morphological Intervention Programs*

In their meta-analysis based on 30 studies, Goodwin and Ahn (2013) found positive intervention effects on morphological knowledge, phonological awareness, vocabulary, decoding, and spelling, but not reading comprehension or fluency. An ongoing discussion concerns the age at which students can profit from morphological interventions. Interestingly, about half of the studies analysed by Goodwin and Ahn did not start before Grade 4, perhaps based on the assumption that morphological information might be too complex for younger children. However, effect sizes were larger for preschoolers and early elementary school children than for students beyond Grade 2. Similarly, in their systematic review of 22 morphological intervention studies, Bowers et al. (2010) found that training effects for younger children (preschool – 2nd grade) were as large or even larger than those for older children (3rd – 8th grade). This is in line with Bowers and Bowers’ (2018) dedicated call for early morphological instruction (see also Bowers et al., 2010). The authors argued that early morphological instruction can support children’s understanding as to why sometimes the same letter cluster has different pronunciations (e.g., “heal – health”) and at other times different letter clusters have the same pronunciation (e.g., “bald – balled”). Knowledge of frequent stems and affixes may support children in their decoding of unfamiliar words by breaking them down into morphemic sub-units. However, current meta-analyses and reviews are almost exclusively based on studies in English (Bowers et al., 2010; Goodwin & Ahn, 2010, 2013) and therefore do not necessarily generalize to more transparent orthographies like German.

In German-speaking countries, reading and spelling instruction in Grades 1 and 2 is heavily based on phonics (Landerl, 2017). Although phoneme-based sounding out is reliable and successful (though slow and laborious) for reading, it is insufficient for spelling. Simply translating each sound of a spoken word with a phonologically adequate grapheme leads to phonologically plausible but incorrect spellings unless morpheme constancy is regarded. Still, the curriculum for the first two school years contains only very basic morphology instruction. For example, children are encouraged to derive unknown spellings from similar familiar words (e.g., derive plural “Straßen” – Engl.: streets – from singular “Straße”) and spell nouns with an uppercase first letter, according to the German capitalization rule.

One program in particular was designed for poor spellers in late elementary school (Grade 4 and beyond). The MORPHEUS program (Kargl & Purgstaller, 2010) was developed at the same time as the English Structured Word Inquiry (SWI) program (Bowers & Kirby, 2010) and has many conceptual similarities. Both programs introduce children to the concept

of word stems and affixes and encourage them to identify, segment and (re)combine high frequency morphemes. While the objective of the English SWI is to communicate different approaches to investigate written words (combining etymology, morphology and phonology), the German MORPHEUS program focuses on morphology and teaching the most common morphemes to improve children's spelling. In previous intervention studies, MORPHEUS was mostly delivered as a two- to five-week program, in small group settings. Children received teaching lessons complemented by homework consisting of paper-and-pencil as well as computerized tasks. A number of small-scale studies provided evidence that this program can enhance spelling (Kargl et al., 2008; Schneeberger et al., 2011) as well as reading proficiency (Gebauer, Fink, Kargl, et al., 2012; Weiss et al., 2010) in typical as well as poor spellers in Grades 3 to 8 compared to untrained control groups.

### The Present Study

The first aim of the present study was to investigate whether in the transparent German orthography, younger children in Grade 2 can also profit from a morpheme-based intervention. In Grade 2, reading and spelling instruction mostly focus on grapheme-phoneme and phoneme-grapheme translations, supplemented by only minor morpheme-based information for spelling (see above). Children are encouraged to sound out words during reading and phonologically adequate spellings are often accepted, even if they deviate from the orthographically correct spellings. Children at this stage have clear mastery of the alphabetic principle. Here we asked if explicit morphological instruction via the MORPHEUS program leads to improved spelling and reading skills in children as young as Grade 2.

While we expected the intervention to induce larger gains in written language processing skills compared to a control group receiving standard school instruction, it is unclear if any such training effects are directly related to gains in morphological knowledge, or due to an increased focus on spelling practice during the training period. Thus, our second aim was to investigate to what extent any training effects can directly be related to changes in explicit as well as more implicit morphological processing skills. The explicit measure was a standard morphological awareness task requiring children to adapt verbally presented nonwords to a given sentence frame (e.g., "1. Gestern habe ich gewornt. 2. Es wäre gut, wenn du heute \_\_\_ (wornt)." – Engl.: 1. Yesterday I wored. 2. It would be great, if you could \_\_\_ (worn) today.). Given that MORPHEUS trains participants' morphological awareness, gains were to be expected in this task. Indeed, earlier studies found improvements in this type of task after morpheme-based intervention (Apel et al., 2013; Arnbak & Elbro, 2000; Kargl et al., 2008).

Two tasks were used to investigate more implicit mechanisms of morphemic processing, a nonword reading task and a masked primed lexical decision task. In the nonword reading task, children read aloud existing stems that were either combined with an existing prefix or suffix (e.g., "wahrlein" – Engl.: truelet) or with a matched letter cluster that did not constitute an affix (e.g., "wahrnauf" – Engl.: truenauf). We expected faster and more accurate response times to affixed compared to non-affixed nonwords following training.

The masked primed lexical decision task was based on a widely replicated nonword priming paradigm (Beyersmann, Casalis, et al., 2015; Beyersmann, Grainger, et al., 2015; Beyersmann, et al., 2021; Hasenäcker et al., 2016) using four different prime conditions: a suffixed word condition (e.g., "flexible – flex"), a suffixed nonword condition (e.g., "flexify – flex"), a non-suffixed nonword condition (e.g., "flexint – flex"), and an unrelated control condition (e.g., "faulty – flex"). Embedded word priming effects were expected prior to training (Beyersmann, Grainger, et al., 2015; Beyersmann, et al., 2021; Hasenäcker et al., 2020). We further hypothesized that depending on the underlying mechanisms, training may lead to three different priming patterns. 1. If explicit practice in recognizing, storing, and retrieving frequent stems and affixes boosts morphological processing, the training group should show a more marked increase in priming effects for the affixed conditions, compared to the non-affixed conditions. 2. If the morpheme training mostly impacts embedded word activations, priming in all related conditions should increase from pre- to post-test in the training group. 3. If children do not benefit from the training, the priming pattern should not differ at pre- and post-test in both groups.

### Method

#### Participants

Ninety-three German-speaking second-graders participated in this study. The participants were selected from three different schools. The intervention was carried out during language lessons. In two of the schools, children ( $n = 52$ ,  $M_{\text{age}} = 8.29$  years,  $SD_{\text{age}} = 0.50$ ) from four classrooms received the eight-week intervention. In the third school, children functioned as a business-as-usual control group receiving standard language classes (six classrooms;  $n = 41$ ,  $M_{\text{age}} = 8.33$  years,  $SD_{\text{age}} = 0.47$ ). The children from the two schools did not differ in the nonverbal IQ-test given at pre-test (description, see below, intervention: mean z-score = 0.09,  $SD = 0.86$ ; control: mean z-score = -0.10,  $SD = 0.67$ ;  $t(88) = 1.116$ ,  $p = .267$ ).

A total of 122 children completed the classroom administered tasks. However, children who did not speak German at first-language level ( $n_{\text{intervention}} = 15$ ,

$n_{\text{control}} = 5$ ), had identified special needs ( $n_{\text{intervention}} = 2$ ,  $n_{\text{control}} = 3$ ) or whose parents did not agree to contributing their children's data to the research project ( $n_{\text{intervention}} = 2$ ,  $n_{\text{control}} = 2$ ) were excluded.

A power analysis was conducted to ensure that the sample size was large enough to find the expected training effect on spelling and thereby reduce the probability of a type II error. Based on recent studies administering a similar intervention (Gebauer, Fink, Kargl, et al., 2012; Kargl et al., 2008; Schneeberger et al., 2011) we set our expected effect size ( $\eta_p^2$ ) to be at least .11. According to a power analysis calculated in G\*power (Faul et al., 2007) not more than 56 participants are necessary to reduce the type II error to .05. Our sample size clearly exceeded this  $N$ .

All parents provided written consent for assessment and/or training procedures, which were also approved by the local school authorities.

### Procedure

Pre-tests were carried out in January, followed by an eight-week training period (intervention group) or regular German classes (control group), followed by post-tests. A measure of nonverbal IQ was administered at pre-test only, while all other tasks were carried out before and after the intervention period. One of the reading measures (sentence reading), spelling and non-verbal intelligence were administered in the classroom (50-60 minutes total duration), while the remaining reading tasks (word and nonword reading), morphological awareness and a lexical decision task (with morphological priming) were assessed individually (30-45 minutes total duration) in a quiet room in the school. The time between the group and one-on-one sessions varied between one and five days.

### Tasks

#### Non-Verbal Intelligence

The subtests series and *classification* of the German version of the Culture Fair Test (CFT 20-R, Scale 1; Weiß, 2006) were given as classroom tests to control for group differences in children's non-verbal intelligence. The series completion task required children to identify the item that completed a series of shapes (5 minutes). In the classification task, children had to find the item that violated a logical rule (5 minutes). The number of correct items per subtest was z-transformed and averaged to obtain one IQ-score for each child.

#### Spelling

The spelling test of the Lese- und Rechtschreibtest (SLRT-II; Moll & Landerl, 2010) was used to evaluate children's spelling proficiency. To make possible

training effects more observable the longer format with 48 words (standardized for third- and fourth-graders) was used (the standard version for second-graders terminates after 24 words). Children were given a booklet containing 48 sentences with one missing word in each sentence. The experimenter first read out the missing word, then the sentence and then repeated the missing word. Children's task was to fill in the gap. In addition to the number of correctly spelt items, three scores were generated: correctly spelt morphological stems, orthographic stems and affixes. Stems were considered morphological if knowledge of related words could facilitate spelling (e.g., "Bäume <- Baum" – Engl.: trees <- tree). The SLRT-II provides two parallel versions (A & B), which were counterbalanced during pre-test. At post-test the order of versions was reversed to minimize repetition effects.

### Reading – Fluency

Reading fluency was assessed using the Salzburger Lese-Screening (SLS 1-4; Mayringer & Wimmer, 2003) and the Lese- und Rechtschreibtest (SLRT-II; Moll & Landerl, 2010). The SLS 1-4 was administered in class. Participants were given three minutes to silently read as many sentences as possible. After every sentence, a check mark or a cross had to be circled depending on whether or not the sentence was correct (e.g., "Strawberries are blue."). The dependent measure was the number of sentences marked correctly. The SLRT-II was carried out one-on-one. Children were asked to read aloud a list of words and a list of nonwords as fast as possible for one minute each. The dependent measure was the number of items read correctly within the time limit. The two parallel forms were counterbalanced and all children received different forms at pre- and post-test.

#### Reading of Morphologically Complex Stimuli

To investigate training effects on reading fluency, four reading lists with morphologically complex items were created. Each list contained 24 nonwords consisting of existing stems combined with real or pseudo-affixes. The stem was either paired with (1) a prefix (e.g., "umfinden"), (2) a pseudo-prefix (e.g., "arfinden"), (3) a suffix (e.g., "gesetzisch") or (4) a pseudo-suffix (e.g., "gesetzucht"). For the prefixed verb stimuli, the infinitive morpheme "-en" was added to the stem to create possible verb forms (in German, verb forms without a suffix do not exist in the present tense). A parallel form was generated for each list by reversing item order. The parallel forms were counterbalanced during pre-test and children were always given the other parallel form during post-test. They were instructed to read out loud all items as fast as possible, without making mistakes. The score was the number of correctly read nonwords per minute. Cronbach's alpha across the four item lists was .98 at pre- as well as post-test.



### Morphological Awareness

This task was based on a German test of morphological awareness for children aged 10 or older (Test zur Erfassung der morphematischen Bewusstheit, TMB; Kargl et al., 2006). Although the original test is administered in print, we made the test more accessible for our younger participants by presenting the stimuli verbally. Sentences containing a nonword were presented. Children then had to manipulate the nonword in such a way that it would complement the next sentence (e. g.: "1. Gestern habe ich gewornt. 2. Es wäre gut, wenn du heute \_\_\_ (wornt)." – Engl.: Yesterday I worked. 2. It would be great, if you could \_\_\_ (work) today.). All items required the morphologically adequate manipulation of a pre- and/or suffix. Correctly derived nonwords mirrored either pseudo-nouns (derivation, inflection), -adjectives (derivation, building the comparative and/or superlative) or -verbs (past participle, inflection). Mistakes in pronunciation of the stem were not rated. The score was the number of correctly manipulated words (max. 24). Cronbach's alpha was .83 at pre-test and .77 at post-test.

### Masked Morphological Priming

A lexical decision paradigm with masked morphological priming was used to assess the relevance of morphological segments in written word processing. Materials and procedure were adapted from a study by Beyersmann et al. (2021). Two sets of target words each containing 48 items were selected from the childLex corpus (Schroeder et al., 2015). Target words in the prefix condition were the infinitive forms of verbs and target words in the suffix condition were nouns or adjectives. Each target was preceded by an affixed word prime (e.g., prefix: "mitdenken – DENKEN" – Engl.: think along – THINK; suffix: "steinchen – STEIN" – Engl.: little stone – STONE), an affixed nonword prime (e.g., prefix: "hindenken – DENKEN" – Engl.: think towards – THINK; suffix: "steinkeit – STEIN" – Engl.: stoniness – STONE), a non-affixed nonword prime (e.g., prefix: "kardenken – DENKEN" – Engl.: karthink – THINK; suffix: "steinucht – STEIN" – Engl.: stonel – STONE), and an unrelated prime (e.g., prefix: "karhasten – DENKEN" – Engl.: carpaint – THINK; suffix: "piratucht – STONE" – Engl.: pirate – STONE). The prefixed words included the prefixes "auf-", "mit-", "ab-", and "an-", the suffixes "-chen", "-haft", "-heit", and "-lich", which were each repeated 12 times.

Nonword primes were created using the target word (e.g., "denken" – Engl.: think, "stein" – Engl.: stone), and combining it with a prefix (e.g., "hin-" – Engl.: towards) or suffix (e.g., "-keit" – Engl.: -ity), such that the whole letter string was not a real word (e.g., prefix: "hindenken" – Engl.: think towards; suffix: "steinkeit"

– Engl.: stoniness). Non-affixed nonword primes were created by combining the same word (e.g., "denken" – Engl.: think; "stein" – Engl.: stone) with a common, non-morphemic letter-sequence (e.g., "kar-"; "-ucht"), such that the whole letter string was not a real word ("kardenken" – Engl.: karthink; "steinucht" – Engl.: stonel). Unrelated primes were non-affixed nonwords and orthographically unrelated to the target. All nonwords were orthographically legal and pronounceable. The four prime conditions were matched on length.

For the purpose of the lexical decision task, 98 nonword targets (48 for the prefixed materials and 48 for the suffixed materials) were created from words by replacing one or two letters (e.g., "laufen" – Engl.: to run -> "laupen"; "ganz" – Engl.: whole -> "galz"). In both the prefixed and suffixed materials, nonword targets were matched to real word targets on length. To mimic the structure of the primes preceding real word targets, primes preceding nonword targets were selected in a similar fashion. Each nonword target was preceded by four different types of primes, by combining the nonword targets with an affix (e.g., "anlaufen – LAUPEN"; "galzhaft – GALZ"), with a different affix (e.g., "umlaufen – LAUPEN"; "galzisch – GALZ"), with a non-morphemic letter sequence (e.g., "emlaufen – LAUPEN"; "galztern – GALZ"), and by combining an unrelated nonword with a non-morphemic letter sequence (e.g., "emliezen – LAUPEN"; "zelptern – GALZ").

At pre- and post-test each participant completed the prefixed and suffixed masked priming experiments. Within each set of materials, target order was randomized. To ensure that every child encountered each target and prime only once, four lists per condition were generated and counterbalanced across participants. The presented lists at pre-test were always different from the lists a child received at post-test.

Targets were presented at the center of a laptop screen. Each trial started with a 500 ms forward mask of hash keys (#####) followed by the prime (50 ms) in lowercase and then the target in uppercase. The target remained present until a response had been made. The children were instructed to indicate as quickly and accurately as possible whether the presented target was a word or not by pressing either the "K" (yes) or "D" (no) button.

### Morpheme-Based Training

The administered morpheme-based spelling training was an adaptation of the computer-aided training program MORPHEUS (Kargl & Purgstaller, 2010). The training was provided in the classroom by two graduate students who were trained and supervised

by the developers of the program. At the start of the intervention, all participants received a folder in which they collected their weekly worksheets. Additionally, a quarter to a third of the time they worked on tablets with an especially designed app. The central aim of the program was to increase children's awareness of the morphological structure of words (prefix, stem, suffix) and thereby learn that words are often composed of familiar elements. The basic idea was to familiarize children with the correct spellings of a limited number of high-frequency morphemes, in order to enable them to correctly spell a multitude of morphologically complex words containing those morphemes (Bowers & Kirby, 2010; Scheerer-Neumann, 1979).

The training phase lasted eight weeks. On two days each week altogether three training lessons (one double and one single lesson) were held during language lessons in school and children received materials for about one hour of homework per week. The progress was checked regularly by reviewing worksheets and controlling whether all previous chapters on the tablets were completed. Chapters with an accuracy rate of 75 % or lower were repeated. After the first half of the training there was a one-week break from school for all participating children.

The training consisted of three levels, which were composed of overall 14 chapters (see online Appendix for exemplary tasks). Level one (chapters 1 and 2) introduced the concepts of word families and word classes (nouns, verbs, adjectives), including the spelling rule of capitalization for nouns. Children learned to identify words from the same word family (i.e., including the same stem) in a set of presented words and to build derived words by combining given prefixes, stems, and suffixes. First, they learned that nouns (referred to as "name words" in German – meaning that nouns often notify the name of objects) always start with an uppercase letter (e.g., "Tür", "Freund", "Haus" – Engl.: door, friend, house) and are often accompanied by an article ("die", "der", "das" – Engl.: the). Verbs describe what someone does and adjectives how someone/something is. Verbs and adjectives are not capitalized. Subsequently, the newly acquired knowledge was practiced by assigning different words to the corresponding word class.

Level two (chapters 3 to 6) introduced word families with stems that involve variation of vowel pronunciation and spelling across word forms (e.g., "ich sprech-e – du sprich-st"; "Baum – Bäum-e" – Engl.: I speak – you speak; tree – trees). Children also learned that certain suffixes ("-keit", "-chen", "-ung", "-er", "-erin") always notify nouns. For example, when the adjective stem "dunkel" (Engl.: dark) is combined with the suffix "-heit" (Engl.: -ness) it becomes the

noun "Dunkelheit" (Engl.: darkness) and must be capitalized. Tasks in these chapters were combining prefixes, stems, and/or suffixes into words, segmenting words into their constituent morphemes, identifying the stems in word forms with vowel change ("Bäume → Baum" – Engl.: trees → tree) and memorizing high-frequency morphemes by repeatedly reading given prefixes and stems.

Level three was by far the largest, consisting of chapters 7 to 13. Each of those chapters focused on word stems containing specific letter combinations that are frequently misspelt. These letter clusters typically concerned orthographic marking of vowel length (i.e., double consonants, "tz", or "ck" following short vowels; "ß", "ie", silent "h", or double vowels marking long vowels). Children practiced to identify vowels in word stems (but not those present in pre- or suffixes, e.g., "be-komm-en"; Engl.: get). The short vowel in this stem is marked by a subsequent double consonant and this orthographic pattern is consistent across all word forms including this stem, even when the next morpheme also starts with a consonant (e.g., "er kommt" – Engl.: he comes). Children also learned that certain suffixes (e.g., "-lich") can be used as indicators for adjectives, which are not capitalized. Across chapters, children were familiarized with high frequency stems (including specific orthographic patterns) and used them to build and segment morphologically complex words, including inflectional and derivational processes involving pre- as well as suffixing of nouns, verbs and adjectives, and also compounding (e.g., "Spiel-platz" – Engl.: play-ground), which is highly productive in German. The overarching concept was to practice the principle of morpheme constancy exemplified by stems, pre- and suffixes with high frequency of occurrence. In Chapter 14 previously learned contents were recapped and revised.

#### **Business-as-usual control group**

During the training period, the control group received standard German language lessons based on the workbook *Karibu 2* (Eichmeyer et al., 2013). The dominant structural unit in this program is the syllable: Children are encouraged to segment polysyllabic words in German texts into syllabic units during reading and spelling. To practice spelling monosyllabic words ending in a stop consonant, which is always devoiced in German pronunciation, children were asked to extend words to make voiced consonants perceptible in intervocalic position (e.g., "Dieb – Diebe" – Engl.: thief – thieves). This strategy helps to identify if a syllable final consonant needs to be doubled, as double consonants are ambisyllabic between two vocalic syllable nuclei (e.g., "kommt – kommen" – Engl.: comes – come). This word extension strategy usually involves morphological processes (e.g., noun pluralization or

verb inflection), but is not explicitly alluded to children. The program contains some morphological aspects as required by the national curriculum: Children are taught the concept of “name words” (nouns) and that they are consistently spelt with a capital first letter. In the last third of the one-year program, children typically start building words in the context of verb inflection (e.g., “ich male, du malst, er malt” – Engl.: I paint, you paint, he paints). Even later in the program, the terminology “word family” is introduced for words including the same stem (for more details regarding the administration of Karibu 2 over the course of the school year see Eichmeyer & Zoltan, 2013). The Karibu 2 program is supplemented by unspecific semantically structured language exercises (e.g., seasons or months of the year, the weather, emotions) that involve reading and writing short texts.

## Results

### *Standardized Tests of Spelling and Reading.*

The main goal of our training was to improve children’s spelling skills. Earlier studies (Gebauer, Fink, Kargl, et al., 2012; Weiss et al., 2010) had also demonstrated gains in reading. Thus, we first investigated whether the training was indeed able to induce specific gains in the standardized measures of spelling and reading. Pre- and post-test scores are displayed in Table 1. ANOVAs with time (pre-/post-test) as within-subjects factor and group (intervention/control) as between-subjects factor were run in SPSS Statistics 26.0.

Children in both groups showed improvements in their overall spelling skills (number of words spelt correctly) from pre- to post-test,  $F(1,85) = 47.10, p < .001, \eta_p^2 = .357$ . Importantly, this improvement was larger in the training than in the control group,  $F(1,85) = 5.58, p = .020, \eta_p^2 = .062$ . There was no main effect of group,  $F(1,85) = 1.39, p = .240, \eta_p^2 = .016$ . Planned comparisons did not reveal group differences at pre-test,  $t(85) = 0.51, p = .611; d = 0.11$ , or post-test,  $t(85) = 1.85, p = .067; d = 0.39$ .

Depending on the administered parallel form (A or B; SLRT-II) 54 % or 58 % of the words in the spelling test were morphologically complex in that they consisted of two or more morphemes. In order to get a more specific impression whether the improvements in spelling were directly related to morphological skills, we analysed the number of correctly spelt affixes and stems separately (see Table 1). For stems, we differentiated between morphological stems the spelling of which could be derived based on morphological information (e.g., “Bäume <- Baum” – Engl.: trees <- tree; A: 14 of 48 stems; B: 16 out of 49 stems) and orthographic stems, which had to be accessed from orthographic memory (e.g., “wohnen” – Engl.: to reside/to live; A: 34 out of 48 stems; B: 33 out of 49 stems).

There was an interaction between type of stem, time and group,  $F(1,84) = 9.17, p = .003, \eta_p^2 = .098$ . Consequently, time and group effects regarding spelling of morphological and orthographic stems were analysed separately. Spelling of both, morphological and orthographic stems improved over time, morphological:  $F(1,84) = 41.98, p < .001, \eta_p^2 = .333$ ; orthographic:  $F(1,84) = 37.85, p < .001, \eta_p^2 = .311$ , but a group x time interaction was only observed for morphological stems,  $F(1,84) = 8.97, p = .004, \eta_p^2 = .096$ , and not orthographic stems,  $F(1,84) = 0.41, p = .526, \eta_p^2 = .005$ . There also was a main effect of group for morphological,  $F(1,84) = 5.81, p = .018, \eta_p^2 = .065$ , but not orthographic stems,  $F(1,84) = 0.46, p = .501, \eta_p^2 = .005$ . Planned comparisons on morphological stems confirmed that there were no group differences at pre-test,  $t(84) = 1.18, p = .242; d = 0.26$ , but at post-test the intervention group spelt more morphological stems correctly than the control group,  $t(84) = 3.15, p = .002; d = 0.69$ . No significant effects were observed for affixes (all  $p$ s > .05).

All three standardized reading measures indicated overall improvements from pre- to post-test (sentence reading:  $F(1,83) = 74.08, p < .001, \eta_p^2 = .472$ ; word reading:  $F(1,88) = 112.28, p < .001, \eta_p^2 = .561$ ; nonword reading:  $F(1,88) = 66.29, p < .001, \eta_p^2 = .430$ ). However, no time x group interaction was observed for sentence reading,  $F(1,83) = 1.02, p = .316, \eta_p^2 = .012$ , and word reading,  $F(1,88) = .27, p = .603, \eta_p^2 = .003$ , and the time x group interaction for nonword reading,  $F(1,88) = 4.61, p = .034, \eta_p^2 = .050$ , was due to a larger improvement in the control than in the experimental group. The group effect was also not significant for any of the reading measures (sentence reading:  $F(1,83) = 2.13, p = .148, \eta_p^2 = .025$ ; word reading:  $F(1,88) = 0.12, p = .729, \eta_p^2 = .001$ ; nonword reading:  $F(1,88) = 0.09, p = .764, \eta_p^2 = .001$ ). Thus, there was no evidence for training-related improvements in the standardized reading measures.

### *Reading of Morphologically Complex Stimuli*

Separate ANOVAs were run for the prefix- and the suffix conditions. Affix type (pseudo-/real affix) and time (pre-/post-test) were within-subjects factors and group (intervention/control) the between-subjects factor. Distributions were normalized by winsorizing scores more than 3 SDs above the mean to the reading time corresponding to 3 SDs above the mean (this affected only 10 scores altogether). Number of nonwords read correctly per minute and condition are displayed in Table 2.

For the (pseudo-) prefixed lists there was a main effect of affix type,  $F(1,88) = 167.87, p < .001, \eta_p^2 = .656$ , as well as time,  $F(1,88) = 138.31, p < .001, \eta_p^2 = .611$ , but not group,  $F(1,88) = 0.04, p = .841, \eta_p^2 = .000$ . Stems with existing prefixes were read more efficiently than stems with



**Table 1***Standardized Spelling and Reading Tests: Means and Standard Deviations for Pre- and Post-Test*

Task	Intervention		Control		Pre-test	Post-test
	Pre-test	Post-test	Pre-test	Post-test		
	M (SD)	M (SD)	M (SD)	M (SD)		
Spelling	17.85 (10.51)	23.14 (10.98)	16.74 (9.70)	19.32 (8.30)	1 42	3 44
Morphological stems	37.33 (21.26)	53.24 (24.17)	32.50 (15.40)	38.35 (18.23)	0.00 92.86	0.00 93.75
Orthographic stems	55.84 (22.75)	63.18 (21.42)	52.07 (19.67)	61.11 (18.32)	6.06 94.12	14.71 97.06
Affixes	79.60 (10.56)	79.23 (12.73)	76.02 (10.03)	74.46 (9.41)	36.67 100	51.72 100
Sentence reading	22.38 (7.09)	26.02 (8.76)	19.53 (7.46)	24.13 (7.58)	7 37	10 46
Word reading	42.12 (16.53)	49.86 (18.71)	40.46 (14.27)	49.00 (19.92)	13 80	11 98
Nonword reading	30.67 (9.36)	32.92 (10.08)	29.24 (9.53)	33.10 (10.92)	10 55	12 64

Note: Spelling: Number of correctly spelt words. Morphological stems, orthographic stems, affixes: % correctly spelt morphemes. Sentence reading: Read sentences in 3 minutes. Word and nonword reading: Number of read items in 1 minute.

pseudo-prefixes (prefix:  $M = 33.58$ ,  $SD = 14.39$ ; pseudo-prefix:  $M = 24.35$ ,  $SD = 10.17$ ) and both groups read more items per minute correctly after than before the training period (pre-test:  $M = 25.61$ ,  $SD = 11.25$ ; post-test:  $M = 32.32$ ,  $SD = 13.25$ ). Importantly, there was an interaction affix type  $\times$  time,  $F(1,88) = 17.49$ ,  $p < .001$ ,  $\eta_p^2 = .166$ , and affix type  $\times$  time  $\times$  group,  $F(1,88) = 6.06$ ,  $p = .016$ ,  $\eta_p^2 = .064$ . Follow-up ANOVAs separately for the two groups revealed a significant affix type  $\times$  time interaction for the intervention,  $F(1,48) = 25.88$ ,  $p < .001$ ,  $\eta_p^2 = .350$ , but not for the control group,  $F(1,40) = 1.26$ ,  $p = .268$ ,  $\eta_p^2 = .031$ , showing that from pre- to post-test, only the intervention group showed larger improvements for prefixed than pseudo-prefixed nonwords.

For the (pseudo-) suffixed lists we found main effects of affix type,  $F(1,88) = 76.79$ ,  $p < .001$ ,  $\eta_p^2 = .466$ , and time,  $F(1,88) = 86.16$ ,  $p < .001$ ,  $\eta_p^2 = .495$ , but not group,  $F(1,88) = 0.27$ ,  $p = .605$ ,  $\eta_p^2 = .003$ . There was a significant interaction affix type  $\times$  time,  $F(1,88) = 5.10$ ,  $p = .026$ ,  $\eta_p^2 = .055$ , but no other interactions were significant ( $F_s(1,88)$  between 0.20 and 0.54, all  $p_s > .1$ ). Again, stems with existing suffixes were read more efficiently than stems with pseudo-suffixes (suffix:  $M = 22.84$ ,  $SD = 9.78$ ; pseudo-suffix:  $M = 19.45$ ,  $SD = 8.03$ ), performance was better after than before the training (pre-test:  $M = 19.14$ ,  $SD = 8.21$ ; post-test:  $M = 23.15$ ,  $SD = 9.72$ ), and improvements were larger for suffixed compared to pseudo-suffixed items. However, in the suffixed condition we did not see any evidence for effects directly related to the training.

### Morphological Awareness

Children showed better performance in the morphological awareness task during post- than pre-test,  $F(1,88) = 133.85$ ,  $p < .001$ ,  $\eta_p^2 = .603$  (see Table 2). However, there was no time  $\times$  group interaction,  $F(1,88) = 0.00$ ,  $p = .965$ ,  $\eta_p^2 = .000$ , and no main effect group,  $F(1,88) = 0.01$ ,  $p = .917$ ,  $\eta_p^2 = .000$ . Thus, we have no evidence that the improvement in task performance was related to the morphological training.

### Masked Morphological Priming

Because this task turned out to be quite difficult for our young participants, we decided to exclude four children (one from the training and three from the control group) who were obviously still struggling with basic reading (percentile 10 or below on word reading during post-test). Children who performed at or below guessing rate (50 % accurate) in the lexical decision paradigm were also excluded (pre-test: one child each from the prefix and suffix condition; post-test: seven children from the prefix and six from the suffix condition). Three words from the suffix condition and two words from the prefix condition were excluded from analysis because of generally low accuracy rates. The analysis of response times (RTs) was based on correct responses on word targets. RTs below 300 ms and above 7000 ms were excluded as we assumed that they resulted from task unrelated factors (2.34 % of correct responses). To meet the normality assumptions

**Table 2**

*Reading of Morphological Complex Stimuli and Morphological Awareness: Means and Standard Deviations for Pre- and Post-Test*

Task	Intervention		Control			
	Pre-test	Post-test	Pre-test	Post-test	Pre-test	Post-test
	M (SD)	M (SD)	M (SD)	M (SD)	Min Max	Min Max
Prefix	28.71 (14.41)	39.11 (15.98)	29.64 (12.68)	36.39 (15.82)	6.58 72.00	9.62 84.71
Pseudo-prefix	21.68 (9.17)	26.61 (10.50)	21.72 (10.08)	27.08 (12.19)	4.64 51.43	4.62 62.61
Suffix	20.55 (9.34)	25.67 (11.24)	20.13 (8.35)	24.77 (11.08)	5.22 46.45	4.17 55.00
Pseudo-suffix	17.99 (8.15)	21.89 (8.18)	17.50 (7.90)	20.26 (9.13)	5.75 40.00	3.33 47.50
Morphological awareness	12.37 (5.39)	17.06 (4.51)	12.29 (4.87)	16.95 (3.26)	2 23	5 24

Note: Reading lists: Numbers of correctly read nonword per minute (after winsorizing). Morphological awareness: Means and standard deviations of correct responses.

data-points with standardized residuals bigger than 2.5 in absolute values were excluded (see Baayen, 2008) and reactions times were transformed (-1/RT; see Kliegl et al., 2010).

Training effects in accuracy and RTs of the lexical decision task were analysed with (generalized) linear mixed-effects models using the statistic software R (R Core Team, 2019) with the packages lme4 (Bates et al., 2015), car (Fox & Weisberg, 2019) and lmerTest (Kuznetsova et al., 2017). Suffix and prefix conditions were analysed separately. Fixed factors were prime type (affixed word, affixed nonword, non-affixed nonword, unrelated), time (pre-/post-test), and group (intervention/control), subject and item were random factors. In order to find the optimal number of factors fixed factors and interactions were only included if they improved the model's fit in a backward stepwise model selection procedure.

Mean response accuracy was 89.79 % (SD = 30.27 %) in the prefix and 89.64 % (SD = 30.47 %) in the suffix condition. Neither in the prefix, nor in the suffix condition main effects of time or group, or the interactions time x group or prime type x time x group were significant.

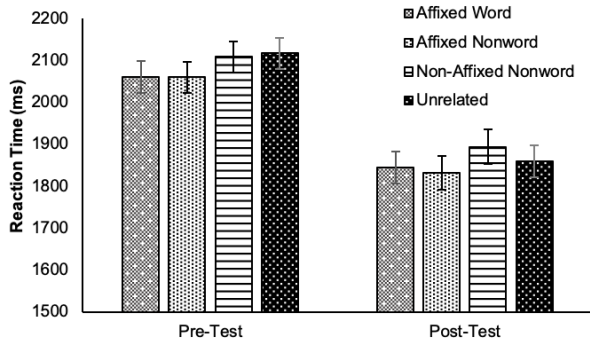
Mean RTs averaged across groups at pre- and post-test are presented separately for prefix- and suffix conditions in Figures 1 and 2. Mean RTs separately for the intervention and control groups are presented in Figure 3 and 4. In the prefix condition there was a main effect time,  $\chi^2(1) = 259.12, p < .001$ . Children were faster at post-test than pre-test. Prime type was marginally significant,  $\chi^2(3) = 7.25, p = .064$ . Post-hoc contrasts

showed that the affixed word,  $z = 2.02, p = .044$ , and affixed nonword condition,  $z = 2.42, p = .016$ , but not the non-affixed nonword condition,  $z = 1.23, p = .219$ , differed from the unrelated condition. There was no difference between the affixed word and the affixed nonword condition,  $z = -0.41, p = .680$ , the affixed word and non-affixed nonword condition,  $z = 0.78, p = .434$ , and between the affixed nonword and non-affixed nonword condition,  $z = 1.19, p = .234$ . There was no main effect group but there was an interaction time x group,  $\chi^2(1) = 17.76, p < .001$ . The groups did not differ at pre-test,  $z = -0.06, p = .530$ , or post-test,  $z = 0.43, p = .665$ , but the decrease in reaction times was larger in the intervention than in the control group, i.:  $z = -16.09, p < .001$  (265 ms) vs. c.:  $z = -8.28, p < .001$  (180 ms). No other interactions were significant.

In the suffix condition we observed main effects of time,  $\chi^2(1) = 478.08, p < .001$ , and prime type,  $\chi^2(3) = 9.69, p = .021$ , with lower reaction times at post- than pre-test. Post-hoc contrasts showed that the affixed word,  $z = -15.33, p < .001$ , but not the affixed nonword,  $z = 1.56, p = .119$ , and the non-affixed nonword condition,  $z = 1.42, p = .155$ , differed from the unrelated condition. Again, there were no differences between the affixed word and affixed-nonword condition,  $z = 1.44, p = .150$ , the affixed word and the non-affixed nonword,  $z = 1.59, p = .113$ , and the affixed nonword and the non-affixed nonword condition,  $z = 1.42, p = .887$ . The group effect was not significant but there was an interaction time x group,  $\chi^2(1) = 30.84, p < .001$ . Like in the prefix condition there was a stronger decrease in overall RTs in the intervention than in the control group (300 vs. 229 ms). There were no group differences at pre-test,  $z = -0.02, p = .985$ , and post-test,  $z = 1.32, p = .186$ , and

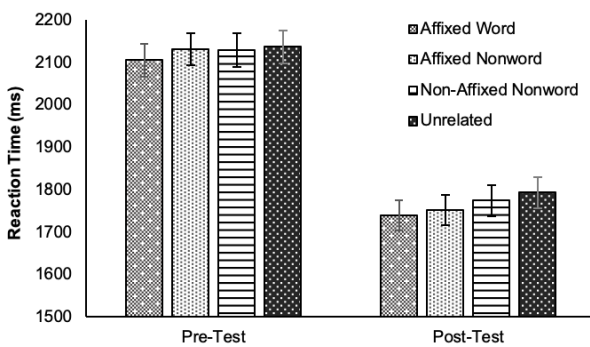
reaction times decreased in both groups, *i.*:  $z = -21.88, p < .001$ ; *c.*:  $z = -12.04, p < .001$ . No other interactions were significant.

**Figure 1**  
*Reaction Times—Prefix Condition*



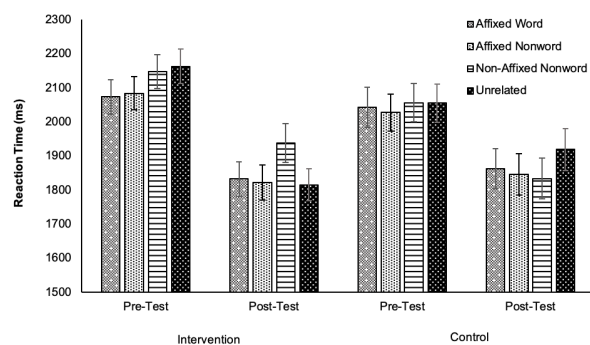
Note: Mean RTs (lexical decision task) averaged across groups at pre- and post-test. Each of the four bars represents the mean reaction time after “seeing” one out of the four possible masked primes (affixed word, affixed nonword, non-affixed nonword, unrelated). Error bars show standard errors.

**Figure 2**  
*Reaction Times—Suffix Condition*



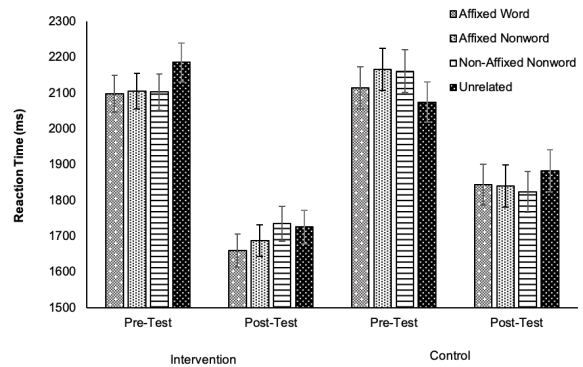
Note: Mean RTs (lexical decision task) averaged across groups at pre- and post-test. Each of the four bars represents the mean reaction time after “seeing” one out of the four possible masked primes (affixed word, affixed nonword, non-affixed nonword, unrelated). Error bars show standard errors.

**Figure 3**  
*Reaction Times—Prefix Condition*



Note: Reaction times (lexical decision task) after a morpheme-based training or alternatively before and after eight weeks of regular German classes. Each of the four bars represents the averaged reaction time after “seeing” one out of the four possible masked primes (affixed word, affixed nonword, non-affixed nonword, unrelated). Error bars show standard errors.

**Figure 4**  
*Reaction Times—Suffix Condition*



Note: Reaction times (lexical decision task) after a morpheme-based training or alternatively before and after eight weeks of regular German classes. Each of the four bars represents the averaged reaction time after “seeing” one out of the four possible masked primes (affixed word, affixed nonword, non-affixed nonword, unrelated). Error bars show standard errors.

**Discussion**

Morpheme-based instruction programs have recently attracted an increasing amount of attention among researchers and practitioners (Allen & Lembke, 2020; Apel et al., 2013; Bowers & Bowers, 2017, 2018; Bowers et al., 2010; Gebauer, Fink, Filippini, et al., 2012; Goodwin & Ahn, 2013; Zhang & Zou, 2020). Explicit phonics-based instruction helps children to understand and apply the alphabetic principle of Indo-European orthographies (Ball & Blachman, 1988, 1991; Bus & van Ijzendoorn, 1999; Ehri et al., 2001; Fischer & Pfost, 2015). More recently, it has been proposed that explicit training of morphological knowledge may support children’s reading development (Arnbak & Elbro, 2000; Bowers & Bowers, 2017; Bowers & Kirby, 2010). Crucially, in many orthographies the morphological principle tends to overrule the alphabetic principle.

The current study investigated the efficiency of a morpheme-based instruction in German-speaking second-graders and expanded the research literature in several ways. First, the vast majority of training studies published in international scientific journals were carried out in the phonologically opaque English orthography and only very few studies focused on more transparent orthographies like German (Kargl et al., 2008; Schneeberger et al., 2011; Walter et al., 2007; Weiss et al., 2010). German orthography is more consistent than English, and is also characterized by a richer morphology and strong reliance on morphological knowledge within the spelling system. Second, while several studies have demonstrated the efficiency of morpheme-based interventions for older German children (Gebauer, Fink, Filippini, et al., 2012; Gebauer, Fink, Kargl, et al., 2012; Kargl et al., 2008; Schneeberger et al., 2011; Walter et al., 2007; Weiss et al., 2010), we investigated if children in Grade 2, who

typically receive a largely phonics-based teaching approach, would also benefit from explicit morphemic instruction.

An important further aim of our study was to investigate if the intervention program would more generally boost children's spelling or reading skills. As morphological interventions in school-aged children typically involve the reading and spelling of morphemes and words, it is possible that training effects are caused by memory effects. That is, children may memorise the trained words and word parts and use this knowledge during written language processing. Training effects would then be evident due to the fact that the word material used during training overlaps with that occurring in pre- and post-test.

### *Training Effects on Spelling and Reading*

The first important finding was that second-graders who received eight weeks of morpheme-based instruction showed significantly stronger improvements in spelling than children who received the business-as-usual teaching. Note that the morpheme-based training was provided as part of children's language classes, so the amount of teaching was comparable between the two groups. In contrast, the standard Grade 2 curriculum in Austria is more strongly focused on phoneme segmentation and blending. This finding thus extends earlier research on morphological intervention in older German students (Schneeberger et al., 2011; Walter et al., 2007). It further suggests that children as young as Grade 2 can profit from morphologically structured spelling instruction in a phonologically transparent orthography (for converging evidence from English, see Bowers et al., 2010; Goodwin & Ahn, 2013).

A more fine-grained analysis of the spelling results showed that the larger improvement in the training group was mostly related to more accurate spellings of word stems for which morphological information was helpful. Our data suggest that the intervention helped children to understand the principle of morpheme constancy. Orthographic stems, the spelling of which must be fully memorized (e.g. that the word form "wohnt" is spelt with a so-called "silent h" because all words containing the stem "wohn-" – Engl.: to reside, are spelt that way), did not show a training-specific improvement. Moreover, affix spellings did not improve from pre- to post-test, probably because the test contained mostly phonologically simple affixes, which were already spelt correctly at pre-test.

For reading, we also saw significant improvements from pre- to post-test. However, the improvements in the training group were not larger than those in the control group. Earlier evaluations of morphological intervention on older German speaking students

reported mixed effects on reading. Some studies reported training effects on sentence reading efficiency and comprehension, but only compared to control groups not receiving any intervention (Gebauer, Fink, Kargl, et al., 2012; Weiss et al., 2010), while others failed to replicate training effects on reading (Gebauer, Fink, Filippini, et al., 2012; Schneeberger et al., 2011). Since the focus of the current intervention was on spelling, the absence of a training effect on reading was less surprising. The finding that nonword reading improved more in the control than in the training group is likely due to the phonics-oriented instruction these children received during the training period. While children in the intervention group were instructed to use their morphological knowledge during spelling and reading, children in the control group practiced sounding out and blending phonemes, which seemed to particularly benefit their nonword reading efficiency. Thus, our findings provide tentative evidence that phonics-based teaching more effectively builds phonological decoding skills than a morpheme-based instruction.

For the two standardized word reading measures (silent sentence reading and oral word reading fluency), both teaching approaches were equally efficient. However, these reading measures were not specifically designed to test for morpheme-based reading strategies. They probably required a mix of phoneme-, morpheme-, and word-based strategies and may not be sufficiently sensitive to specific improvements in either of these strategies. In the morpheme-based reading tasks, we contrasted conditions in which an existing stem was combined with a real pre- or suffix or with a matched letter cluster that did not correspond to a morpheme. An impact of morphology on reading was directly evident in these tasks as all children found it easier to read the affixed compared to the matched pseudo-affixed condition, and both groups showed stronger improvements from pre- to post-test in the affixed compared to the pseudo-affixed conditions. Interestingly, the results revealed a specific training-related improvement in this task: In the prefix condition, the difference in improvement between prefixed and pseudo-prefixed conditions was larger in the intervention than the control group. We did not see such a specific training-related change in the suffix condition, where both groups showed comparable training effects across conditions (suffixed and pseudo-suffixed). It is possible that the position of suffixes at the end of items was less prominent for children processing words in a left-to-right fashion. The finding of a specific, training related effect on prefixes, is, however, promising.

### *Training Effects on Morphological Awareness*

A standard morphological awareness task in which children had to adapt word forms to a given

sentence was used to assess training effects on explicit morphological knowledge. While both groups performed clearly better during post- than pre-test, we did not see any evidence for intervention-related improvements. There are several reasons for the absence of an intervention effect on morphological awareness. First, the control group received some degree of morphological training as part of their language classes (even though this was not a specific focus). It is possible that this input was already sufficient to build their morphological awareness skills to a similar extent as the training group. Second, during training, only existing, high-frequency stems were combined with existing affixes, while in the morphological awareness task, pseudo-stems were used, which were (obviously) not in children's lexicon. It is possible that these pseudo-stems confused our relatively young participants. It should also be noted that our morphological awareness task was adapted from a written version for older students (TMB; Kargl et al., 2006). Indeed, an earlier study with poor and typical readers in Grades 3 to 8 reported larger improvements in training compared to control groups (Kargl et al., 2008) for the written form of the same task. In the current study, we used a spoken version of the task as children's spelling skills were probably not yet sufficiently developed for the written version. It is not clear if the verbal task version was sufficiently sensitive to capture changes in morphological awareness.

### *Training Effects on Morphological Processing*

For the first time, we included a masked priming paradigm to investigate children's morphological segmentation skills during written language processing. Two findings are particularly important. First, we observed an overall larger improvement in visual word recognition speed in the intervention compared to the control group, suggesting that students became more proficient at rapidly identifying printed words. Second, there was no evidence for training related changes in priming patterns that would indicate a boost in children's morphological processing skills. It is possible, that the masked priming paradigm was not sufficiently sensitive to capture small effect changes after an eight-week training period. More intensive, longer-term training may be required to induce more robust changes. The present findings suggest that the MORPHEUS intervention program did not elicit any changes in children's automatic morphological segmentation processes. However, the program did provide a boost to participants' overall word recognition speed.

Despite the absence of a training effect on masked priming, the observed priming effects shed light on how young readers decompose morphologically complex stimuli into their morphemic constituents during

reading. Significant priming effects in the two prefixed conditions suggest that children rapidly decomposed letter strings into prefix + stem, regardless of whether the morpheme combinations formed a real word (e.g., "mitdenken – DENKEN" – Engl.: think along – THINK) or a nonword (e.g., "hindenken – DENKEN" – Engl.: think towards – THINK). If the stem occurred in combination with a non-morphemic unit, priming effects were not observed (e.g., "kardenken – DENKEN"; Engl.: karthink – THINK). This pattern is similar to findings in skilled readers (Beyersmann, Ziegler, et al., 2016; Rastle & Davis, 2008; Rastle et al., 2004) who seemingly decompose words into their components by rapidly and automatically "stripping off" any orthographic string that resembles an affix, therefore also referred to as affix-stripping (Traft & Forster, 1975). In the suffix condition, however, only primes with a semantic relation to the target (word primes) induced priming effects. This pattern is different from other studies with children (Beyersmann, Grainger, et al., 2015; Beyersmann et al., 2021; Hasenäcker et al., 2016, 2020) and is suggestive of whole word processing (Giraudo & Grainger, 2001). These differences in priming patterns may be due to inherent differences of prefixes and suffixes regarding their semantic information. While manipulating both prefixes and suffixes can change the meaning of a word entirely, several prefixes like those used in this study ("auf-", "mit-", "ab-", "an-") can occur as separate prepositions and carry meaning on their own. Suffixes on the other hand need a stem to convey meaning (e.g., "-ly" typically marks an adverb). Therefore, prefixes have been described as having a quasi-lexical status (Beyersmann, Ziegler, & Grainger, 2015). In addition, if processing follows the reading direction (left to right), prefixes, at the beginning of a word, might be processed and detached first. Suffixes would only be detectable after stem identification. Our findings suggest that morphological processing may not strictly follow a whole-word-then-segmentation route or vice versa but rather allows for the parallel processing of whole words and morphemes. This would be in line with Grainger and Beyersmann's (2017) framework of a multiple-route model of word processing. Grainger's and Beyersmann's theoretical framework builds on the idea that embedded stems and also affixes can be activated via an entirely non-morphological process, with direct mappings from orthography to the lexical level (see also, Beyersmann et al., 2019).

The exact processes involved in and developmental trajectories of morphological priming should be subject to future research. In any case, the here reported morphological priming effects demonstrate that children as young as Grade 2 already make use of the structure inherent in morphologically complex words.



## Limitations

One limitation of our study is that it was not possible to randomly assign students to the experimental or control groups, because the training was delivered as part of children's language classes. For administrative reasons, we also decided against random assignment of classrooms. The classrooms that received the intervention were selected so that they were reachable for the two graduate students who carried out the teaching lessons.

A further limitation is that the business-as-usual control group also received some degree of morpheme-based instruction, which corresponded to the Grade 2 curriculum. The capitalization rule of German orthography is morpheme-based (all nouns are capitalized), such that the recognizing "name words" has always been part of the early spelling curriculum. Word building and word family exercises, such as recognizing the identical stem in word families, occurred occasionally in the control group, while it was the central focus within the morphology-intervention group. Thus, the two groups clearly differed in the amount of morpheme-based input. Note that we found more marked training effects on spelling in spite of the (reduced) occurrence of morphological instruction in the control group. As argued earlier, it is possible that the minimal morpheme-based input in the control group was sufficient to induce improvements in morphological processing similar to the trainings group, thus explaining the lack of group differences in these tasks. It is also possible that the eight-week training period was too short to reveal effects on morphological processing. More extensive training may be necessary to fully benefit from the morpheme-intervention program. Finally, it would have been interesting to carry out a follow-up assessment one or two months after the end of the intervention in order to test for long-term effects. Unfortunately, the time-line of the project and the academic year prevented such an additional assessment.

## Conclusions

The current study showed significant training-related improvements in orthographic spelling as well as in oral reading of morphologically complex words. These findings suggest that the MORPHEUS intervention program led to an overall boost in children's orthographic knowledge. At the same time, we were unable to provide direct evidence of training-related changes to children's explicit or implicit morphological processing skills. This may be due to methodological limitations of our tasks and study design, but it is also possible that the improvements in spelling and reading were not induced by improvements in morphological processing, but rather by more general factors like gaining knowledge on frequent letter patterns or

orthographic learning of individual words rather than morphemes.

## Acknowledgement

Elisabeth Beyersmann was supported by a Discovery Early Career Researcher Award (DECRA) by the Australian Research Council (DE190100850).

Reinhard Kargl discloses that he is the co-author of the commercial intervention program MORPHEUS.

The study design followed the ethical standards as laid down in the 1964 Declaration of Helsinki and its later amendments.

All parents provided written consent for assessment and/or training procedures, which were also approved by the local school authorities.

All authors agreed with the content and gave explicit consent to submit.

## References

- Allen, A. A., & Lembke, E. S. (2020). The effect of a morphological awareness intervention on early writing outcomes. *Learning Disability Quarterly*. <https://doi.org/10.1177/0731948720912414>
- Amenta, S., & Crepaldi, D. (2012). Morphological processing as we know it: An analytical review of morphological effects in visual word identification. *Frontiers in Psychology*, 3, Article 232. <https://doi.org/10.3389/fpsyg.2012.00232>
- Apel, K. (2014). A comprehensive definition of morphological awareness: Implications for assessment. *Topics in Language Disorders*, 34(3), 197–209. doi:10.1097/TLD.000000000000019
- Apel, K., Brimo, D., Diehm, E., & Apel, L. (2013). Morphological awareness intervention with kindergartners and first- and second-grade students from low socioeconomic status homes: A feasibility study. *Language, Speech, and Hearing Services in Schools*, 44(2), 161–173. [https://doi.org/10.1044/0161-1461\(2012/12-0042\)](https://doi.org/10.1044/0161-1461(2012/12-0042))
- Arnbak, E., & Elbro, C. (2000). The effects of morphological awareness training on the reading and spelling skills of young dyslexics. *Scandinavian Journal of Educational Research*, 44(3), 229–251. <https://doi.org/10.1080/00313830050154485>
- Ball, E. W., & Blachman, B. A. (1988). Phoneme segmentation training: Effect on reading readiness. *Annals of Dyslexia*, 38(1), 208–225. <https://doi.org/10.1007/BF02648257>

- Ball, E. W., & Blachman, B. A. (1991). Does phoneme awareness training in kindergarten make a difference in early word recognition and developmental spelling? *Reading Research Quarterly*, *26*(1), 49–66.
- Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge University Press.
- Bates, D., Maechler, M., Bolker, B., & Steve Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. <http://dx.doi.org/10.18637/jss.v067.i01>
- Berko, J. (1958). The child's learning of English morphology. *WORD*, *14*(2–3), 150–177. <https://doi.org/10.1080/00437956.1958.11659661>
- Berninger, V. W., Abbott, R. D., Nagy, W., & Nagy, W. (2010). Growth in phonological, orthographic, and morphological awareness in Grades 1 to 6. *Journal of Psycholinguistic Research*, *39*(2), 141–163. <https://doi.org/10.1007/s10936-009-9130-6>
- Beyersmann, E., Casalis, S., Ziegler, J. C., & Grainger, J. (2015). Language proficiency and morpho-orthographic segmentation. *Psychonomic Bulletin & Review*, *22*(4), 1054–1061. <https://doi.org/10.3758/s13423-014-0752-9>
- Beyersmann, E., Castles, A., & Coltheart, M. (2012). Morphological processing during visual word recognition in developing readers: Evidence from masked priming. *The Quarterly Journal of Experimental Psychology*, *65*(7), 1306–1326. <https://doi.org/10.1080/17470218.2012.656661>
- Beyersmann, E., Cavalli, E., Casalis, S., & Colé, P. (2016). Embedded stem priming effects in prefixed and suffixed pseudowords. *Scientific Studies of Reading*, *20*(3), 220–230. <https://doi.org/10.1080/10888438.2016.1140769>
- Beyersmann, E., Grainger, J., Casalis, S., & Ziegler, J. C. (2015). Effects of reading proficiency on embedded stem priming in primary school children. *Journal of Experimental Child Psychology*, *139*, 115–126. <https://doi.org/10.1016/j.jecp.2015.06.001>
- Beyersmann, E., Grainger, J., & Castles, A. (2019). Embedded stems as a bootstrapping mechanism for morphological parsing during reading development. *Journal of Experimental Child Psychology*, *182*, 196–210. <https://doi.org/10.1016/j.jecp.2019.01.010>
- Beyersmann, E., Mousikou, P., Javourey-Drevet, L., Schroeder, S., Ziegler, J. C., & Grainger, J. (2020). Morphological processing across modalities and languages. *Scientific Studies of Reading*, *24*(6), 500–519. <https://doi.org/10.1080/10888438.2020.1730847>
- Beyersmann, E., Mousikou, P., Schroeder, S., Javourey-Drevet, L., Ziegler, J. C., & Grainger, J. (2021). The dynamics of morphological processing in developing readers: A cross-linguistic masked priming study. *Journal of Experimental Child Psychology*, *208*. <https://doi.org/10.1016/j.jecp.2021.105140>
- Beyersmann, E., Ziegler, J. C., Castles, A., Coltheart, M., Kezilas, Y., & Grainger, J. (2016). Morpho-orthographic segmentation without semantics. *Psychonomic Bulletin & Review*, *23*, 533–539. <https://doi.org/10.3758/s13423-015-0927-z>
- Beyersmann, E., Ziegler, J. C., & Grainger, J. (2015). Differences in the processing of prefixes and suffixes revealed by a letter-search task. *Scientific Studies of Reading*, *19*(5), 360–373. <https://doi.org/10.1080/10888438.2015.1057824>
- Bowers, J. S., & Bowers, P. N. (2017). Beyond phonics: The case for teaching children the logic of the English spelling system. *Educational Psychologist*, *52*(2), 124–141. <https://doi.org/10.1080/00461520.2017.1288571>
- Bowers, J. S., & Bowers, P. N. (2018). The importance of correctly characterizing the English spelling system when devising and evaluating methods of reading instruction: Comment on Taylor, Davis, and Rastle (2017). *Quarterly Journal of Experimental Psychology*, *71*(7), 1497–1500. <https://doi.org/10.1177/1747021818759477>
- Bowers, P. N., & Kirby, J. R. (2010). Effects of morphological instruction on vocabulary acquisition. *Reading and Writing: An Interdisciplinary Journal*, *23*(5), 515–537. <https://doi.org/10.1007/s11145-009-9172-z>
- Bowers, P. N., Kirby, J. R., & Deacon, S. H. (2010). The effects of morphological instruction on literacy skills: A systematic review of the literature. *Review of Educational Research*, *80*(2), 144–179. <https://doi.org/10.3102%2F0034654309359353>
- Bus, A. G., & van IJzendoorn, M. H. (1999). Phonological awareness and early reading: A meta-analysis of experimental training studies. *Journal of Educational Psychology*, *91*(3), 403–414. <https://doi.org/10.1037/0022-0663.91.3.403>

- Carlisle, J. F. (2000). Awareness of the structure and meaning of morphologically complex words: Impact on reading. *Reading and Writing: An Interdisciplinary Journal*, 12(3), 169–190. <https://doi.org/10.1023/A:1008131926604>
- Casalis, S., & Louis-Alexandre, M. (2000). Morphological analysis, phonological analysis and learning to read French: A longitudinal study. *Reading and Writing: An Interdisciplinary Journal*, 12(3), 303–335. <https://doi.org/10.1023/A:1008177205648>
- Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. Harper & Row, Publishers.
- Dawson, N., Rastle, K., & Ricketts, J. (2018). Morphological effects in visual word recognition: Children, adolescents, and adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44(4), 645–654. <http://dx.doi.org/10.1037/xlm0000485>
- Deacon, S., & Kirby, J. (2004). Morphological awareness: Just “more phonological”? The roles of morphological and phonological awareness in reading development. *Applied Psycholinguistics*, 25(2), 223–238. <https://doi.org/10.1017/S0142716404001110>
- De Simone, E., Beyersmann, E., Mulatti, C., Mirault, J., & Schmalz, X. (2021). Order among chaos: Cross-linguistic differences and developmental trajectories in pseudoword reading aloud using pronunciation Entropy. *PLoS ONE*, 16(5), Article e0251629. <https://doi.org/10.1371/journal.pone.0251629>
- Ehri, L. C., Nunes, S. R., Willows, D. M., Schuster, B. V., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel’s meta-analysis. *Reading Research Quarterly*, 36(3), 250–287. <https://doi.org/10.1598/RRQ.36.3.2>
- Eichmeyer, A., Gönning, M., Kunze, H., Warnecke, A., von Werder, K., & Zoltan, G. (2013). *Karibu 2: Sprachbuch* [Karibu 2: Language book]. E. DORNER.
- Eichmeyer, A., & Zoltan, G. (2013). *Karibu 2: Material für Lehrerinnen und Lehrer* [Karibu 2: Material for teachers]. E. DORNER.
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G\*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. <https://doi.org/10.3758/BF03193146>
- Fischer, M. Y., & Pfost, M. (2015). Wie effektiv sind Maßnahmen zur Förderung der phonologischen Bewusstheit [How effective are measures to increase phonological awareness]? *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 47(1), 35–51. <https://doi.org/10.1026/0049-8637/a000121>
- Fox, J., & Weisberg, S. (2019). *An {R} Companion to Applied Regression* (Third Edition). Sage. <https://socialsciences.mcmaster.ca/jfox/Books/Companion/>
- Gebauer, D., Fink, A., Filippini, N., Johansen-Berg, H., Reishofer, G., Koschutnig, K., Kargl, R., Purgstaller, C., Fazekas, F., & Enzinger, C. (2012). Differences in integrity of white matter and changes with training in spelling impaired children: A diffusion tensor imaging study. *Brain Structure and Function*, 217(3), 747–760. <https://doi.org/10.1007/s00429-011-0371-4>
- Gebauer, D., Fink, A., Kargl, R., Reishofer, G., Koschutnig, K., Purgstaller, C., Fazekas, F., & Enzinger, C. (2012). Differences in brain function and changes with intervention in children with poor spelling and reading abilities. *PLoS ONE*, 7(5), Article e38201. <https://doi.org/10.1371/journal.pone.0038201>
- Giraudo, H., & Grainger, J. (2001). Priming complex words: Evidence for supralephical representation of morphology. *Psychonomic Bulletin & Review*, 8(1), 127–131. <https://doi.org/10.3758/BF03196148>
- Goodwin, A. P., & Ahn, S. (2010). A meta-analysis of morphological interventions: Effects on literacy achievement of children with literacy difficulties. *Annals of Dyslexia*, 60(2), 183–208. <https://doi.org/10.1007/s11881-010-0041-x>
- Goodwin, A. P., & Ahn, S. (2013). A meta-analysis of morphological interventions in English: Effects on literacy outcomes for school-age children. *Scientific Studies of Reading*, 17(4), 257–285. <https://doi.org/10.1080/10888438.2012.689791>
- Grainger, J., & Beyersmann, E. (2017). Edge-aligned embedded word activation initiates morpho-orthographic segmentation. In B. H. Ross (Ed.), *The Psychology of Learning and Motivation: Vol. 67* (pp. 285–317). Elsevier Academic Press.
- Hasenäcker, J., Beyersmann, E., & Schroeder, S. (2016). Masked morphological priming in German-speaking adults and children: Evidence from response time distributions. *Frontiers in Psychology*, 7, Article 929. <https://doi.org/10.3389/fpsyg.2016.00929>

- Hasenäcker, J., Beyersmann, E., & Schroeder, S. (2020). Morphological priming in children: Disentangling the effects of school-grade and reading skill. *Scientific Studies of Reading, 24*(6), 484–499. <https://doi.org/10.1080/10888438.2020.1729768>
- Heathcote, L., Nation, K., Castles, A., & Beyersmann, E. (2018). Do 'blacheap' and 'subcheap' both prime 'cheap'? An investigation of morphemic status and position in early visual word processing. *Quarterly Journal of Experimental Psychology, 71*(8), 1645–1654. <https://doi.org/10.1080/17470218.2017.1362704>
- Kargl, R., & Landerl, K. (2018). Beyond phonology. *Topics in Language Disorders, 38*(4), 272–285. doi:10.1097/TLD.0000000000000165
- Kargl, R., & Purgstaller, C. (2010). *MORPHEUS: Morphemunterstütztes Grundwortschatz-Segmentierungstraining* [Morpheme-based vocabulary-segmentation training]. Hogrefe.
- Kargl, R., Purgstaller, C., Weiss, S., & Fink, A. (2006). *Test zur Erfassung der morphematischen Bewusstheit (TMB)* [Test to measure morphological awareness (TMB)]. Unveröffentlichtes Testverfahren [unpublished]. Lese-Rechtschreibinstitut Graz.
- Kargl, R., Purgstaller, C., Weiss, S., & Fink, A. (2008). Effektivitätsüberprüfung eines morphemorientierten Grundwortschatz-Segmentierungstrainings (MORPHEUS) bei Kindern und Jugendlichen [Examining the effects of a morpheme-based vocabulary-segmentation training]. *Heilpädagogische Forschung, 34*(3), 147–156.
- Kirby, J. R., Deacon, S. H., Bowers, P. N., Izenberg, L., Wade-Woolley, L., & Parrila, R. (2012). Children's morphological awareness and reading ability. *Reading and Writing: An Interdisciplinary Journal, 25*(2), 389–410. <https://doi.org/10.1007/s11145-010-9276-5>
- 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. <https://doi.org/10.1080/13506280902986058>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in linear mixed effects models. *Journal of Statistical Software, 82*(13), 1–26. <https://doi.org/10.18637/jss.v082.i13>
- Landerl, K. (2017). Learning to read German. In L. Verhoeven, & C. Perfetti (Eds.), *Learning to read across languages and writing systems* (pp. 299–323). Cambridge University Press.
- Levesque, K. C., Kieffer, M. J., & Deacon, S. H. (2017). Morphological awareness and reading comprehension: Examining mediating factors. *Journal of Experimental Child Psychology, 160*, 1–20. <https://doi.org/10.1016/j.jecp.2017.02.015>
- Longtin, C.-M., Segui, J., & Hallé, P. A. (2003). Morphological priming without morphological relationship. *Language and Cognitive Processes, 18*(3), 313–334. <https://doi.org/10.1080/01690960244000036>
- Mayringer, H., & Wimmer, H. (2003). *Salzburger Lese-Screening für die Klassenstufen 1-4* [Salzburg reading-screening for Grades 1-4]. Huber.
- Moll, K., & Landerl, K. (2010). *Lese- und Rechtschreibtest (SLRT-II)* [Reading and spelling test (SLRT-II)]. Huber.
- Morris, J., Porter, J. H., Grainger, J., & Holcomb, P. J. (2011). Effects of lexical status and morphological complexity in masked priming: An ERP study. *Language and Cognitive Processes, 26*(4–6), 558–599. <https://doi.org/10.1080/01690965.2010.495482>
- Mousikou, P., Beyersmann, E., Ktori, M., Javourey, L., Crepaldi, D., Ziegler, J. C., Grainger, J., & Schroeder, S. (2020). Orthographic consistency influences morphological processing in reading aloud: Evidence from a cross-linguistic study. *Developmental Science, 23*(6). <https://doi.org/10.1111/desc.12952>
- Mussar, R., Sénéchal, M., & Rey, V. (2020). The development of morphological knowledge and spelling in French. *Frontiers in Psychology, 11*. Article 146 <https://doi.org/10.3389/fpsyg.2020.00146>
- Quémart, P., Casalis, S., & Colé, P. (2011). The role of form and meaning in the processing of written morphology: A priming study in French developing readers. *Journal of Experimental Child Psychology, 109*(4), 478–496. <https://doi.org/10.1016/j.jecp.2011.02.008>
- Rastle, K., & Davis, M. H. (2008). Morphological decomposition based on the analysis of orthography. *Language and Cognitive Processes, 23*(7–8), 942–971. <https://doi.org/10.1080/01690960802069730>

- Rastle, K., Davis, M. H., & New, B. (2004). The broth in my brother's brothel: Morpho-orthographic segmentation in visual word recognition. *Psychonomic Bulletin & Review*, 11(6), 1090–1098. <https://doi.org/10.3758/BF03196742>
- R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>
- Rispens, J. E., McBride-Chang, C., & Reitsma, P. (2008). Morphological awareness and early and advanced word recognition and spelling in Dutch. *Reading and Writing: An Interdisciplinary Journal*, 21(6), 587–607. <https://doi.org/10.1007/s11145-007-9077-7>
- Scheerer-Neumann, G. (1979). *Intervention bei Lese-Rechtschreibschwäche. Überblick über Theorien, Methoden und Ergebnisse* [Interventions for reading and spelling difficulties. Overview of theories, methods and results]. Kamp.
- Schiff, R., Raveh, M., & Fighel, A. (2012). The development of the Hebrew mental lexicon: When morphological representations become devoid of their meaning. *Scientific Studies of Reading*, 16(5), 383–403. <https://doi.org/10.1080/10888438.2011.571327>
- Schmalz, X., Beyersmann, E., Cavalli, E., & Marinus, E. (2016). Unpredictability and complexity of print-to-speech correspondences increase reliance on lexical processes: More evidence for the orthographic depth hypothesis. *Journal of Cognitive Psychology*, 28(6), 658–672. <https://doi.org/10.1080/20445911.2016.1182172>
- Schneeberger, B., Kargl, R., Purgstaller, C., Kozel, N., Gebauer, D., Vogl, J., Rohrer, S., & Fink, A. (2011). Förderung von Kindern und Jugendlichen mit Problemen im Schriftspracherwerb [Supporting children and teenagers with difficulties in written language acquisition]. *Zeitschrift für Heilpädagogik*, 12, 476–483.
- Schroeder, S., Würzner, K.-M., Heister, J., Geyken, A., & Kliegl, R. (2015). childLex–Eine lexikalische Datenbank zur Schriftsprache für Kinder im Deutschen [childLex – A lexical database for written language for children - in German]. *Psychologische Rundschau*, 66(3), 155–165. <https://doi.org/10.1026/0033-3042/a000275>
- Traft, M., & Forster, K. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior*, 14(6), 638–647. [https://doi.org/10.1016/S0022-5371\(75\)80051-X](https://doi.org/10.1016/S0022-5371(75)80051-X)
- Ulicheva, A., Harvey, H., Aronoff, M., & Rastle, K. (2018). Skilled readers' sensitivity to meaningful regularities in English writing. *Cognition*, 195. <https://doi.org/10.1016/j.cognition.2018.09.013>
- Walter, J., Schliebe, L., & Barzen, S. (2007). Evaluation eines morphemorientiert-strategischen Rechtschreibtrainings in schulischen Fördergruppen mit Grundschulern der 3. Klasse [Evaluation of a morpheme-oriented strategic spelling training with third graders]. *Heilpädagogische Forschung*, 33(3), 143–154.
- Weiß, R. H. (2006). *Grundintelligenztest Skala 2 (CFT 20-R)* [Intelligence scale]. Hogrefe.
- Weiss, S., Grabner, R. H., Kargl, R., Purgstaller, C., & Fink, A. (2010). Behavioral and neurophysiological effects of morphological awareness training on spelling and reading. *Reading and Writing: An Interdisciplinary Journal*, 23(6), 645–671. <https://doi.org/10.1007/s11145-009-9177-7>
- Zhang, H., & Zou, W. (2020). Morphological intervention in promoting higher-order reading abilities among college-level second language learners. *Sustainability*, 12(4), 1465. <https://doi.org/10.3390/su12041465>



**Appendix**

Examples for Training Tasks (Tablet and Paper-Pencil Tasks)

**Figure 1A**

Level 1: Identifying Words From the Same Word Family (Tablet).

rufen	vorlesen	hören
lesen	gehört	Leser
zuhören	anrufen	zurufen

LES	RUF	HÖR

Note: Children place each word in a specific box, depending on the stem morpheme, e.g., lesen, vorlesen, Leser (Engl.: to read, read to sb., reader) are placed into the box for the stem morpheme LES.

**Figure 2A**

Level 1: Assigning Different Words to the Corresponding Word Class (Paper-Pencil Task).

Ü12: Ordne die Wörter richtig zu.

lese, gut, Tag, rot, Lesebuch, rufe

Namenwörter

der	das
-----	-----

Tunwörter

ich	ich
-----	-----

Wiewörter

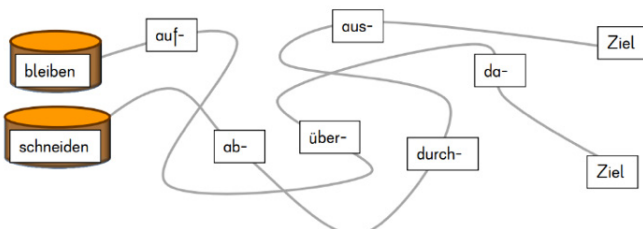
--	--

Note: Presented words are written into the corresponding row: noun (Namenwort, Engl.: name word; e.g., der Tag – Engl.: the day), verb (Tunwort, Engl. do-word; e.g., lese – Engl.: read) or adjective (Wiewort, Engl: how-word; e.g., rot – Engl.: red).

**Figure 3A**

Level 2: Combining Prefixes and Words Into Derived Words (Paper-Pencil Task).

Ü4: Wenn du einer Linie des Labyrinthes folgst, findest du Vorsilben, die mit dem Wortstamm ein sinnvolles Wort ergeben. Schreibe die gefundenen Wörter auf die Zeilen.



bleib:	schneid:

Note: The stem morpheme at the beginning of the line can be combined with each of the prefixes along the line, e.g., bleiben - aufbleiben, überbleiben ... (Engl.: to stay – stay awake, remain ...).

**Figure 4A**

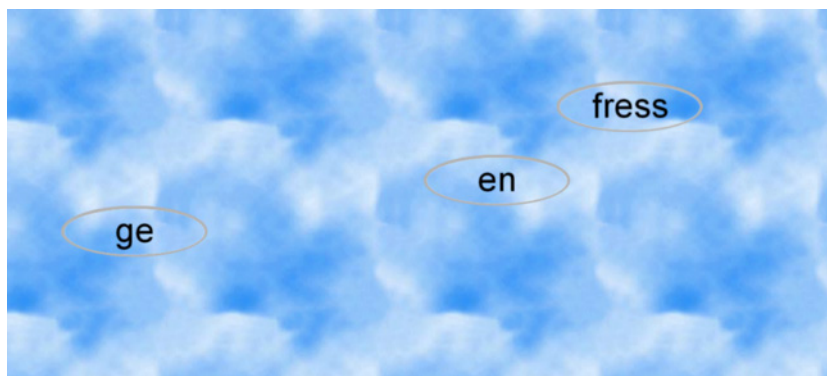
Level 3: Sorting Words According to Stem + Identifying (Short) Vowels Followed by Double Consonant (Paper-Pencil Task).

Ü5: Hier haben sich die Wörter von zwei Wortfamilien vermisch.  
 Trage sie in die richtigen Zeilen ein.  
 Schreibe den Selbstlaut in den Kreis und die Doppelbuchstaben in die Vierecke.

fallen      Stimme      Zufall      Abfall  
 stimmen      Einfall      Stimmung      gestimmt

**Figure 5A**

Level 3: Assembling Morphemes (Tablet).



Note: A prefix (ge), stem (fress) and suffix (en) need to be tapped in the right order to build a morphologically complex word (gefressen; Engl.: eaten).

**Figure 6A**

Level 3: Segmenting Morphologically Complex Words Into Prefix, Stem, and Suffix (Paper-Pencil Task).

Ü6: Trage die Wörter und ihre Bestandteile in die Tabelle ein.

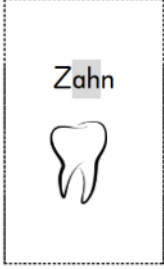

verbieten, vorspielen, anbieten, verspielen, mitspielen

Vorsilbe	Wortstamm	Nachsilbe
ver	biet	en

**Figure 7A**

Level 3: Compounding (Paper-Pencil Task).

Ü15: Zusammengesetzte Namenwörter. Bilde Wörter, indem du die Namenwörter zu einem Wort verbindest.

	Bürste	
	Arzt	
	Pasta	
	Rad	

Note: Stems are combined into compounds and written down (e.g., Zahnbürste, Zahnarzt, ...; Engl.: toothbrush, dentist, ...), each containing the noun Zahn (Engl.: tooth).

**Figure 8A**

Level 3. Morpheme Constancy (Paper-Pencil Task).

Ü11: Bilde Wörter und markiere die Doppelbuchstaben

zer + reißen	_____
aus + suchen	_____
vor + rennen	_____
auf + fressen	_____
auf + fallen	_____

Note: Children practice to combine and correctly spell words in which the prefix ends and the stem starts with the same letter.