Too Little Morphology Can Kill You: The Interplay Between Low-Frequency Morpho-Orthographic Rules and High-Frequency Verb Homophones in Spelling Errors

Dominiek Sandra

Abstract Many orthographies represent the morphological structure of words, i.e., keep the spelling of a morpheme constant despite variability in pronunciation (e.g., *cats, dogs*). Experimental work strongly suggests that this structure plays a beneficial role in both visual word recognition and spelling. Readers apparently decompose words into their constituent morphemes for the sake of lexical access. Moreover, early on, spellers rely on a word's morphological structure to derive its spelling (e.g., *picked, called*). However, morphologically complex words can also be a spelling hurdle, more particularly, when different morphological structures yield different spellings (i.e., morpho-orthographic representations) with the same pronunciation, i.e., grammatical homophones. The error risk on these homophones is codetermined by the token frequencies of the homophones, the rule's type frequency, and properties of working-memory. The focus in this chapter is on a salient error type in the spelling of Dutch verb homophones but is extended to other languages as well.

Keywords Homophones · Homophone intrusions · Morphological decomposition · Morphological awareness · Spelling errors · Rule frequency · Homophone dominance · Working memory

D. Sandra (✉)
Department of Linguistic, University of Antwerp, Antwerp, Belgium
e-mail: dominiek.sandra@uantwerpen.be

R. Levie et al. (eds.), *Developing Language and Literacy*, Literacy Studies 23,
https://doi.org/10.1007/978-3-030-99891-2_8

1 The Key Role of the Phonological and Morphological Principles in Alphabetic Orthographies

Since the birth of psycholinguistics, the major focus has been on the recognition of written words and the recognition and production of speech. This is not surprising. Spoken language research involves the primary function of language. In contrast to written language, speech is a naturally evolved human skill. Hence, the study as to which mental processes and structures enable our ability to produce and understand spoken language sheds light on the nature of the mental infrastructure that has evolved to make language possible.

When turning to the written modality, the focus was and still is on written word recognition. This is not surprising either. Although psycholinguistics before the Chomskyan era targeted totally different issues than those during and after that period, their major focus was on recognition processes and their underlying memory structures. Understanding highly automatic mental processes is a scientific challenge for cognitive psychology, and word recognition is such a process.

Because word spelling neither involves the primary function of language (speech) nor is a fully automatic processes, it has escaped the attention of many researchers of language. Another reason may be the difficulties that are involved in studying production processes. Measuring word recognition processes, like the accuracy in the recognition of briefly flashed words on a tachistoscope or reaction speed in a lexical-decision task, is easier than the study of spelling processes.

The present paper sets out from the conviction that some issues in writing do provide insight into the interface between language and cognition, in this case, word spelling (e.g., Sandra & Fayol, 2011). As Bar-On and Kuperman (2019) write: "the 'weak spots' of a language where spelling errors are abundant can expose specific mechanisms of word learning and processing" (p. 1121, see also Protopapas et al., 2013). The focus will be on one of those weak spots in the spelling of a particular language: homophonous verb forms in Dutch that are formed by regular inflectional rules of the concatenative type. However, it will become clear that the same problems emerge in several alphabetic orthographies. Hence, these spelling problems point to language-independent cognitive components in spelling.

The Dutch orthography is an alphabetic system. Its basic principle is "spell what you hear", i.e., the basis of all alphabetic orthographies. However, this phonological principle is not applied across the board. The second foundation of this spelling system is a morphological principle, which means that the spelling of a morpheme remains constant across all words or word forms in which it occurs. There would be no need for the latter principle if it did not clash from time to time with the phonological principle. In such cases, it overrides this basic principle. For instance, even though a [t] is heard in the Dutch word for dog, i.e., *hond,* [hɔnt], it is spelled with <d> because the [d]-sound that can be heard in the plural [hɔndən] is devoiced in word-final position. This dual encoding of linguistic structure in the orthography is not restricted to Dutch. It occurs in many languages, English and French only being two other examples. For instance, the past tense suffix in *picked, called, ended* is

pronounced differently but is consistently spelled as <ed>. An example from French is the case of silent letters. For instance, the final letter of *galop* ('gallop') is <p>, although it is not pronounced – [gɑlo] – because a [p] is heard in *galoper* [gɑlope] ('to gallop').

Both the phonological and the morphological principles confront young children with cognitive hurdles when learning to read and spell. Phonemes are the most abstract building blocks of spoken words. This explains why alphabetic writing systems were the last to emerge in the history of writing (Britannica, 2021). It also makes them the least accessible phonological units, which is why phoneme awareness develops later than all other phonological awareness skills (Anthony & Francis, 2005). Children find it hard to identify the sounds in, for instance, *cat.* This is counterintuitive for fluent readers, but we can re-experience the difficulty of this deceptively simple segmentation task when attempting to segment a word in a typologically different language: in our perception, the word's sounds blend into a seamless sound stream. Learning to read and write changes the children's perception of spoken words (Morais & Kolinsky, 2002), as they no longer hear an impenetrable sound stream but the individual sounds. The consensus in the literature is that phoneme awareness is a two-way street: a precursor of this skill that exists before literacy acquisition facilitates the phonological recoding of written words, which in turn enhances phoneme awareness. This results in a self-reinforcing, interactive process of word decoding (Share's self-teaching hypothesis Share, 1995), which eventually yields fluent technical readers and spellers.

When the orthography of a language is also governed by a morphological principle, children face another challenge. They must learn when this principle overrides the basic phonological principle of their orthography. As the review below will show, this requires a form of metalinguistic awareness that takes time to develop. Moreover, it may result in written word forms where the conflict between the two principles leads to persistent spelling problems. This will be the central theme of this text. However, let's first have a look at the good news.

The morphological principle recurs in the orthography of ~~so~~ many languages. This makes it likely that there is a reason why preserving morphemes' spellings takes precedence over representing a word's sounds. The most plausible explanation that comes to mind is that a constant spelling of the basic meaning-carrying units[1] of words facilitates word recognition and makes it easier to spell morphologically related words. It seems easier to recognize a word like *musician* (derived from *music*) than, for instance, *musishian,* the French word *galop* ('gallop', derived from *galoper*) than *galo* or *galeau,* or the Dutch word *hond* ('dog', singular of *honden*) than *hont.* Similarly, the use of a morphological relationship may facilitate the spelling of a word. For different languages, the people who were involved in designing spelling rules must have followed the same rationale: the principle "spell the same what means the same" is beneficial for readers and/or spellers. This makes sense as

[1] Note that I also refer to inflectional affixes, which have a very abstract meaning (e.g., 3r person singular past tense), i.e. their grammatical function.

the purpose of language is to communicate meaning. It is almost self-evident to express basic meaning units by giving them the same orthographic representation. However, if the orthographic representation of morphological structure indeed facilitates written word recognition and word spelling, how could it ever be harmful? This question seems to undermine the goal of my paper from the start. Why put something in the spotlight that seems impossible when following the above line of reasoning?

Before embarking on my endeavor, I must first present evidence from the literature that what seems logical is indeed true, i.e., that the morphological structure of written words is useful for visual word recognition and word spelling. Before doing so, a short note on the scope of this paper is in order. The literature on the role of morphemes in reading and, to a lesser extent, in spelling, is massive. It is an impossible task to write a solid review in a single chapter. In view of my goal to discuss the harm that morphological spelling rules may cause, i.e., regular Dutch verb homophones, I will restrict myself to inflected word forms and only refer to derived words when it is unavoidable.

Note that inflection is the territory where one would expect the largest benefits of morphological spelling rules. Inflectional affixes are a means for generating grammatical variants of a single word – plurals (*cats*), verb forms (*works, worked, working*), different grammatical cases (nominative, accusative, etc.). Being grammatical variants of a word, these forms result from mechanistic affixing operations (e.g. add *-s, -ed, -ing*). This is the case in Germanic and Romanic languages but also in a typologically different language like Hebrew (e.g., Ravid, 2001). This makes these word forms fully predictable (ignoring irregulars), both at the level of form and meaning. In contrast, derivational affixes are used to create new words. Neither their existence (e.g., *deep-depth* but not *steep-stepth*) nor their meaning (*revolve-revolution*) is predictable.

In Sect. 2, I will address some major findings on the role of morphology in visual word recognition. This is where I will also discuss experiments with derived words, as these have been the popular word type in this line of research. However, as will become clear, it is plausible that the insights into the processing of derived words, almost by implication, generalize to regularly inflected word forms. In Sect. 3, I will discuss the evidence bearing on the role of morphemes in spelling. Taken together, Sects. 2 and 3 will reveal the beneficial nature of morphological structure in written words. Finally, in Sect. 4, I will present evidence, essentially from spelling errors on Dutch verb forms but also from identical phenomena in other languages, that morphological structure can be harmful too. In Dutch, the facts show that even analytically transparent morphological spelling rules remain a difficult hurdle when it comes to their application, even for the best spellers. They are harmful because they do not fit the basic cognitive principles of the language user. That is why the study of word spelling can shed light on the interface between language and cognition.

2 The Role of Morphological Structure in Visual Word Recognition

The interest in morphology in psycholinguistics emerged in the mid-seventies of the last century. Murrell and Morton (1974) demonstrated that a short period of word memorization improved participants' recognition of a briefly flashed word when it was morphologically related to one of the words on the study list (e.g., *cars-car*). No facilitation was found when the word shared the same letter sequence but was morphologically unrelated (*card-car*). Their conclusion that the lexical representation of a morpheme is accessed in word recognition was supported by the results in Taft and Forster's (1975) seminal paper on the visual recognition of derived words. In their model, the morphemic representation of the stem provides access to the central lexicon, where the lexical representation of the derived word is stored. Hence, both the stem and the whole word are accessed.

Note that, from a linguistic perspective, it would have been strange if derived words did not have their own representation in the mental lexicon, as it is often impossible to infer a derived word's meaning from the meanings of the stem and the affix(es) (e.g., *revolution, revolver,* which are both etymologically related to *revolve*). Linguists would be more surprised to learn that the stem, which often cannot be used to compute the meaning of the derived word, is nonetheless automatically accessed. However, upon closer inspection, such a semantic perspective cannot support hypotheses on orthographic processing, as the early stages of lexical processing are unlikely to be affected by higher-level information about word meaning.

Stem access also occurs for regularly inflected word forms (e.g., *finds*). Later experiments by Taft (1979) demonstrated an effect of stem frequency for regularly inflected word forms matched on the frequency of the whole form (e.g., *sized* vs. *raked*). However, for my later arguments, Taft's most important finding was that high-frequency inflected word forms were recognized faster than low-frequency ones matched on stem frequency (e.g., *things* vs. *worlds*). This indicates that even regularly inflected forms are stored in the mental lexicon. As will become clear, this finding will support my claims in Sect. 4, as it will help explain why a particular type of morphologically complex words can be harmful in spelling. Again, linguists would not be surprised by Taft's finding that regularly inflected word forms are morphologically decomposed, as these forms result from the mechanistic application of affixation rules. However, they would be surprised to learn that the full forms are stored as well, as the meaning of these forms can be computed on the fly.

Using the technique of frequency manipulation, Baayen et al. (1997) reported similar findings for the singular form of Dutch nouns. Lexical decision times to the singular were determined by the summed frequency of both the singular and the plural, which consisted of the stem and the plural suffix *-en.* This finding, too, indicates that a regularly inflected form is decomposed into its stem and suffix, which increases the frequency of the stem representation. The frequency of the plural itself

also determined response speed, which supports the existence of separate lexical representations for these fully regular inflected forms as well.[2]

Using a different technique, i.e., morphological priming, Stanners et al. (1979), also concluded that inflected forms are morphologically decomposed in the process of lexical access. When priming the stem with an inflected form (*pours-pour, burned-burn, lifting-lift*), using an average of 10 intervening items between prime and target, they obtained equally strong priming effects as for identity priming (e.g., *pour-pour*). Several subsequent experiments confirmed full priming of regularly inflected forms on their stem (Fowler et al., 1985; Napps, 1989; Sonnenstuhl et al., 1999). These findings support the idea of prelexical morphological decomposition, which results in access to the stem representation. This type of experimental design makes it impossible to decide whether these items also have full-form representations. To do that, the inflected form should have been the target. Smaller priming of, for instance, *pour-pours* compared to *pours-pours* would suggest the existence of a full-form representation.

Despite the popularity of the priming paradigm, researchers started to question its validity for the study of lexical processing and representation. A clever experiment by Oliphant (1983) highlighted this. In a classical priming experiment, he found the typical facilitation effect for identical word repetition. However, when the repeated words appeared as primes in the instructions for the experiment, which had to be read aloud, he found no repetition effect. He concluded that effects in this paradigm depend on participants' conscious access to the primes and, hence, do not shed light on unconscious processes in the automatic process of lexical access. In their seminal paper Forster and Davis (1984) doubted the validity of visible primes as well, even on theoretical grounds. They argued that it is hard to explain why low-frequency words show stronger facilitation from identity priming than high-frequency words. This would mean that the repetition of low-frequency words would soon wipe out the typical frequency effect in word recognition experiments. However, this is at odds with the fact that frequency is a robust predictor of word recognition times. Hence, something had to be wrong with the priming paradigm itself. The authors convincingly demonstrated that the effects from visible primes are plagued by episodic memory effects and hence are problematic for the study of the mental lexicon. Episodic memory contains the memory traces for specific events we have experienced. For instance, our recollection that we saw a word or a related word on a screen a short time ago and that we made a certain response. Forster and Davis argued that the visible prime not only contacted a memory trace in the mental lexicon but also left an episodic memory trace, which was apparently more salient for high-frequency than for low-frequency words. In a series of experiments, they demonstrated that presenting the primes immediately before the target, but so briefly

[2] In contrast to Taft (1979) these authors did not find evidence for full-form representations of regular verb forms ending in *-en.* Baayen et al. do not offer a conclusive explanation and emphasize the homographic nature of the Dutch

suffix *-en,* which is used to mark the plural of verbs and nouns (and considerably more often for verbs).

(i.e., 60 ms)[3] that participants could not identify them, resulted in equally strong priming effects for high-frequency and low-frequency words. They concluded that a masked prime does not leave an episodic memory trace and can only initiate a lexical access process. Whatever the word's frequency, the prime offers the same head-start in lexical processing. They ruled out other explanations for their findings and concluded that the masked priming effect was purely lexical. Their paper demonstrated that masked priming is a powerful technique for the study of the lexical access process.

The masked priming technique revived interest in prelexical morphological decomposition. A crucial paper was published by Longtin et al. (2003), who reported equal facilitation on response times when visual targets were primed by masked transparent derivations (*gaufrette-gauffre,* 'wafer'-'waffle'), opaque derivations (*fauvette-fauve,* 'warbler'-'wildcat'), or pseudo- derivations (*baguette-bague,* 'little stick'-'ring'). They found inhibition when the prime was the concatenation of a stem spelling and a letter sequence that did not match a suffix (*abricot-abri, 'apricot'-'shelter'*). Importantly, pseudo-derivations were not derivations at all but words with a so-called surface morphological structure: a concatenation of the orthographic sequences of a stem and a suffix in a monomorphemic word. The authors interpreted their data as evidence for a prelexical process of blind morphological decomposition, which only operates when the orthographic string is the concatenation of a potential stem and a potential suffix. This process is blind to the true morphological status of the segmented parts, and as a prelexical process, by definition, can have no access to information about form-meaning units. As the ending *cot* in *abricot* does not match the spelling of a French suffix, the process does not segment the letter string into *abri* + *cot,* so that the lexical representation of *abri* is not primed.

The Longtin et al. paper put Taft and Forster's concept of prelexical morphological decomposition in the spotlight again. Longtin and Meunier (2005) found further support for this process, using masked pseudowords that consisted of a non-existing combination of a stem and suffix. These pseudowords were either interpretable (*rapidifier-rapide,* 'quickify'-'quick') or not (*sportation-sport,* 'sport' + verbal suffix). Their control items were the concatenation of a stem and an existing orthographic word ending that could not be a suffix (e.g., *rapiduit-rapide, 'quick'* + non-suffix). They only obtained facilitation on lexical decision times from pseudowords with a morphological surface structure (potential stem + potential suffix), i.e., from *rapidifier* and *sportation* but not from *rapiduit*.

Rastle et al. (2004) reported evidence from English that confirmed Longtin et al.'s notion of blind prelexical morphological decomposition. In a masked priming paradigm with a lexical decision task, they found equally strong facilitation for pseudo-derivations (*corner-corn*) and semantically transparent derivations (*cleaner-clean*). As in the Longtin et al. study, primes with the same orthographic overlap but with no surface morphological structure (*brothel-broth*) produced no facilitation.

[3] Many experiments that were published later used shorter interstimulus intervals, e.g., 50 ms or sometimes smaller.

Some studies also provided evidence for early semantic effects in the processing of morphologically complex words (Diependaele et al., 2005; Feldman et al., 2015). Note that such findings do not contradict the existence of a bottom-up process of blind morphological decomposition. This process may initially be blindly driven by morpho-orthographic information but quickly interact with partially activated lexical representations through a process of top-down activation.

Closer to our focus on inflected word forms, the masked priming technique was also applied to inflected word forms and has been used in combination with ERP, MEG, and fMRI data. Using 50 ms masked primes in combination with ERP measurements, Royle et al. (2012) found that masked regular past tenses in French facilitate lexical decision times on their stem (e.g., *cassait-casse,* 'broke'-'break'). In contrast, they found null effects for prime-target pairs that were semantically related (synonyms) or only orthographically related (e.g., *cassis-casse,* 'blackcurrant'-'break'). The morphological priming condition also left a unique signature on the ERP data for the morphological pairs: an early morpho-orthographic effect at 250 ms post onset (N250) and a strong effect around 400 ms post onset (N400) in the morphological condition only. However, the problem of morphological processing has not yet been fully unraveled. The title of Leminen et al. (2019) paper – "Morphological processing in the brain: The good (inflection), the bad (derivation) and the ugly (compounding)" – emphasizes our incomplete understanding of morphological processing in visual word recognition. Yet, for inflected word forms, they find a lot of converging evidence in the data from neuroimaging techniques. They conclude that most EEG and MEG studies indicate that regular forms are accessed and decomposed earlier than irregular ones and mobilize different memory systems: procedural versus declarative memory. Furthermore, EEG, MEG and fMRI data suggest a different topographical distribution of the activation patterns triggered by these two types of inflected word forms: the processing of the regular forms involves areas that are typically involved with the procedural memory network.[4]

The literature on visual word recognition suggests an important role of morphology in the processing of morphologically complex words, both for inflected word forms and derived words. The evidence favoring a process of blind prelexical decomposition is quite convincing. Even though the concept originated in the domain of derivational morphology (Taft & Forster, 1975), it is supported by several findings for inflected word forms. For instance, the frequency effects of stems and whole word forms (Taft, 1979; Baayen et al., 1997) and the masked priming effects in the neuroimaging data discussed by Leminen et al. (2019) are compatible with the process of blind morphological decomposition proposed by Longtin and

[4] But see Morris and Stockall (2012), who report equal morphological priming from regular and irregular past tense primes (*sold-sell* vs. *walked-walk*) at the N250 component in an ERP experiment with 50 ms masked primes. A striking result, as past tenses that do not share the orthographic form of their stem have the same effect as those that do. The authors conclude: "For this to be possible, the morphological relationship between 'sold' and 'sell' must be accessible to early stages of form based, pre-semantic processing." (p. 91).

colleagues and Rastle and colleagues. This also makes sense from a theoretical perspective. It would be difficult to defend the existence of blind morphological decomposition for derived words while questioning such a process for inflected forms, as the latter can be generated mechanistically. The question is, of course, ultimately an empirical one – and a strong empirical argument favoring this position is provided by the blind morphological decomposition effects, which are insensitive to the semantic transparency of a derived word (*gaufrette* vs. *fauvette*), the distinction between derived and pseudo-derived words (*gaufrette* vs. *baguette*), and the distinction between existing words and non-existing words like *rapidifier* and *sportation.* If the process is operational whenever the written word has a surface morphological structure, it should also decompose regularly inflected forms.[5] At the same time, several findings indicate that there is also a full-form representation of inflected forms, even though their existence and meaning can be predicted, simply as the result of exposure frequency. Apparently, repeating a regular word form sufficiently often yields a full-form orthographic representation, just as is the case with monomorphemic words.

3 The Role of Morphological Structure in Spelling

3.1 Morphological Relations between Words are Beneficial at an Early Age

If morphological structure matters for word reading, what about word spelling? Do spellers also rely on a word's morphological make-up? There is a lot of evidence that they do, in some cases even from a very early age.

Rebecca Treiman and coworkers demonstrated that many young children soon benefit from a word's morphological structure to spell a word. In American English, the phoneme /t/ is pronounced as the flap [t] in intervocalic position, as in words like *duty* and *dirty.* Treiman et al. (1994) found that children between 5 and 8 years old made fewer errors on words like *dirty,* which are derivations in which a stem is followed by a suffix (*dirt* + *y*) than on matched mono-morphemic words like *duty,* which offer no such help. Young children seem to be able to rely on their morphological awareness that *dirty* is derived from *dirt* and, for that reason, is spelled with a *t.*

Treiman and Cassar (1996) contrasted inflected words like *tuned* and *faced* (mostly regular past tenses and a few plurals) with matched monomorphemic ones like *brand* and *feast.* Young children often experience problems with the spelling of consonant clusters, often leaving out the first consonant. Cassar and Treiman wondered whether these omissions were caused only by phonological factors (e.g.,

[5] Note that a blind decomposition process that is tuned to the presence of orthographic sequences matching affixes makes no predictions with respect to compounds.

difficulties in perceiving the vowel and a subsequent liquid or nasal as two distinct speech sounds) or also by children's knowledge of the words' morphological structure. They found fewer omission errors of the first consonant in inflected word forms, across several experimental tasks. This result, too, indicates that morphological structure can help even young children to spell correctly (for similar results, see Bourassa and Treiman, 2008).

However, there is a methodological problem with pairs like *brand* vs. *tuned.* The items are matched on their final consonant cluster, but there is a systematic frequency mismatch: the stem of inflected items occurs in several inflected word forms, making it a higher-frequency sequence than the corresponding letter string in the control words. Deacon and Bryant (2006a) remedied this possible contamination by comparing inflected and derived words to monomorphemic controls matched on the stem's orthographic pattern, like *rocked* vs. *rocket.* The items were pronounced in a sentence context, then repeated in isolation, upon which the children had to write the letter sequence preceding the final consonant cluster (which was provided, e.g., *rock* in ___*ed* vs. ___*et*). The 6- to 8-year-old children performed better on the bimorphemic items than on their controls. Apparently, these children had previously noticed the stem in derived words and now used this knowledge to their advantage when having to spell its letter pattern. They could do so in derived words but not in monomorphemic words. Deacon and Bryant (2006b) reported the same outcome for 7- to 9-year-old children: the same letter sequence (e.g., *turn*) was spelled better when it was the stem in a derived word (*turning*) than when it had no morphemic status (*turnip*).

Comparable findings were reported for French. In this language too, young children have been shown to make use of morphological relations in spelling. In French, the final letter of written words is often not pronounced (e.g., *tabac,* [taba], 'tabacco'). Sénéchal (2000) and Sénéchal et al. (2006) compared two sets of words ending in a silent letter: those whose silent letter could be recovered from the pronunciation of derived words, and those that did not allow this (controls). For instance, the *p* in *galop* ([galo], 'gallop') can be recovered from *galoper* ([galope], 'to gallop'), whereas the *c* in *tabac* can only be spelled from memory. The children, who were between 9 and 10 years old (Grade 4), spelled a silent letter more often in the morphological condition, thus demonstrating their ability to rely on morphological relations between words for spelling purposes at an early age.

Pacton et al. (2012) followed the reverse rationale. They used words whose spelling mismatched the spelling that was predicted by this strategy. For instance, the French word *numéro* ('number') should be misspelled as *numérot* if spellers apply a morphological strategy. Indeed, the derived verb *numéroter* ('to assign numbers to'), in which a [t] sound can be heard, suggests that a *t* must be spelled when the root is not followed by a suffix. As predicted, good spellers made overgeneralization errors reflecting the use of such a morphological strategy. Incidentally, note that this is a case where morphology is harmful (see Sect. 4).

Another rationale was used by Casalis et al. (2011). Morphological relationships should help in the spelling of orthographically ambiguous vowels, i.e., vowels with different orthographic realizations across words. For instance, the vowel sound in

lait, i.e., [ɛ], can be spelled as *ai* (*lait,* 'milk), *è* (*très,* 'very'), *ê* (*fête,* 'party'), or *ei* (*neige,* 'snow'). The authors dictated derived words like *laitage* ('dairy product') and controls like *failaise* ('cliff') in Grade 3 (age: ± 8 years) and Grade 4. The ambiguous vowel was spelled correctly significantly more often in derivations than in controls. This finding converges with the above findings that young children already rely on morphological relationships during spelling.

Pacton and colleagues reported interesting experiments with pseudowords. A potential problem with Deacon and Bryant's study was the presentation of the test items. Presenting the final letters (e.g., ___ed) may have triggered a morphological strategy. Pacton et al. (2013) also matched the frequency of the critical orthographic sequence in derived and control words by presenting 8-year-olds with pseudowords in an orthographic learning paradigm. This method enabled full control over the exposure frequency to the critical letter sequence. The pseudowords were presented in the context of a series of short stories, which children had to read silently. Control items appeared seven times in a story (e.g., *modoit*), whereas the critical items appeared five times as a monomorphemic item (e.g., *vensoit*) and twice with a real French suffix (e.g., *vensoite, vensoitiste*). Grade 3 children benefited from the morphological relationships. Note that the critical items were seen less often (as isolated letter strings) than the control items. Apparently, the morphological relationships overruled this difference in whole-word frequency.

Pacton et al. (2018) used an extra control condition to rule out another interpretation. The effect might be due to the presence of an extra phonological cue for the spelling of the silent letter in the derived words (e.g., *vensoitiste*) and/or the presentation of each critical item in several words (i.e., three). Using the same technique, they compared how well Grade 3 and Grade 5 children learnt the spelling of items like *coirard* in a so-called opaque condition (seven presentations of *croirard*), a morphological condition (five presentations of *croirard* and two derived words: *coirarde* and *coirardage*), and an orthographic control condition (five presentations of *croirard* and two words in which a non-suffix was appended to the pseudoword: *coirardume* and *coirardore*). Even though the morphological and orthographic conditions were matched on the two potentially confounding factors in Pacton et al. (2013), the children in the morphological condition (in both grades) again performed better in a forced-choice task with three alternative spellings (e.g., *coirard, coirars, coirar*).

Another approach was taken by Bar-On and Kuperman (2019) and Gahl and Plag (2019), who looked for patterns in spelling errors. Their studies bring us closer to home, as they adopted the same rationale as we did in our studies on 'the Dutch spelling problem': a systematic study of error patterns should shed light on the factors that trigger the errors and determine their nature. Gahl and Plag (2019), like Sandra (2010), emphasize that this is the same methodological approach as the one in studies of speech errors, which were also used as a means for discovering the representations, processes, and temporal dynamics of the production process (e.g., Fromkin, 1971; Garrett, 1975). Bar-On and Kuperman studied the erroneous intrusion of a vowel letter in Hebrew words and found that the majority did not disrupt

the word's morphological structure. Gahl and Plag studied errors in the spelling of suffixes like *-able/-ible* (e.g., *acceptible, accessable*) and *-ence/-ance* (e.g., *avoidence, occurrance*), and found that the error risk was determined by the strength of the morphological boundary, i.e., the ease with which the stem could be segmented from the suffix (e.g., due to a high ratio of stem frequency to whole word frequency). Importantly, these spelling errors did not reveal a preference for the higher-frequency suffix (which matches our findings in Dutch, cf. Sandra et al., 1999). The major error determinant was the segmentability of the derived words.

3.2 The Importance of Morphological Awareness

Many researchers who have studied the beneficial impact of morphological relationships on young children's spelling have also addressed the question whether this was caused by (or at least correlated with) their level of morphological awareness. The term 'morphological awareness' is rather vague, as is the term 'phonological awareness' (Uppstad & Tønnessen, 2007). Different researchers operationalize it in different ways. Sometimes, it is measured by asking participants to infer a morphological rule from an example and apply it to another word (e.g., *help-helped, live-?* or *run-runner, teach-?*). Other researchers measure it by asking participants to identify morphemes in words (e.g., *teach* is a part of *teacher*). Despite such differences, several researchers have demonstrated that children who obtain good scores in a morphological awareness task more often make use of morphological relationships in spelling, even when the pronunciation of a morpheme varies across words. This has been found for (a) suffixes whose pronunciation varies as the result of a preceding sound, like the English past tense suffix *-ed* (Nunes et al., 1997a, b), (b) silent word-final letters in French (Sénéchal, 2000; Sénéchal et al., 2006), (c) orthographically ambiguous vowels in the stem of morphologically complex words (Casalis et al., 2011), and (d) the stem of derived words (Deacon & Bryant, 2006b). Casalis et al. (2011) demonstrated that morphological awareness cannot be reduced to phonological awareness but adds a significant and independent contribution to the spelling of morphologically complex words. Interestingly, Deacon and colleagues reported results indicating a positive impact of morphological awareness on spelling in general, i.e., not only on words whose spelling depends on morphological relationships (Deacon & Bryant, 2006b). Perhaps being sensitive to words' morphological structure or being consciously aware of their morphemes implies a general interest in words, which includes morphological relations but is not restricted to it.

Several studies have investigated whether literacy skills are better in schools whose curriculum includes an explicit instructional goal to train children's morphological awareness. Carlisle (2010) reviewed the literature on the relationship between the instruction of morphological awareness and the key components of literacy achievement, i.e., phonological awareness, orthographic development (visual word recognition and spelling), and meaning. She found only seven studies that addressed effects on orthographic development. There are indications that

working towards this goal in teaching English improves spelling performance. However, the sample is small and factors like the age of the target group make it difficult to arrive at a firm conclusion. Plausibly, an explicit focus on words' morphological structure must come at the right age. Which age is the right age seems to depend on the morphological complexity of the language. Carlisle emphasizes that all four studies on Chineseconsistently showed that morphological awareness instruction improves both character reading and writing, even at a very early age: from kindergarten to Grade 4.

Research by Ravid and coworkers indicates that morphological awareness is also related to the typology of the language (e.g., Ravid, 2001, 2012; Ravid & Bar-On, 2005, Gillis & Ravid, 2006). Hebrew is a morphologically rich language: all verb forms and most nouns and adjectives consist of a root and a morphological pattern of affixes, the former being a discontinuous sequence of three or four consonants, which is interdigitated with vowels, i.e., the pattern (Ravid, 2001). Children who acquire Hebrew as their native language quickly catch on to spoken words' morphological structure and thus develop the skill of attending to it. They bring these morphological skills to their spelling performance. For instance, Ephratt (1997) found that gradeschoolers who were asked to color three letters of their own choice in a word consistently colored root letters, and Ravid and Bar-On (2005) found stronger priming from genuine roots than from identical pseudo-roots. This heightened root sensitivity makes children considerably better in using morphological cues for spelling neutralized phonological segments compared to children with a morphologically poorer native language, like Dutch (Gillis & Ravid, 2006).

4 When Morphology Hurts

Notwithstanding (a) the observation that young children already attempt to apply morphological relationships when spelling and (b) the reports that morphological awareness improves spelling performance on morphologically related words, one should not conclude that spelling errors on these words always betray a weak morphological awareness. Indeed, sometimes morphology hurts spellers with a high level of morphological awareness. There is evidence that some morphologically complex words cause unexpectedly many spelling errors, even among those with a high level of morphological awareness. Sometimes, the nature of some morpho-orthographic spelling rules does not seem to fit general cognitive principles that are used for spelling, and, consequently, is a source of persistent spelling errors.

In the following paragraphs I will summarize the experimental work that several of my collaborators and I have carried out on what is the greatest stumbling block in written Dutch: the spelling of (some types of) regular verb forms. The phenomenon and its causes in our cognitive infrastructure are fascinating, especially because this problem is not restricted to this language. It surfaces in several orthographies, in different disguises, and thus sheds light on general, i.e., not language-specific, cognitive phenomena. Whether these aspects of our cognition cause spelling

problems or not depends entirely on the nature of the spelling rules of the language. A considerable spelling hurdle emerges when there is a clash between the two major spelling principles behind the orthography of an alphabetic language: the phonological and morphological principles. It is a hurdle at which even the best spellers fall from time to time, and this is precisely what makes it interesting for researchers interested in language and cognition. How can that be?

4.1 The Morphological Principle in Dutch: Uniform Stem Spelling and Analogical Suffix Spelling

Table 1 exemplifies the spelling rules for regular verb forms in Dutch. The examples in columns 3–6 highlight their transparency, at least from a descriptive point of view. They show what spellers need to know. Present tense: 1st person singular =

Table 1 The spelling of Dutch verb forms, as a function of the phonological properties of the stem

Grammatical function			Stem-final phoneme		
		no *d*/*no t*	*d*	*t*	no *d*/no *t* weak prefix
Infinitive		werken	leiden	testen	bedoelen
		wɛrkən	lɛidən	tɛstən	bədulən
		(to work)	(to guide)	(to test)	(to mean)
Present	1st singular	werk	leid	test	bedoel
		wɛrk	lɛit	tɛst	bədul
Present	2nd singular	werkt	leidt	test	bedoelt
		wɛrkt	lɛit	tɛst	bədult
Present	3rd singular	werkt	leidt	test	bedoelt
		wɛrkt	lɛit	tɛst	bədult
Present	plural	werken	leiden	testen	bedoelen
		wɛrkən	lɛidən	tɛstən	bədulən
Past	singular	werkte	leidde	testte	bedoelde
		wɛrktə	lɛidə	tɛstə	bəduldə
Past	plural	werkten	leidden	testten	bedoelden
		wɛrktən	lɛidən	tɛstən	bəduldən
Past Participle		gewerkt	geleid	getest	bedoeld
		ɣəwɛrkt	ɣəlɛit	ɣətɛst	bədult
Imperative		werk	leid	test	bedoel
		wɛrk	lɛit	tɛst	bədul

stem, 2nd and 3rd person singular[6] = stem + <t>, all persons in the plural = infinitive. Past tense: stem + <te > when a voiceless consonant is heard in the infinitive (spelled as *p, t, k, f, s, ch*),[7] otherwise stem + <de>. Past participle: stem + <t> or <d>, depending on the consonant in the past tense suffix. Imperative: stem. There is no more to it.

In more than 90% of the cases (type-wise and token-wise, Sandra & Van Abbenyen, 2009) the suffix that must be spelled can be derived from the verb form's pronunciation, as illustrated by the verb *werken* ('to work) in Table 1. For instance, the final sound in *werkt* is pronounced, like the suffix sound in the English verb form *works.* Similarly, the two sounds of the past tense suffix are pronounced: *werkte*, [wɛrktə], 'worked'. The correspondence between pronunciation and spelling also holds in verb forms whose stem-final phoneme triggers the past tense allomorph [də], spelled <de> (e.g., daalde, [da:ldə], 'descended'). This is also the case, for instance, when the 2nd person singular in the present tense is mentioned before its subject: one does not hear a [t], and, hence, one does not spell <t> (e.g., *werk je,* [wɛrk jə], 'work you' vs. *je werkt,* [jə *wɛrkt*], 'you work).

When using a verb like *werken* as an example, readers who are unfamiliar with Dutch might think that Dutch verb forms are always spelled in accordance with the principle "spell what you hear". However, the above morphology-based description, in terms of stem and suffix, and the verbs *leiden* ('guide'), *testen* ('to test'), and *bedoelen* ('to mean') in Table 1 clearly show that this is not the case. At the stem level, this becomes clear in the spelling of inflected verb forms whose stem ends in the phoneme /d/ (e.g., *ik leid,* 'I guide'). Due to final devoicing a [t] is heard, i.e. [lɛit]. Spellers must recover the infinitive to decide whether they must spell a *t* or a *d:* we spell a *d* in *ik leid* because the infinitive *leiden* is pronounced with a [d], i.e., [lɛidən]. The principle that causes orthographic constancy of the stem implicitly adopts the idea of underlying phonemes: the phoneme /d/ underlies the sound [t] in *leid.* Obviously, it does not matter whether underlying phonemes are ghost entities that have originated in linguists' analytical minds rather than objective facts like those on which theories in physics and biology are founded.[8] Whatever one's take on this matter, the Dutch spelling rules require the stem to have a constant spelling across all verb forms in the inflectional paradigm.[9] This is known as the 'principle of uniformity' in Dutch spelling.

---

[6] The 2nd person singular is spelled without a <t> if the verb form precedes the subject.

[7] Language users have no problem in choosing between *-te* and *-de,* as the consonant is the result of an automatic phonological process of progressive assimilation, i.e., the first phoneme of the suffix is voiced/voiceless when the last stem phoneme is voiced/voiceless.

[8] I do not reject this kind of linguistic analysis. However, it is important to remain aware that this analysis has shaped our perception of language and can yield the erroneous belief that this perception matches the reality of language. Rather, this 'received' knowledge reflects a linguistic, hence, analytical, perspective on language.

[9] There are systematic exceptions. The voiced consonants /v/ and /z/ in verbs like *blijven* ('to stay') and *reizen* ('to travel') are not written as *v* and *z* in word-final (devoiced) position, as the morphological principle would require, but as their devoiced counterparts *f* and *s* (e.g., *blijf,* 'stay', *reis,* 'travel'). Despite this inconsistency, spellers make no errors on these verb forms.

The morphological principle also holds at the suffix level. For instance, the 3rd person singular of the present tense of *leiden* is *leidt.* It is pronounced as [lɛit], due to final devoicing of the stem-final consonant, but so is the form *leid.* However, *leidt* is spelled with <t> at the end, the orthographic marker of a suffix, because a [t] sound follows the stem in a form like *maakt,* [ma:kt] ('makes'). Hence, the spelling *leidt* reflects the application of the morphological principle at the stem level and at the suffix level. The principle of preserving the spelling of the suffix also determines the spelling of the past tense. For instance, the past tense of *leiden* is *leidde,* pronounced as [lɛidə] – the double *d* is pronounced as [d] (degemination) – because it is the concatenation of the stem *leid* and the suffix allomorph *-de.* The suffix is spelled as *-de* in analogy with a past tense like *belde,* [bɛldə], 'called', where the sound sequence [də] is heard after the stem. Hence, application of the morphological principle yields the doublet <dd> in past tense forms of verbs whose stem spelling ends in <d>. The same analogical reasoning applies to a past tense form like *testte,* pronounced as [tɛstə] ('tested'): the stem spelling *test* is concatenated with the suffix spelling *-te* because the sound sequence [tə] is heard after the stem in a past tense like *maakte,* [ma:ktə] ('made').[10] Not surprisingly, the morphological principle that governs the spelling of suffixes is known as the 'analogical principle'.

To summarize, the spelling of regular Dutch verb forms is driven by a morphological principle, which comprises two subprinciples: one at the stem level and one at the suffix level. The principle of uniformity stipulates that the spelling of the stem remains constant across all regular verb forms in which it occurs. The principle of analogy stipulates that the spelling of the suffix is spelled in analogy with its spelling in inflected forms with the same grammatical function of verbs whose suffix can be heard. From an analytic perspective, this is all very straightforward and transparent.

4.2 When Phonology and Morphology Clash: The Tragedy of Regular Verb Homophones

Despite this descriptive simplicity, many errors on Dutch verb forms are made. The real problem is the persistence of the errors, as not only beginning spellers make them but experienced spellers as well. One encounters them in newspapers, subtitles, headlines in television journals, exams and papers written by students studying (even language students), and occasionally … their professors. Some people make more of these errors than others (see below) but we should learn a humble lesson from the errors' remarkable tenacity: some morpho-orthographic rules can be simple in description but difficult in practice. This is surprising, as the rules are taught

[10] This analogical principle is not applied if it would yield the doublet <tt> in final position. An inflected form like *testt* is orthographically impossible in Dutch: reduplicated consonants never occur in word-final position. Hence, a general orthographic principle overrules this morphological principle for this subset of verbs. Spellers have no problems with this inconsistency.

at a young age (Grade 4, about age 10) and errors in them are laden with stigma from elementary school to university level. Even in society at large, there is virtually zero tolerance for spelling errors on verb forms. The irony: despite this, everybody makes them, at least occasionally.

The spelling of regular verb forms in Dutch presents us with a paradox: errors on them are persistent even though (a) they are descriptively easy, (b) they are taught at an early age, (c) students are reminded of their importance throughout their school career, and (d) there is a stigma on them in society. Even though many people like to link these errors to a general norm relaxation in society and a tendency to lower the bar in education, such an account cannot explain that a doctoral dissertation with the significant title "The tragedy of the verb forms" (van der Velde, 1956) was published more than 60 years ago. Nor can it explain why highly experienced spellers, like journalists and teachers at all educational levels, cannot avoid making these errors from time to time. Clearly, there must be something about these simple rules that nonetheless makes them difficult. This phenomenon, which is obviously a pain in the neck of language purists, makes it even more interesting for psycholinguists, as a systematic investigation of these errors can shed light on our cognitive infrastructure for spelling. This is what the Dutch story has taught us.

4.2.1 Verb Homophones

This error persistence is also remarkable against the background of children's early sensitivity to morphological awareness (see Sect. 3.1.). The Dutch 'verb tragedy' seems to fly in the face of what is known in the international literature. This literature would predict a difficult learning curve (as for the English past tense, e.g., Nunes et al., 1997a, b) but not the fact that these errors are so persistent that they even survive in texts of highly educated spellers. Clearly, the errors' resistance against prolonged efforts in spelling education and a considerable amount of stigmatization highlights an aspect of our cognitive infrastructure that has remained under the research radar for a long time. This is what makes them so intriguing: whereas problems with morpho-orthographic spelling rules in other languages often (largely) disappear when children grow older (provided they follow a typical literacy development), the errors we studied continue to plague even experienced spellers from time to time.

We started our investigation (Sandra et al., 1999) from a simple observation: these errors cluster around certain verb types. More particularly, they typically occur when the application of the morpho-orthographic rules and the phenomenon of word-final devoicing interact to cause verb homophones. Two verb types yield such grammatical homophones. Verbs whose final stem phoneme is /d/, spelled as *d* as the result of the principle of uniformity, yield homophones in the 1st vs. 2nd and 3rd person singular of the present tense (Type 1). See Table 1, where *leid* and *leidt* are spelled in accordance with the morphological principle (stem uniformity and suffix analogy), but are both pronounced as [lɛit], due to devoicing of stem-final /d/. Second, verbs whose stem-final phoneme is not /d/ and with a so-called weak prefix

yield homophones in the 3rd person singular present tense and the past participle (Type 2). For instance, application of the morphological spelling rules to the verb *bedoelen* ('to mean') yields the homophone pair *bedoelt-bedoeld* ('means'-'(has) meant', see Table 1). Final devoicing yields the pronunciation [bədult] for the form ending in <d> as well, thus masking the difference caused by the word forms' morpho-orthographic spelling. As a result, typical verb spelling errors in Dutch are homophone substitutions. At the same time, such errors also occur in verb forms with partial homophones in their inflectional paradigm. For instance, past participles ending in <d>, like *gedroomd* ('(has) dreamed') are partially homophonous with the 3rd person singular present tense (*droomt,* 'dreams'). Spelling errors like *gedroomt,* i.e., a non-existing form, are common (cf. Surkyn et al., 2021, for the cognitive factors behind them). Hence, the first take-home message is that the confrontation between transparent morpho-orthographic rules and homophony creates one of the 'weak spots' in (Dutch) spelling that Bar-On and Kuperman (2019) refer to.

4.2.2 Three Cognitive Factors behind the Dutch Verb Tragedy

Working memory

Our research revealed that these errors are driven by the operation of three factors (Frisson & Sandra, 2002; Sandra et al., 2004, 2010; Sandra et al., 1999; Sandra & Van Abbenyen, 2009; Surkyn et al., 2020, 2021; Verhaert et al., 2016; see also Sandra, 2007, 2018, 2020 for a discussion of the data from a theoretical perspective, but with different accents than here). The first factor is working memory. Sandra et al. (1999) found that an increase in the distance between the verb homophone and the word that determines its spelling (the subject for the present tense, the auxiliary verb for the past participle) increases the error rates. This demonstrates that the application of these rules consumes working memory resources. The determinant of the suffix spelling must be kept in working memory until the verb form is spelled. However, it can be lost by then, as this information fades away as time goes by. Another possibility is that retrieval of the memory trace is still possible but cannot be accomplished in time. Schmitz et al. (2018), in a study of spelling errors on Dutch verb homophones on Twitter, reported the same adjacency effect. They also found that more errors were made in the evening and at night than in the morning and that the error risk was larger in longer tweets. The latter effects suggest that whatever causes a reduction of attentional resources in working memory leads to an increase in the errors.

Homophone dominance

The second factor is the homophones' orthographic representations in long-term memory. In several experiments we found that most errors are due to intrusions of the higher-frequency homophone. This suggests that, when the rules cannot be used (in time) to determine the verb form's spelling in working memory, the

higher-frequency homophone is automatically retrieved and spelled. To be sure, this will more often result in the correct spelling than an error, as the higher-frequency spelling is most likely to occur. However, even though 'choosing' the higher-frequency form is a good 'strategy' in a probabilistic domain,[11] it is a bad one in a rule-governed domain, as rules are deterministic by nature. Hence, spellers' tendency to spell the higher-frequency form when running into working memory problems often causes a spelling error. The result is that most errors are made on the lower-frequency homophone. For instance, for verbs with homophones like *leid* and *leidt* (Type 1), most errors are made in the 1st person when the 3rd person spelling is more frequent but in the 3rd person when the spelling of the 1st person has a higher frequency. The same effect occurs for the Type 2 verbs mentioned earlier (e.g., Sandra et al., 1999, see also Assink, 1985). We have dubbed this phenomenon the effect of homophone dominance.

This effect was also demonstrated in cases where the spelling rule for an inflected form is poorly known, as in the case of the Dutch informal imperative. If spellers do not know the spelling rule, there is no point in mobilizing working memory resources. This, then, should be an ideal scenario for the retrieval process and the frequency-dependent intrusion errors that it gives rise to. This was confirmed by the results reported by Sandra (2010). The spelling of the informal imperative for Type 1 verbs matches the stem spelling, i.e., ends in <d>. Significantly more intrusion errors occurred when the verb's <dt> homophone was the higher-frequency one (e.g., *wordt,* 'becomes', for *worden,* 'to become').

The studies mentioned above were controlled experiments, requiring participants to write from dictation under time-pressure. Thus, we taxed their working memory to magnify the effect and to thus guarantee sufficient statistical power for detecting error patterns. However, we recently found evidence for the same effect in an anonymized chat corpus, consisting of more than 400,000 posts voluntarily provided to us by teenagers between 14 and 20 years old (> 2.3 million tokens, comprising 5804 and 2441 verb forms of Types 1 and 2, respectively; Surkyn et al., 2020). This finding is important, as it confirms the experimental data in an everyday writing situation. Interestingly, even though chatters are known not to observe the traditional spelling rules (e.g., *kisssssss, yolo, w8*[12]), they displayed the same effect of homophone dominance for Type 1 and Type 2 verbs. Schmitz et al. (2018) reported the same findings for both verb types in an analysis of Twitter data. Our chat corpus contained many more homophone intrusions (about 30% for both verb types) in comparison to Schmitz et al.'s corpus of tweets (about 7%), a difference that is probably related to differences between these two types of social media. More importantly, despite this difference in overall error rates, the error patterns (effect of homophone frequency) were the same. These error patterns were targeted in Surkyn et al.'s study. An important finding was that the error pattern was not

[11] Obviously, the preference for the higher-frequency form is not a conscious choice but is due to the higher accessibility of its orthographic representation in long-term memory.

[12] yolo: you only live once; w8: wait

affected by significant differences in the error rates of socially defined subgroups (Gender: male/female, Age: younger/older, and Educational Track: general, technical, vocational): the effect of homophone dominance did not interact with the effects of these social variables. This is exactly what is predicted by the model that we derived from all our experimental studies. Working memory acts as an error *trigger,* whereas the frequencies of the orthographic representations of the verb homophones in long-term memory determine the error *pattern.* Social variables, which may affect the knowledge of the spelling rules, their speed of application, or the spelling attitude, may have an impact on the attentional process in working memory – and, hence, affect the number of errors (which was indeed the case for all three social variables) – but not on the automatic process of orthographic retrieval. Hence, the error pattern should be the same.

Sandra and Van Abbenyen (2009) found that the frequency of the higher-frequency homophones was not only determined by the frequency of the verb homophones themselves. The frequency of a homographic homophone of one of the verb spellings co-determined the error-risk. For instance, 1st person singular verb forms like *bloed* ('blood') and *dood* ('kill'), whose spelling (and pronunciation) is also used for a (semantically related) noun and/or adjective, contributed to the orthographic frequency of the <d> spelling, as evidenced by the error pattern.

Interestingly, in a recent study (Surkyn et al., 2021), we found that the effect of frequency is not restricted to the level of full inflected forms but extends to the sub-lexical level. In our study of spelling errors of partially homophonous past participles (e.g., *gedroomt* instead of *gedroomd*, 'dreamed', due to the partial homophony with *droomt,* 'dreams') we found significant effects of two additional frequency variables. First, chatters made fewer errors when the correct <d> spelling received more support from other forms in the verb's inflectional paradigm (e.g., *droomde,* 'dreamed). This measure of <d> support was calculated as the (logarithm of the) ratio of the token frequency of the <d> spelling over the token frequency of the <t> spelling. Second, they made fewer errors when the correct <d> spelling received more support from the token frequency of the bigram straddling the morphological boundary, taking all inflected forms of all verbs into account. This measure of bigram support was calculated as (the logarithm of) the ratio of the token frequency of the bigram ending in <d> (e.g., *md* in *gedroomd, 'dreamed'*) over the token frequency of the bigram ending in *t* (e.g., *mt* in *komt, 'comes'*). These two factors accounted for independent variance in the error data. Importantly, both measures pick up aspects of the morpho-orthographic spelling of Dutch verb forms, as they both reflect the analogy principle in the spelling of Dutch verb forms: at the intra-paradigmatic level and at the inter-paradigmatic level. Interestingly, the morpho-orthographic cause of this relationship between the error risk and the effect of bigram support is highlighted by the fact that the effect is only found when the count is made over verbs but disappears when including nouns and adjectives as well. In keeping with these findings, Sandra (2010) and Sandra and Van Abbenyen (2009) also found evidence of sub-lexical homophonous patterns straddling the morpheme boundary, more particularly, in past tense forms.

Low-frequency spelling rules

The third factor that is responsible for intrusions of verb homophones is the frequency with which the morpho-orthographic rules must be applied. Perhaps, the importance of this factor should have been emphasized more strongly in earlier publications (but see Sandra & Van Abbenyen, 2009). Indeed, this factor amounts to another important frequency effect: rules that are not needed often enough in spelling cannot become automatic in their application. At this point, the distinction between the principle of uniformity and the principle of analogy is crucial. The former does not cause spelling problems in young adolescents. Even in our analysis of verb spelling errors in their chats, we (Surkyn et al., 2020) observed almost no phonetic spellings like *vint* for *vind* (1st person singular from *vinden,* 'to find'), pronounced [vInt]. We attested only 14 such errors on a total of 1665 homophone intrusions with a stem-final <d> (0.84%). If even adolescents do not make such errors on a medium where observing spelling rules is not considered important and medium-specific spellings are often used, we can safely conclude that the principle of uniformity does not pose a challenge. The fact that we observed this in a dataset where about 1 verb homophone out of 3 was misspelled reinforces this conclusion. As mentioned earlier, this is likely due to this principle being applied across the board in Dutch spelling. The phoneme /d/, spelled <d>, must often be recovered from a final [t]-sound, by means of a morphological relationship, in order to spell the singular of many nouns (*land-landen,* 'country'-'countries'), the uninflected form of many adjectives (*goed-goede,* 'good'), and all regularly inflected forms of verbs with a stem-final <d> (*leid,* 'lead', *leidt,* 'leads', 'led', *leidde,* '(has) led', *geleid*). The frequent application of this principle causes a form of overtrained behavior, which enables automatic application.

In contrast, the principle of analogy is used very infrequently. Estimates based on CELEX (Baayen et al., 1995)[13] show that verb homophones must be spelled very seldom. For instance, homophones from verbs with a stem-final <d> have an occurrence frequency of less than 10%, both type-wise and token-wise (Sandra & Van Abbenyen, 2009). Especially the low type frequency (5%) means that the paucity of homophones from different verbs in everyday writing situations makes it almost impossible for spellers to form a solid representation of the analogical spelling rules, whether it be in the form of an abstract rule or a population of sufficiently frequent exemplars. Cognitive scientists are not surprised when a principle whose use is so limited does not lead to automatic application. This explains why an attentional process in working memory is required, which fails whenever this memory system runs out of resources. In our speeded dictation tasks, this was due to a combination of the time-pressure and (in one condition in Sandra et al., 1999) the distance between the verb homophone and the word that determines its suffix spelling (the latter always preceded the homophone). However, this finding can probably be

[13] CELEX was used instead of SUBTLEX-NL (Keuleers et al., 2010) because only the former contains information on a verb form's grammatical function (e.g., 1st person singular present tense).

generalized to any factor that depletes working memory resources. In everyday writing situations, this can obviously also be time-pressure and the distance between the two grammatically related words, but it can also be an overload of working memory as the result of the complexity of the writing process itself. Indeed, the latter involves a focus on the meaning of the text, its cohesion and structure, the formulation of the sentence, lexical choices, and the spelling of the selected words.

The interplay of three cognitive factors

The conclusion emanating from this research is that morpho-orthographic rules can cause considerable problems. The frequency with which a morpho-orthographic rule is needed for spelling is a major determinant of spelling success. This fits in with our knowledge about the importance of frequency in cognition, both in language processing and in other domains. This explains why the principle of uniformity, i.e., stem constancy, causes no spelling problems in Dutch. It leaves its stamp on so many word spellings – across word classes, i.e., nouns, adjectives, and verbs – that it can be swiftly applied to the morpho-orthographic spelling of the stem (i.e., <d> in stem-final position, as in *leid* or the inflected form *leidt*). Very strong evidence in favor of this is the finding that even teenagers do not spell such stems phonetically in their chat messages (see above).

In contrast, the principle of analogy, which is almost exclusively applied in the spelling of verb homophones, is needed so infrequently that even experienced spellers do not seem able to automate it. The resulting need to allocate the analogical rules attentional resources creates an error risk, especially when some factor taxes the speller's working memory resources: time-pressure, words separating the subject and the verb homophone, a division of attention over several aspects of writing, etc. When this happens, spellers apparently do what they are used to doing in other domains: they have recourse to the highest-frequency event. This is another frequency effect, but this time at the item-specific level. As mentioned earlier, spellers' preference for the higher-frequency form is not a conscious decision. As in many other domains, their behavior is unconsciously steered by what they have encountered or done most often. In the absence of the ability to apply the rule in time, they unknowingly rely on the distribution of the homophone's orthographic forms they have been exposed to, a result of implicit statistical learning.

In short, two frequency factors dominate the scene: rule frequency and item frequency. In the absence of sufficiently frequent rule application, even experienced spellers fail to apply the rule when their working memory resources are depleted, and then spell the most frequent item. The balance between rule frequency and item frequency determines spelling success.

4.2.3 Converging Evidence from Different Alphabetic Orthographies

The French connection

Very similar findings were reported in French by Largy et al. (1996, see also Fayol et al., 1994). Despite some differences between their study and ours, there is a

remarkably strong convergence between the observations in Dutch and French. Largy et al.'s participants had to write down sentences that had just been read and simultaneously recall a list of five words or count the number of clicks they heard during dictation. The use of a concurrent task (dual-task paradigm) was meant to create cognitive overload in working memory. In their critical sentences, like *Les chimistes prennent des liquides. Ils les filtrent* ('The chemiste take liquids. They filter them.'), Largy et al. found homophone intrusions like *filtres,* which is the plural form of the noun *filtre* rather than the plural verb form (*filtrent,* 'filter'). As in our experiments in Dutch, the errors were characterized by a frequency effect: more intrusions were made when the noun homophone was more frequent than the verb homophone. Note that the intruder and the target word belonged to different grammatical classes.[14] Apparently, when working memory is overloaded, the 'pressure' from the higher-frequency homophone can be so strong that the homophonic intruder is not only an error from the perspective of the grammatical context (like the errors we studied) but also from the perspective of the target word's grammatical category.

Although the commonalities are striking, there are a few differences between our studies and Largy et al.'s study. An important difference concerns the occurrence of the errors in everyday writing situations. Whereas the homophone intrusions in Dutch are a persistent and notorious error, Largy et al. remark that their homophone intrusions seldom occur outside the laboratory. They must be experimentally induced by creating a considerable overload in participants' working memory. A second difference concerns the trigger of the errors. Whereas the errors in Dutch and French were both triggered by working memory overload, those in French were likely induced by the presence of another grammatical homophone in the sentence. The French direct object *les* ('them') before the critical verb is homographic and homophonic with the French plural form of the definite article. This may have created a noun bias that facilitated the retrieval of a higher-frequency noun homophone. Finally, Largy et al. studied verb-noun homophones, whereas we studied homophones of the same verb. Differences in the occurrence frequencies of the French and Dutch errors in ordinary writing situations – very rare (Largy et al.) vs. relatively common – are related to all these discrepancies.

The commonalities between the studies are more important. They emphasize that the Dutch findings are not restricted to the specific orthography of one language. Rather, the converging cross-linguistic data show that descriptively simple morpho-orthographic spelling rules can cause spelling errors when working memory is sufficiently overloaded. This opens the door for the intrusion of the wrong grammatical homophone, which is most likely to be the higher-frequency spelling.

Our findings that homophone intrusions in Dutch are not restricted to word forms but extend to the sublexical level (Sandra & Van Abbenyen, 2009; Surkyn et al., 2021) are also supported by research in French. Pacton and Fayol (2003) found that

[14] As mentioned earlier, the effect of homophone dominance in Dutch is co-determined by the frequency of homographic noun and adjective homophones (Sandra & Van Abbenyen, 2009).

children who were tested in Grade 3 (~8 years old) and Grade 5 (~11 years old) are biased to spell the homophonic word-final phoneme /ã/ as *ent,* which is the most frequent spelling pattern for this pronunciation. As this homophone is written as *ent* in adverbs, as part of the adverbial suffix *-ment* (e.g., *calmement*, 'calm') and *ant* in past participles (*regardant,* 'watching'), most homophone intrusions occur in past participles. Importantly, Grade 3 children made these errors on word forms with a high and low frequency, whereas Grade 5 children continued to make them on low-frequency forms. The latter finding is important: if Grade 5 children's spelling were only determined by rule application, form frequency would have no effect.

The findings reported by Pacton et al. (2005) further emphasize the importance of sublexical homophony. The French sound [o] can have different spellings across words (e.g., *piano,* 'piano'; *manteau,* 'coat'; *escargot,* 'snail'), as can the sound [ɛt] (e.g., *planète,* 'planet'; *défaite,* 'defeat'; *assiette,* 'plate'; *conquête,* 'conquest'). Both sound patterns also correspond to the spelling of the masculine and feminine diminutive suffixes, respectively: *-eau* (*éléphanteau,* 'baby elephant') and *-ette* (*fillette,* 'little girl'). The authors found that children in Grades 2, 3, and 5, but also adult spellers, more often chose the latter spelling sequences when the preceding context triggered the diminutive ('a little /vitar/ is a /vitaro/'). However, the probability that they do so was determined by their familiarity with the orthographic pattern straddling the morpheme boundary. For instance, they made more errors when they had to spell pseudowords like [vitafo] and [sorivet] than [vitaro] and [soritet] because the letter patterns *feau* and *vette* occur less often than *reau* and *tette* (see Sandra, 2010; Surkyn et al., 2021, for similar effects of cross-morphemic letter patterns).

Take-home message from the Dutch and French data

The convergence of the Dutch and French data, shows that children, but also adults, do not simply apply even simple morpho-graphic rules but are sensitive to the frequency of all kinds of orthographic patterns in words: at the level of full forms but also at the sublexical level, even when the pattern crosses a morpheme boundary. Clearly, the written input results in the emergence of varying associative strengths between a particular pronunciation and its possible spellings, apparently with complete disregard for grammatical information. These associations and their occurrence frequency are stored, whether it be in the form of multiple mappings between a pronunciation and several spellings or a cluster of exemplars. Obviously, this also occurs when such storage can have counterproductive effects on spelling. Upon reflection, this is what one should expect, as our brain cannot 'know' during learning whether an association between pronunciation and spelling will turn out to be helpful or harmful. Instead, it diligently stores what it is exposed to and excels in keeping track of recurring associations, i.e., patterns in the input. Such statistical learning naturally yields orthographic representations of the inflected forms of regular grammatical homophones, even though these are superfluous from a linguistic perspective, and representations of regularities (in contrast to rules) in the mappings between a single pronunciation and multiple possible spellings. Spellers'

frequency-sensitive access to this stored information is the source of many grammatical homophone intrusions.

More cross-linguistic convergence

Not unexpectedly, problems with morpho-orthographic spelling rules also turn up in other languages (see Bryant et al., 1999). In English, it takes children time to learn that the regular past tense ending is always spelled as *ed,* despite variations in its pronunciation (*stopped, killed, started*). However, English regular past tenses do not give rise to grammatical homophones. In contrast, the English genitive does (e.g., *the boys have a book* vs. *the boys' book* vs. *the boy's book,*). Bryant et al. (1997) observed that British children between the ages of 9 and 11 still have problems spelling the silent apostrophe. Interestingly, this spelling problem seems quite persistent as well, as evidenced by the authors' remark that "[t]he apostrophe has become a kind of cultural shibboleth: educated people, it is typically assumed, use it well and uneducated people do not. In some circumstances, such as applying for a job or even writing an examination essay, the misuse of apostrophes can be a serious disadvantage." This could be a description of the stigma that rests on spelling errors on regular Dutch verb homophones.

Similar problems occur in Greek. Chliounaki and Bryant (2003) point out that children encounter the same problems with morpho-orthographic spelling rules in a shallow orthography (Greek) as in a deep one (English). This indicates that learning to apply these rules represents a serious cognitive hurdle in many orthographies (but see below for the impact of typological differences). The title of their paper captures this important message: "Different morphemes, same spelling problems: Cross-linguistic developmental studies." Protopapas et al. (2013) arrive at the same conclusion and observe that "difficulties in spelling [Greek] inflectional suffixes, […] persist *through a long period of morphological development,* whereas lexical idiosyncrasies determining word root spellings seem to be mastered more readily" (p. 640, my emphasis).

In Danish, children wrestle with the same problem (Juul & Elbro, 2004). One such problem is reminiscent of the effect of homophone dominance that is so persistent in Dutch. The infinitive and present tense forms of Danish verbs with a stem-final /r/ are grammatical homophones with differently spelled suffixes (*e* vs. *er,* respectively). As in Dutch, this situation causes considerable spelling problems.

4.3 Spelling Errors Affect the Quality of Correct Orthographic Representations

In our accounts of the persistence of spelling errors on regular Dutch verb homophones, we have focused on the role of working memory, rule frequency and relative homophone frequency. Interestingly, recent research suggests the role of yet another frequency factor: the relative frequencies of correct and incorrect spellings

of a word. The findings reported by Rahmanian and Kuperman (2019), in the context of reading experiments, support the idea that being exposed to incorrect homophone spellings (e.g., *inocent, comit, begining*) leads to the storage of these incorrect orthographic representations as well. The authors observed that a larger proportion of incorrect spellings in the input caused longer eye fixation times in sentence reading and longer lexical decision times in isolated word recognition. They interpreted their results in terms of Perfetti's (2007) Lexical Quality Hypothesis: stable lexical representations have consistent associations between their orthographic, phonological, and semantic aspects. Misspellings cause unstable representations because each spelling error reduces the frequency of the correct spelling and because an alternative (homophonic) spelling form is stored. The latter acts as a competitor during word reading (and, quite likely, also word spelling). This view fits in with our above description of the brain as an excellent bookkeeper of the information it is exposed to.

Rahmanian and Kuperman's findings support our studies in Dutch. Before discussing this point, it is important to mention a few differences between their study and ours. As mentioned, these authors did not address spelling performance but the impact of spelling errors on word reading. Moreover, they focused on words whose spelling must be memorized whereas we studied rule-governed (inflected) word forms. Finally, their spelling errors were pseudo-homophones, i.e., spellings that do not exist in English (*inocent*), whereas we focused on spelling errors that were existing homophones. Despite these differences, their findings are directly relevant for an account of our results. Even though the Dutch grammatical homophones for the Type 1 and Type 2 verbs mentioned above are both existing spelling forms, misspellings that are triggered by the higher-frequency homophone disturb the frequency relationship between the two spelling forms and further increase the frequency imbalance.

Interestingly, we found evidence supporting Rahmanian and Kuperman's claim in a recent study on partially homophonic verb forms in chat messages (Surkyn et al., 2021). The verb forms were past participles like *gedroomd* ('dreamed'), which are homophonous with the 3rd person singular of the present tense (*droomt,* 'dreams'). A striking observation was the exceptionally high error rate on the form *gezegd* ('(has) said'), which was misspelled as the non-existing form *gezegt* (containing *zegt,* 'says') in more than 25% of its occurrences (well above the average of 11%). We suggested that the frequent occurrence of this spelling error, for a verb form that itself occurs very frequently in chat messages, causes the development of a competitor orthographic representation. This fits in with Rahmanian and Kuperman's claim that sufficient exposure frequency to an incorrect, non-existent (homophonic) spelling feeds a process of implicit learning of this incorrect spelling, which results in an unstable lexical representation. As Gahl and Plag (2019) remark, my paper on sublexical homophone intrusions also suggested the possibility of incorrect orthographic learning: "It might be argued that the incorrect orthographic representations are also stored in the mental lexicon" (Sandra, 2010, p. 425). Given a chat context, the more a given spelling error is used by other chatters – through a process of accommodation, doubt, or yet another factor – the stronger the competition from this rival spelling will become. It is unclear whether learning incorrect

spellings is a medium-specific phenomenon, affects some spellers more than others, and can be avoided (by some types of spellers) in more formal writing contexts. The issue of the medium-specific learning of incorrect spelling forms is a promising avenue for future research.

Here, too, the above remark applies that the brain is indifferent to distinctions like 'correct' or 'incorrect'. It just stores what it is exposed to. This supports an earlier claim I made in several publications: sufficient exposure to regular verb forms leads to an independent full-form representation, like any other word. From an analytical perspective, such a representation is obviously not needed – on the contrary, it is even the source of spelling errors – but there is no reason why the brain would 'care' about this. It just keeps track of repeated inputs (Sandra, 1994).

Homophone intrusions at the sublexical level (Sandra, 2010) are also compatible with Rahmanian and Kuperman's view. Spelling errors in Dutch like *taste* (for *tastte,* 'touched) or *lachtte* (for *lachte,* 'laughed) involve a homophonic intruder that is a non-existing spelling. Quite possibly, such errors reflect the speller's reliance on the higher-frequency spelling of a sound sequence that crosses the morpheme boundary (see also Pacton et al., 2005). It also fits the many demonstrations in the literature that human cognition can better be described as a probabilistic system, which excels in detecting recurrent (but not necessarily constant) patterns, than as a deterministic system, which is superior in rule induction and application (see also Deacon et al., 2008).

Interestingly, further support for the negative impact of misspellings comes from a possibly unexpected source: the impact of digital writing on formal writing situations. Simoës-Perlant et al. (2018a, b) studied the impact of digital writing on writing in a formal context. They studied the rise of two types of intrusion errors since the advent of social media: those that cannot be mistaken as spelling errors (e.g., *svt* for *souvent,* 'often') and those that can be mistaken for ordinary spelling errors (e.g., *quil* for *qu'il,* 'which he'; *c'est à dire* for *c'est-à-dire*, 'that is to say'). When comparing the errors on the same dictation that secondary education students wrote in 1974 (age ~15 years) and in 2012 (matching the groups on spelling level and taking the test at the same time of the year), they found no impact from social media for the former error type but a considerable impact for the latter error type: 3.42% errors in 1974 vs. 13.75% in 2018. They conclude: "Therefore, if the modifications that can be mistaken for misspellings are frequently used in digital writing, this can damage the content of the orthographic lexicon, at least more than when the user only produces standard writing." (p. 172) Clearly, their description of the disturbing effect of competitor spelling forms with the same pronunciation is the same as Rahmanian and Kuperman's. It also converges with our view that the persistence of spelling errors on Dutch verb homophones is due to sufficient exposure to rivaling orthographic representations. Sufficient exposure automatically yields orthographic representations, which affect spellers' output. Simoës-Perlant et al. (2018a, b) also found that digital writing in instant messaging poses a larger threat for adolescents who are poor spellers and whose orthographic knowledge is still not fully consolidated. Weak orthographic representations are more easily overruled by non-standard spellings that are frequently used in instant messaging. This, too, is compatible with

Perfetti's Lexical Quality theory and the above views on how orthographic representations develop.

4.3.1 Morpho-Orthographic Homophones vs. Lexical Homophones: Shared Error Sources

It is very interesting to see that our findings on Dutch grammatical homophones and the results in other languages that are compatible with it run quite parallel with the findings on spelling errors on lexical homophones. White et al. (2008) reported that English homophones like *beech* and *beach* show an effect of homophone dominance: more errors are made on the lower-frequency spelling (*beech*). Moreover, more errors were made when the preceding sentence contained a prime word that was orthographically related to the incorrect spelling, both for the dominant and subordinate spelling forms. For instance, although more errors were made on *beech* than on *beach,* the errors on both forms increased in sentences like *After presenting her speech on animal rights, Sue went to the beach to relax* and *The teacher was most proud of the beech tree in his garden.* This priming effect did not interact with the frequency of the target homophone. The authors conclude (a) that the effect of frequency dominance reveals the frequency-sensitive storage of lexical homophones, which are accessed by the lexical route, and (b) that the priming effect reveals an interaction between the lexical route and the non-lexical route, in which a spelling pattern is assembled through phoneme-grapheme correspondences (otherwise the higher-frequency homophone would be insensitive to the priming effect).

Our findings that the same effect of homophone dominance occurs for regularly inflected verb forms in Dutch strongly suggests that at least the higher-frequency spelling is stored like a monomorphemic form and thus, from the perspective of lexical access, functions like a lexical homophone. White et al.'s suggestion that multiple access 'routes' are operational at the same time can also explain why errors on Dutch verb homophones include intrusions from the lower-frequency spelling on the higher-frequency one. If frequency dominance were the only error determinant, such intrusions could not be explained, as they are excluded both by the morpho-orthographic rule that is appropriate in the grammatical context and by a frequency account. Possibly, these atypical intrusions are due to preceding words that prime the lower-frequency spelling. Another explanation might be that a frequency-sensitive, probabilistic system is subject to activity in a phoneme-grapheme assembly route, and, hence, does not always 'select' the higher-frequency spelling pattern.

Tsai et al. (2011) found that young Taiwanese children's tendency to make spelling errors on lexical homophones in Chinese correlates with measures of their attentional skills. Their suggestion that the errors are a signature of problems with attention, as evidenced by children with attention deficits, is at least compatible with our claim that errors on grammatical homophones in experienced spellers reflect a depletion of the attentional resources in working memory. It seems that homophones, both lexical and grammatical ones, require specific attention. This attentional component might not be needed for grammatical homophones if the

morpho-orthographic rules yielding them could become automatized, due to a sufficiently large number of target types and tokens. This is how we explained the virtually non-existent spelling errors with respect to the principle of uniformity (i.e., spelling invariance of the stem). Research in a morphologically rich language that makes frequent use of morpho-orthographic rules might shed light on this question. There is at least one study that has focused on this issue.

In a cross-linguistic comparison between the spelling abilities of Hebrew and Flemish (i.e., Dutch-speaking) children in Grades 1–6 (~6 to ~12 years old), Gillis and Ravid (2006) found large differences between the two language groups. Whereas the Hebrew children rapidly used morphological cues to spell words correctly, the Flemish group experienced many problems with words whose spelling was only recoverable when using morphological cues (due to stem uniformity and/or analogy at the suffix level). Even in Grade 6, many errors were still made, although the results indicated a growing awareness of the Dutch morphological principle. The authors attribute this discrepancy between the languages to differences between the rich morphological structure of spoken Hebrew words (Ravid, 2012) versus the sparse inflectional morphology in spoken Dutch. Consequently, unlike Dutch-speaking children, young Hebrew children soon learn to attend to a word's morphological structure. Although this explanation pertains to differences between the spoken languages, the authors suggest that this typological difference has implications for spelling as well. Quite likely, a morphologically rich language soon causes a high degree of morphological awareness, which in turn helps children attend to the morpho-orthographic buildup of words. Gillis and Ravid's view is compatible with our claim that Dutch verb homophones yield so many errors because the distribution of verb forms in Dutch hardly requires the application of the analogical principle, as most inflectional suffixes can be spelled 'by ear'. This stands in contrast to the need to frequently appeal to the principle of uniformity, which probably explains why errors are seldom made in the stem part of inflected verb forms, even in adolescent's chat messages. This cross-linguistic comparison supports the third factor that is responsible for the persistence of verb spelling errors in Dutch: morpho-orthographic rule-frequency, or, perhaps more generally, the frequency with which the morphological spelling principle is applied. It also suggests that the importance of morphological structure in the spoken language makes it easier for children to learn and apply morpho-orthographic spelling rules. This underscores the importance of cross-linguistic studies, both in typologically similar languages with similar orthographies (Dutch, English, Danish, French) and in typologically different languages (Hebrew, Greek).

4.3.2 Early Morphological Awareness vs. Errors on Regular Verb Homophones: An Inconsistency?

Our findings that the effect of homophone dominance in Dutch, even in experienced spellers, is notoriously persistent and is apparently immune to all types of instructional methods (rule-based, analogy-based, algorithmic-based) – as even

experienced spellers make the errors – seems to fly in the face of the many studies that have demonstrated that (a) there is a positive correlation between morphological awareness and success in the application of morpho-orthographic rules, (b) a beneficial effect of morphological intervention has been shown, and (c) in the domain of inflectional morphology, grammatical awareness affects the error rate on verb homophones. It also seems at odds with demonstrations of an early morphological awareness in children (Treiman et al., 1994; Cassar & Treiman, 1996, Sénéchal, 2000; Sénéchal et al., 2006).

To begin with, there is abundant evidence that morpho-orthographic rules for spelling inflected forms create a considerable hurdle in many languages. Learning these rules is considerably more difficult, and, hence, slower than learning the mappings between phonemes and graphemes. Insight into words' morphological structure seems to be a prerequisite to guarantee effective rule application. First, this is indicated by the frequent finding (in several languages) that children's morphological awareness correlates with their ability to apply morpho-orthographic rules. Children who can detect the morphological relationship between words (for instance, to derive a silent letter) or who perform well on tasks targeting morphological awareness (e.g., analogy task) have been shown to spell morphologically complex words better than children who score worse in this respect (Sénéchal, 2000; Sénéchal et al., 2006). Note that this not only applies to the ability to analyze a word's internal morphological structure (which may help, for instance, for silent letters or diminutives in French, i.e., words whose spelling does not depend on another word in the sentence) but also pertains to the ability to determine which word determines the spelling of an inflectional suffix, i.e., grammatical awareness. Recent work by Chamalaun et al. (2021) has revealed that grammatical awareness in 11- to 18-year-old secondary school students is a reliable predictor of the spelling correctness of Dutch verb homophones, i.e., the same type of homophones that my collaborators and I studied. Adolescents who were able to identify the grammatical function of a verb form (e.g., present tense, past participle) were significantly more successful in correctly spelling this verb form. Second, the observation that children's performance on these word forms improves as they grow older and that the amount of this improvement is predicted by their degree of grammatical awareness indicates that the ability to spell these word forms correctly requires insight into grammatical concepts. This insight underlies the invariant spelling of the stem (e.g., Deacon & Bryant, 2006b) and the analogical spelling of the suffix or another grammatical marker like the apostrophe in the English possessive (for the English past tense *-ed:* Nunes et al., 1997a, b; for the English possessive: Bryant et al., 1997; for Danish; Juul & Elbro, 2004; for Greek: Chliounaki & Bryant, 2003). Third, intervention studies that explicitly attempt to improve children's morphological awareness have reported beneficial effects. Children with spelling problems obtained higher spelling scores when an intervention program targeted their morphological awareness and explicitly focused on the morpho-orthographic rules (Kirk & Gillon, 2009). Good et al. (2015) showed that this was even the case for children with language impairment, who generalized this morphological awareness to novel words.

In sum, there can be no doubt that the application of morpho-orthographic rules depends on a sufficiently developed morphological awareness.

Still our findings for Dutch verb homophones prove that, when certain conditions are met, analytically simple morpho-orthographic spelling rules may cause persistent spelling problems, even for highly educated people who write texts for professional purposes. Hence, they survive, despite a high level of morphological awareness. Even though the cross-linguistic comparison in Sect. 4.2.3 indicates that grammatical homophones are a problem in other languages as well, a particularly strong manifestation of this problem seems to be found in Dutch. Why?

A first observation is that children's early sensitivity to morphological structure concerns a budding awareness of morphological relatedness at the level of the stem, as shown by Treiman and colleagues and Sénéchal and colleagues. As has been reiterated in this contribution, the morpho-orthographic spelling of Dutch stems, which is governed by the principle of uniformity, does not cause any problems in Dutch either, once it has been learnt. In contrast, spelling inflectional suffixes cause many problems. The 'Dutch problem' is a problem with these suffixes. However, even though such suffixes come with a difficult learning curve in several languages (for instance, English *-ed*, see the work by Nunes and colleagues), the initial difficulties progressively disappear and seldom cause the lasting problem that verb homophones cause in Dutch. I could find only one problem in the literature that echoes the spelling problems for Dutch homophones, or at least the attitude of extreme intolerance that is associated with them: the English apostrophe marker for the possessive (see the earlier citation from Bryant et al., 1997).

Putting everything into perspective, the major culprit is rule frequency. Strong support for this analysis comes from the dissociation in Dutch verb homophones between almost error-free performance with respect to the principle of uniformity and the many problems with respect to the principle of analogy. This supports the idea that only morpho-orthographic rules with a high frequency, i.e., which are applied to a wide variety of words with a high token frequency, can eventually be applied automatically. This account can also explain why even young children who learn to read and spell in Hebrew can readily make use of different types of morphological cues (Gillis & Ravid, 2006). They use morphological spelling rules so frequently that they are quickly able to apply them automatically. However, when rule application cannot become automatic, it requires attentional resources in working memory. Any factor that puts a limit on these resources creates a risk that the rule cannot be applied in time. When this happens, another frequency factor enters the scene and ultimately determines the spelling output: item frequency, more particularly, the homophones' frequencies. Not knowing the answer, spellers (unconsciously) have recourse to the higher-frequency spelling. This effect of homophone dominance has been found at the level of the full form and at the sublexical level.

In short, the answer to the question whether the Dutch findings are not at odds with the rest of the literature is negative. Experienced Dutch spellers apply the morphological principle of uniformity automatically. Moreover, as in other languages,

their degree of grammatical awareness is a predictor of their spelling errors on verb homophones (Chamalaun et al., 2021). The factor that sets the Dutch verb spelling errors apart is rule frequency. Frequent rules have a steep learning curve but eventually lead to automatic rule application. Highly infrequent rules are also difficult to learn but never lead to automatic rule application. In Dutch spelling, the former is the case for the principle of uniformity, the latter for the principle of analogy. This explains why Dutch verb homophones occupy a special place in the landscape of morpho-orthographic spelling errors.

4.3.3 Persistent Errors Do Not Imply One Type of Speller

The persistence of verb homophone intrusions in Dutch provides a window on the cognitive infrastructure that underpins the spelling of Dutch verb homophones. However, it should not make us blind to differences between spellers. Roughly speaking, one can make a clear distinction between two types of spellers: (a) those who have sufficient morphological and grammatical awareness to be able to apply the morpho-orthographic rules but, due to the three factors mentioned above, from time to time make an error on a grammatical homophone and (b) those who have insufficient morphological and/or grammatical awareness to apply these spelling rules. Obviously, in the absence of rule knowledge, working memory need not be overloaded to make a spelling error, as it cannot even start the computational process that implements the spelling rule. Consequently, for spellers of the latter type, the door will always be open to intrusions of the higher-frequency homophone, which means they will make many intrusion errors on verb homophones. These two types of spellers should indeed leave different fingerprints on the error rates. Whereas the former will make an occasional error, the latter will make many errors (even in a single text). Chamalaun et al.'s (2021) finding that spellers' grammatical awareness predicts their spelling success on Dutch verb homophones supports such a distinction. Homophone frequency will shape different error patterns for these two types of spellers. Whereas low rule frequency is the problem of experienced spellers, lack of rule knowledge is the problem of weak spellers.[15]

For practical purposes (e.g., in a school context), it is important to distinguish between these two types of spellers, as the former are the victim of their cognitive infrastructure, so to speak, whereas the latter do not possess the required rule knowledge. From the perspective of this analysis, there is no reason to point the finger to experienced spellers who occasionally fall into an inevitable trap, whereas such a reaction is more understandable when spellers make many such errors. One might want to distinguish a third group: spellers who know the rules but who are sometimes too indifferent to apply them. The prediction would be that such spellers will

[15] Note that the distinction between two types of spellers is a simplification. It would be better to speak of a continuum between good and poor spellers, as some spellers will have a better mastery of the rules (or apply them more swiftly) than others.

make many errors as well, as the result of a negative spelling attitude. Obviously, spelling, like any other form of cognitive behavior, is not only affected by knowledge but by motivation as well.

5 Conclusions

The study of spelling errors on Dutch verb homophones can teach us a lot about human cognition. It shows that morpho-orthographic rules involving inflectional endings whose spelling (a) is determined by another word in the sentence and (b) causes verb homophones for two verb types are difficult to master, like many morpho-orthographic rules in other languages. However, the errors also show that low rule frequency makes it virtually impossible to automatically apply the rules. Hence, these rules continue to require attentional resources in working memory, and when the latter is overloaded, there is a strong tendency to write the higher-frequency homophone spelling. This explains why spellers with a good knowledge of the spelling rules and a willingness to avoid errors occasionally make errors on the lower-frequency form or, as in the case of partially homophonic past participles, rely on intra-paradigmatic and inter-paradigmatic support. It also explains why spellers with a poor knowledge of the spelling rules make many errors, as they can only cling to the higher-frequency spelling. The story of Dutch verb homophones demonstrates why descriptively simple and analytically clear spelling rules, which only involve the concatenation of a stem and a suffix, remain a spelling hurdle for the best spellers and a huge problem for weak spellers. It can safely be predicted that, given the strong impact of rule frequency and item-frequency, in combination with our working memory limitations, these spelling errors are here to stay.

It is important that this conclusion is based on error data on different verb types, which have been studied in both laboratory experiments (speeded dictation) and in analyses of a large chat corpus. It is also important that this conclusion applies to homophone intrusions at both the lexical and sublexical levels.

Understanding the cognitive origin of persistent spelling errors is not only important for spelling researchers. It also has practical consequences. Certainly, it is relevant for teachers, who are confronted with these (highly stigmatized) errors daily. For instance, the above distinction between several types of spellers is relevant for them. This type of research is also relevant at a larger societal level. Perhaps, insight into the almost unavoidable error trap that is created by the concerted action of three cognitive factors may lead to less stigmatization of the occasional error on Dutch verb homophones and to a distinction between such occasional errors and the frequent errors made by weak spellers. The irony is that many of those who condemn an occasional error on verb homophones sin against their own belief, as no one can claim to be immune to the errors.

Obviously, this does not mean that teachers should not focus on morpho-orthographic spelling rules. On the contrary, children and teenagers will not acquire

these rules by means of implicit learning, as errors persist even in the face of considerable teaching efforts. Moreover, implicit learning results in the induction of statistical regularities, which may be helpful on some occasions but act as a jammer on others. High-frequency orthographic patterns are, by definition, probable but nonetheless sometimes grammatically incorrect. Such implicit learning cannot, of course, be 'switched off', as our brain deals in the business of statistical learning all the time. However, as has been shown in many studies, explicit training of morphological and grammatical awareness (see above) is the only way to maximally protect spellers against verb homophone errors that are always lurking around the corner. They will never be able to wipe them out completely, but there is no other way around the problem.

Of course, there is an alternative when a spelling problem is so persistent and is so heavily stigmatized. Spelling rules that are so demonstrably ill fit to the cognitive infrastructure of the user should perhaps be replaced by rules that can be automatically applied. In the case of Dutch, this might amount to the preservation of the principle of uniformity and the rejection of the principle of analogy. Such a change would not lead to a breakdown in adolescents' ability to reflect on language structure. The preservation of the principle of uniformity would help the development of morphological awareness just as much as it does now. Obviously, the quantitative implications of such (or any other) proposal, i.e., the number of affected spelling forms, should be studied first. It remains to be seen whether insights in psycholinguistics carry sufficient weight to have such a far-reaching practical impact. In the case of Dutch verb homophones, installing new rules that can be applied automatically by most users would leave more time for teachers to focus on the essence of writing instruction: the delivery of well-written, coherent, and well-structured texts.

References

Anthony, J. L., & Francis, D. J. (2005). Development of phonological awareness. *Current Directions in Psychological Science, 14*(*5*), 255–259. https://doi.org/10.1111/j.0963-7214.2005.00376

Assink, E. M. H. (1985). Assessing spelling strategies for the orthography of Dutch verbs. *British Journal of Psychology, 76*(*3*), 228–235. https://doi.org/10.1111/j.2044-8295.1985.tb01958

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995).*The CELEX Lexical Database (CD-ROM).* Linguistic Data Consortium, University of Pennsylvania, Philadelphia.

Baayen, R. H., Dijkstra, T., & Schreuder, R. (1997). Singulars and plurals in Dutch: Evidence for a parallel dual-route model. *Journal of Memory and Language, 37*(*1*), 94–117. https://doi.org/10.1006/jmla.1997.2509

Bar-On, A., & Kuperman, V. (2019). Spelling errors respect morphology: a corpus study of Hebrew orthography. *Reading and Writing, 32*(*5*), 1107–1128. https://doi.org/10.1007/s11145-018-9902-1

Bourassa, D. C., & Treiman, R. (2008). Morphological constancy in spelling: A comparison of children with dyslexia and typically developing children. *Dyslexia, 14*(*3*), 155–169. https://doi.org/10.1002/dys.368

Britannica, The Editors of Encyclopaedia. "Learn how the system of writing evolved". *Encyclopedia Britannica*, 29 Apr. 2021., https://www.britannica.com/summary/writing. Accessed 13 May 2021.

Bryant, P., Devine, M., Ledward, A., & Nunes, T. (1997). Spelling with apostrophes and understanding possession. *British Journal of Educational Psychology, 67*(*1*), 91–110. https://doi.org/10.1111/j.2044-8279.1997.tb01229

Bryant, P., Nunes, T., & Aidinis, A. (1999). Different morphemes, same spelling problems: Cross-linguistic developmental studies. In M. Harris & G. Hatano (Eds.), *Learning to read and write: A cross-linguistic perspective* (pp. 112–133). Cambridge University Press.

Carlisle, J. F. (2010). Effects of instruction in morphological awareness on literacy achievement: An integrative review. *Reading Research Quarterly, 45*(4), 464–487. https://doi.org/10.1598/RRQ.45.4.5

Casalis, S., Deacon, S. H., & Pacton, S. (2011). How specific is the connection between morphological awareness and spelling? A study of French children. *Applied Psycholinguistics, 32*(3), 499–511. https://doi.org/10.1017/S014271641100018X

Chamalaun, R., Bosman, A., & Ernestus, M. (2021). The role of grammar in spelling homophonous regular verbs. *Written Language & Literacy, 24*(1), 38–80.

Chliounaki, K., & Bryant, P. (2003). Choosing the right spelling in Greek: Morphology helps. *Revue française de linguistique appliquée, VIII, 35-45*. https://doi.org/10.3917/rfla.081.0035

Deacon, S. H., & Bryant, P. (2006a). Getting to the root: young writers' sensitivity to the role of root morphemes in the spelling of inflected and derived words. *Journal of child language, 33*(2), 401–417. https://doi.org/10.1017/s0305000906007409

Deacon, S. H., & Bryant, P. (2006b). This turnip's not for turning: Children's morphological awareness and their use of root morphemes in spelling. *British Journal of Development Psychology, 24*(*3*), 567–575. https://doi.org/10.1348/026151005X50834

Deacon, S. H., Conrad, N., & Pacton, S. (2008). A statistical learning perspective on children's learning about graphotactic and morphological regularities in spelling. *Canadian Psychology/ Psychologie canadienne, 49*(2), 118–124. https://doi.org/10.1037/0708-5591.49.2.118

Diependaele, K., Sandra, D., & Grainger, J. (2005). Masked cross-modal morphological priming: Unravelling morpho-orthographic and morpho-semantic influences in early word recognition. *Language and Cognitive Processes, 20*(*1–2*), 114–175. https://doi.org/10.1080/01690960444000197

Ephratt, M. (1997). The psycholinguistic status of the root in Modern Hebrew. *Folia Linguistica, 31*(1-2), 77–103. https://doi.org/10.1515/flin.1997.31.1-2.77

Fayol, M., Largy, P., & Lemaire, P. (1994). Cognitive overload and orthographic errors: When cognitive overload enhances subject–verb agreement errors. A study in French written language. *The Quarterly Journal of Experimental Psychology Section A, 47*(2), 437–464. https://doi.org/10.1080/14640749408401119

Feldman, L. B., Milin, P., Cho, K. W., Del Prado, M., Martín, F., & O'Connor, P. A. (2015). Must analysis of meaning follow analysis of form? A time course analysis. *Frontiers in human neuroscience, 9*, 111. https://doi.org/10.3389/fnhum.2015.00111

Forster, K. I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*(4), 680–698. https://doi.org/10.1037/0278-7393.10.4.680

Fowler, C. A., Napps, S. E., & Feldman, L. (1985). Relations among regular and irregular morphologically related words in the lexicon as revealed by repetition priming. *Memory & Cognition, 13*(3), 241–255. https://doi.org/10.3758/BF03197687

Frisson, S., & Sandra, D. (2002). Homophonic forms of regularly inflected verbs have their own orthographic representations: A developmental perspective on spelling errors. *Brain and language, 81*(1-3), 545–554. https://doi.org/10.1006/brln.2001.2546

Fromkin, V. (1971). The Non-anomalous nature of anomalous utterances. *Language, 47*(1), 27–52.

Gahl, S., & Plag, I. (2019). Spelling errors in English derivational suffixes reflect morphological boundary strength. *The Mental Lexicon, 14*(1), 1–36. https://doi.org/10.1075/ml.19002.gah

Garrett, M. F. (1975). The analysis of sentence production. *Psychology of Learning and Motivation – Advances in Research and Theory, 9*, 133–177. https://doi.org/10.1016/S0079-7421(08)60270-4

Gillis, S., & Ravid, D. (2006). Typological effects on spelling development: A crosslinguistic study of Hebrew and Dutch. *Journal of Child Language, 33*(3), 621–659. https://doi.org/10.1017/S0305000906007434

Good, J. E., Lance, D. M., & Rainey, J. (2015). The effects of morphological awareness training on reading, spelling, and vocabulary skills. *Communication Disorders Quarterly, 36*(3), 142–151. https://doi.org/10.1177/1525740114548917

Juul, H., & Elbro, C. (2004). The links between grammar and spelling: A cognitive hurdle in deep orthographies? *Reading and Writing, 17*(9), 915–942. https://doi.org/10.1007/s11145-004-2109-7

Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. *Behavior Research Methods, 42*(3), 643–650. https://doi.org/10.3758/BRM.42.3.643

Kirk, C., & Gillon, G. T. (2009). Integrated morphological awareness intervention as a tool for improving literacy. *Language, speech, and hearing services in schools, 40*(3), 341–351. https://doi.org/10.1044/0161-1461(2008/08-0009)

Largy, P., Fayol, M., & Lemaire, P. (1996). The homophone effect in written French: The case of verb–noun inflection errors. *Language and Cognitive Processes, 11*(3), 217–255. https://doi.org/10.1080/016909696387178

Leminen, A., Smolka, E., Duñabeitia, J. A., & Pliatsikas, C. (2019). Morphological processing in the brain: The good (inflection), the bad (derivation) and the ugly (compounding). *Cortex, 116*, 4–44. https://doi.org/10.1016/j.cortex.2018.08.016

Longtin, C., & Meunier, F. (2005). Morphological decomposition in early visual word processing. *Journal of Memory and Language, 53(1)*, 26–41. https://doi.org/10.1016/j.jml.2005.02.008

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

Morais, J., & Kolinsky, R. (2002). *Literacy effects on language and cognition*. In L. Bäckman & C. von Hofsten (Eds.), *Psychology at the turn of the millennium, Vol. 1. Cognitive, biological, and health perspectives* (pp. 507–530). Psychology Press/Taylor & Francis.

Morris, J., & Stockall, L. (2012). Early, equivalent ERP masked priming effects for regular and irregular morphology. *Brain and Language, 123*(2), 81–93. https://doi.org/10.1016/j.bandl.2012.07.001

Murrell, G. A., & Morton, J. (1974). Word recognition and morphemic structure. *Journal of Experimental Psychology, 102*(6), 963–968. https://doi.org/10.1037/h0036551

Napps, S. E. (1989). Morphemic relationships in the lexicon: Are they distinct from semantic and formal relationships? *Memory & Cognition, 17*(6), 729–739. https://doi.org/10.3758/BF03202634

Nunes, T., Bryant, P., & Bindman, M. (1997a). Morphological spelling strategies: Developmental stages and processes. *Developmental Psychology, 33*(4), 637–649. https://doi.org/10.1037/0012-1649.33.4.637

Nunes, T., Bryant, P., & Bindman, M. (1997b). Learning to spell regular and irregular verbs. *Reading and Writing, 9(5-6)*, 113–135. https://doi.org/10.1007/978-94-017-3054-9_7

Oliphant, G. W. (1983). Repetition and recency effects in word recognition. *Australian Journal of Psychology, 35*(3), 393–403. https://doi.org/10.1080/00049538308258751

Pacton, S., & Fayol, M. (2003). How do French children use morphosyntactic information when they spell adverbs and present participles? *Scientific Studies of Reading, 7*(3), 273–287. https://doi.org/10.1207/S1532799XSSR0703_5

Pacton, S., Fayol, M., & Perruchet, P. (2005). Children's implicit learning of graphotactic and morphological regularities. *Child development, 76*(2), 324–339. https://doi.org/10.1111/j.1467-8624.2005.00848.x

Pacton, S., Deacon, H., Borchardt, G., Danjon, J., & Fayol, M. (2012). Why should we take graphotactic and morphological regularities into account when examining spelling acquisition? In V. Wise Berninger (Ed.), *Past, present, and future contributions of cognitive writing research to cognitive psychology* (pp. 333–358). Psychology Press.

Pacton, S., Foulin, J.-N., Casalis, S., & Treiman, R. (2013). Children benefit from morphological relatedness when they learn to spell new words. *Frontiers in Psychology, 4*, 696. https://doi.org/10.3389/fpsyg.2013.00696

Pacton, S., Afonso Jaco, A., Nys, M., Foulin, J. N., Treiman, R., & Peereman, R. (2018). Children benefit from morphological relatedness independently of orthographic relatedness when they learn to spell new words. *Journal of Experimental Child Psychology, 171*, 71–83. https://doi.org/10.1016/j.jecp.2018.02.003

Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11*(4), 357–383. https://doi.org/10.1080/10888430701530730

Protopapas, A., Fakou, A., Drakopoulou, S., Skaloumbakas, C., & Mouzaki, A. (2013). What do spelling errors tell us? classification and analysis of errors made by Greek schoolchildren with and without Dyslexia. *Reading and Writing, 26*(5), 615–646. https://doi.org/10.1007/s11145-012-9378-3

Rahmanian, S., & Kuperman, V. (2019). Spelling errors impede recognition of correctly spelled word forms. *Scientific Studies of Reading, 23*(1), 24–36. https://doi.org/10.1080/10888438.2017.1359274

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

Ravid, D. (2001). Learning to spell in Hebrew: Phonological and morphological factors. *Reading and Writing, 14*(5-6), 459–485. https://doi.org/10.1023/A:1011192806656

Ravid, D. (2012). *Spelling morphology: The psycholinguistics of Hebrew spelling*. Springer.

Ravid, D., & Bar-On, A. (2005). Manipulating written Hebrew roots across development: The interface of semantic, phonological and orthographic factors. *Reading and Writing: An Interdisciplinary Journal, 18*(3), 231–256. https://doi.org/10.1007/s11145-005-1802-5

Royle, P., Drury, J. E., Bourguignon, N., & Steinhauer, K. (2012). The temporal dynamics of inflected word recognition: A masked ERP priming study of French verbs. *Neuropsychologia, 50*(14), 3542–3553. https://doi.org/10.1016/j.neuropsychologia.2012.09.007

Sandra, D. (1994). The morphology of the mental lexicon: Internal word structure viewed from a psycholinguistic perspective. *Language and Cognitive Processes, 9*(3), 227–269. https://doi.org/10.1080/01690969408402119

Sandra, D. (2007). Skills and representations in learning to spell and In experienced spellers. In G. Jarema & G. Libben (Eds.), *The Mental Lexicon: Core perspectives* (pp. 207–235). Brill. https://doi.org/10.1163/9780080548692_011

Sandra, D. (2010). Homophone dominance at the whole-word and sub-word levels: Spelling errors suggest full-form storage of regularly inflected verb forms. *Language and Speech, 53*(3), 405–444. https://doi.org/10.1177/0023830910371459

Sandra, D. (2018). The role of morphology in reading and writing. In A. Bar-On & D. Ravid (Eds.), *Handbook of communication disorders* (pp. 477–502). De Gruyter Mouton. https://doi.org/10.1515/9781614514909

Sandra, D. (2020). Morphological units: A theoretical and psycholinguistic Perspective. In *Oxford research encyclopedia of linguistics*. (Published online: 28 February 2020) https://doi.org/10.1093/acrefore/9780199384655.013.541

Sandra, D., & Fayol, M. (2011). Spelling errors with a view on the mental lexicon: Frequency and proximity effects in misspelling homophonous regular verb forms in Dutch and French. In *Morphological structure in language processing* (pp. 485–514). https://doi.org/10.1515/9783110910186.485

Sandra, D., & Van Abbenyen, L. (2009). Frequency and analogical effects in the spelling of full-form and sublexical homophonous patterns by 12 year-old children. *The Mental Lexicon, 4*(2), 239–275. https://doi.org/10.1075/ml.4.2.04san

Sandra, D., Frisson, S., & Daems, F. (1999). Why simple verb forms can be so difficult to spell: The influence of homophone frequency and distance in dutch. *Brain and Language, 68*(1–2), 277–283. https://doi.org/10.1006/brln.1999.2108

Sandra, D., Frisson, S., & Daems, F. (2004). Still errors after all those years…. *Written Language & Literacy Written Language and Literacy, 7*(1), 61–77. https://doi.org/10.1075/wll.7.1.07san

Schmitz, T., Chamalaun, R., & Ernestus, M. (2018). The Dutch verb-spelling paradox in social media. *Linguistics in the Netherlands, 35*, 111–124. https://doi.org/10.1075/avt.00008.sch

Sénéchal, M. (2000). Morphological effects in children's spelling of French words. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 54*(2), 76–86. https://doi.org/10.1037/h0087331

Sénéchal, M., Basque, M. T., & Leclaire, T. (2006). Morphological knowledge as revealed in children's spelling accuracy and reports of spelling strategies. *Journal of Experimental Child Psychology, 95*(4), 231–254. https://doi.org/10.1016/j.jecp.2006.05.003

Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition, 55*(2), 151–218. https://doi.org/10.1016/0010-0277(94)00645-2

Simoës-Perlant, A., Lanchantin, T., Gunnarsson-Largy, C., & Largy, P. (2018a). Instant messaging, digital writing and spelling production quality in French. *Lingvisticæ Investigationes, 41*(2), 161–178. https://doi.org/10.1075/li.00018.sim

Simoës-Perlant, A., Gunnarsson-Largy, C., Lanchantin, T., & Largy, P. (2018b). Instant messaging: a threat for poor spellers whose spelling is in the process of being consolidated. *Interacções, 48*, 21–48.

Sonnenstuhl, I., Eisenbeiss, S., & Clahsen, H. (1999). Morphological priming in the German mental lexicon. *Cognition, 72*(3), 203–236. https://doi.org/10.1016/S0010-0277(99)00033-5

Stanners, R. F., Neiser, J. J., Hernon, W. P., & Hall, R. (1979). Memory representation for morphologically related words. *Journal of Verbal Learning & Verbal Behavior, 18*(4), 399–412. https://doi.org/10.1016/S0022-5371(79)90219-6

Surkyn, H., Vandekerckhove, R., & Sandra, D. (2020). From experiment to real-life data: Social factors determine the rate of spelling errors on rule-governed verb homophones but not the size of the homophone dominance effect. *The Mental Lexicon, 15*(3), 422–463. https://doi.org/10.1075/ml.20006.sur

Surkyn, H., Vandekerckhove, R., & Sandra, D. (2021). The impact of analogical effects and social factors on the spelling of partially homophonous verb forms in informal social media writing. *Written Language and Literacy, 24*(1), 1–37.

Taft, M. (1979). Recognition of affixed words and the word frequency effect. *Memory & Cognition, 7*(4), 263–272. https://doi.org/10.3758/BF03197599

Taft, M., & Forster, K. I. (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

Treiman, R., & Cassar, M. (1996). Effects of morphology on children's spelling of final consonant clusters. *Journal of Experimental Child Psychology, 63*(1), 141–170. https://doi.org/10.1006/jecp.1996.0045

Treiman, R., Cassar, M., & Zukowski, A. (1994). What types of linguistic information do children use in spelling? The Case of Flaps. *Child Development, 65*(5), 1318–1337. https://doi.org/10.1111/j.1467-8624.1994.tb00819.x

Tsai, L. H., Meng, L. F., Hung, L. Y., Chen, H. Y., & Lu, C. P. (2011). Coincidence of homophone spelling errors and attention problems in schoolchildren: A survey study. *Research in Developmental Disabilities, 32*(1), 75–80. https://doi.org/10.1016/j.ridd.2010.08.014

Uppstad, P. H., & Tønnessen, F. E. (2007). The notion of 'phonology' in dyslexia research: Cognitivism–and beyond. *Dyslexia: An International Journal of Research and Practice, 13*(3), 154–174. https://doi.org/10.1002/dys.332

van der Velde, I. (1956). *De tragedie der werkwoordsvormen. Een taalhistorische en taaldidactische studie* [The tragedy of verbs. A linguistic-historical and linguisticdidactic study]. Wolters-Noordhoff.

Verhaert, N., Danckaert, E., & Sandra, D. (2016). The dual role of homophone dominance: Why homophone intrusions on regular verb forms so often go unnoticed. *The Mental Lexicon, 11*, 1–25. https://doi.org/10.1075/ml.11.1.01ver

White, K. K., Abrams, L., Zoller, S. M., & Gibson, S. M. (2008). Why did I right that? Factors that influence the production of homophone substitution errors. *The Quarterly Journal of Experimental Psychology, 61*(7), 977–985. https://doi.org/10.1080/17470210801943978