

Spelling errors in English derivational suffixes reflect morphological boundary strength

A case study

Susanne Gahl and Ingo Plag

University of California at Berkeley | Heinrich-Heine Universität
Düsseldorf

To what extent do speakers decompose morphologically complex words, such as *segmentable*, into their morphological constituents? In this article, we argue that spelling errors in English affixes reflect morphological boundary strength and degrees of segmentability. In support of this argument, we present a case study examining the spelling of the suffixes *-able/-ible*, *-ence/-ance*, and *-ment* in an online resource (Tweets), in forms such as <available>, <invisible>, <eloquence>, and <bettermint>. Based on previous research on morphological productivity and boundary strength (Hay, 2002; Hay & Baayen, 2002, 2005), we hypothesized that morphological segmentability should affect the choice between <able> vs. <ible>, <ance> vs. <ence>, and <ment> vs. <-mint>. An analysis of roughly 23,000 non-standard spellings is consistent with that hypothesis, underscoring the usefulness of spelling variation as a source of evidence for morphological segmentability and for the role of morphological representations in language production and comprehension.

Keywords: spelling, misspellings, derivation, affixes, orthography, morphological boundaries, morphology

Pronunciation variation has established itself as a source of evidence on a range of questions in psycholinguistics, including the processing of morphologically complex words (Cohen, 2014; Hay, 2003, 2007; Hay and Baayen, 2005; Kamps, Ernestus, Schreuder, and Baayen, 2005; Plag, 2014; Seyfarth, Garellek, Gillingham, Ackerman, and Malouf, 2017, among many others). For example, Hay (2003, 2007), Plag and Ben Hedia (2018), and Smith, Baker, and Hawkins (2012) found acoustic reduction of affixes correlating with morphological categories. Along similar lines,

Sproat and Fujimura (1993) and Lee-Kim, Davidson, and Hwang (2013) show that the duration and degree of velarization of /l/ at morpheme boundaries in English depends on the strength of the boundary. A smaller literature (e.g. Assink, 1985; Fayol, Largy, & Lemaire, 1994; Largy, 1996; Schmitz, Chamalaun, & Ernestus, 2018) has mined spelling variability for psycholinguistic evidence. In this paper, we argue that misspellings such as <available> and <invisible> also reflect the presence and strength of morphological boundaries, along with several other lexical properties.

Research on pronunciation variation does not characterize variants as 'mispronounced'. Therefore, misspellings might seem less analogous to pronunciation variation and more like 'slips of the pen', i.e. a written analog to speech errors (see e.g. Cutler, 2011; Dell, 1986). However, a characteristic property of speech errors is that talkers, once they become aware of what they have said, consider themselves to have made an error. By contrast, a spelling like <available> may well represent what an individual believes to be the correct spelling of a word. As Sandra (2010, p. 425) points out "it might be argued that the incorrect orthographic representations are also stored in the mental lexicon". Nonstandard spellings may therefore have more in common with pronunciation variants than with speech errors.

Orthographically licensed variation in spelling has already been recognized as reflecting morphological structure, specifically the use of spaces and hyphens in English compounds. Kuperman and Bertram (2013) show that the alternation between concatenated (e.g. <postcard>), hyphenated (<word-play>), and spaced (<trash can>) compounds in English reflects aspects of morphological processing and in turn affects processing during reading (Falkauskas & Kuperman, 2015; Rahmanian & Kuperman, 2019). Specifically, Kuperman and Bertram (2013) argue that variability in compound spelling correlates with differences in morphological boundary strength (Hay & Baayen, 2002, 2005).

Based on the research on compound spelling variability, and the research on keystroke variability in compounds and derived words, we hypothesized that spelling mistakes in English derivational affixes likewise reflect the segmentability of words, i.e. the presence and the strength of morphological boundaries, among other factors.

Following Hay and Baayen (2002, 2005), we consider morphological boundary strength to be a gradient property of morphological boundaries reflecting their salience and the degree of segmentability of complex forms (Hay & Baayen, 2002, 2005). On this view, boundary strength is influenced by multiple parameters such as morphological productivity, semantic transparency, and the relative frequency of base and derived word (Blumenthal-Dramé et al., 2017; Hay & Baayen, 2002; Vannest, Newport, Newman, & Bavelier, 2011). The gradience of morphological boundaries, and the consequences of segmentability for the representation of words and morphemes, are a matter of active debate (Baayen, 2014; Hay & Baayen,

2005). In a recent contribution to that debate, Schmitz et al. (2018) argue that misspellings of Dutch verb forms reflect holistic representations vs. on-the-fly generation of inflectional endings. Effects of morphological boundary strength on spelling behavior thus bear upon key issues in research on lexical processing.

The empirical starting point for our hypothesis about spelling was the observation that pairs of English derivational affixes such as *-able* and *-ible* are often confused with one another. For example, at the time of writing, the Wikipedia page listing frequent misspellings in Wikipedia listed 4277 misspellings of 3077 word forms, many of them morphologically complex (Wikipedia, 2017). Of the 3077 word forms in the list, 2876 appear in the English Lexicon Project (ELP, Balota et al., 2007). Of these 2876 forms, 2152 were multimorphemic, according to the ELP. Evidently, morphologically complex words give rise to many spelling errors. Intriguingly, the errors seem to reflect morphological structure, rather than simple letter confusions: For example, the Wikipedia list includes 60 misspellings of words standardly ending in <able> or <ible>. (In what follows, we use italics to represent morphemes, preceded by a hyphen for suffixes (e.g. *-ible*), and angle brackets to represent strings of letters (e.g. <ible>.) In 25 of these 60 cases, the misspelling differs from the standard spelling only in that <able> is replaced with <ible> or *vice versa* (e.g. <acceptable>, <capible>, <formidable>, <hospitable>, <inevitable>, <liible>, <unavailable>; and <accessible>, <comptable>, <eligible>, <feasable>, <incorruptable>, <incredible>, <infallable>, <irresistable>, <permissable>, <plausible>, <possible>). These examples suggest that writers appear to exchange <ible> and <able>, rather than spelling these endings in some other way, e.g. as <eble> or <ibble>, even though all of these variants would result in forms with identical or nearly identical pronunciations.

We hypothesized that such orthographic morpheme-exchanges reflected morphological boundary strength: Boundaries before *-able* tend to be stronger than those before *-ible*, suggesting that words in *-able* tend to be more segmentable than words in *-ible*. (We discuss this point in greater detail below.) We hypothesized that the degree of segmentability should affect spelling behavior. Against the backdrop of that general hypothesis, we evaluated two specific hypotheses leading to overlapping, but different, testable predictions: The first, which we call the 'segmentability' hypothesis, holds that high segmentability promotes correct affixal spelling. Some evidence supporting that hypothesis comes from case studies of individuals with acquired dysgraphia who produced more correct spellings for multimorphemic forms than monomorphemic ones (see Rapp & Fischer-Baum, 2014, for an overview). The second, which we call the 'typicality' hypothesis, holds that typical instances of affixes should be easier to spell than atypical ones. For example, the word *washable* has several properties typical of words with strong morphological boundaries, making it a typical

instance of a word containing *-able*: Among other things, *washable* is semantically transparent and less frequent than its base, *wash*. By contrast, *available* and *laudable* are atypical words with *-able*, in that they contain weak morphological boundaries: The derived forms are more frequent than their bases (*avail-* and *laud-*), and the semantic relationship between the derived forms and the bases is fairly opaque.

The segmentability hypothesis and the typicality hypothesis make identical predictions about the difficulty of *washable*, *laudable* and *available*, though for different reasons. Both predict *washable* to be an easier spelling target than *laudable* and *available*. The segmentability hypothesis predicts this pattern because *washable* contains the strongest morphological boundary of these three words. The typicality hypothesis predicts the same pattern because the properties of *washable* match those of its suffix (*-able*), whereas *laudable* and *available* are atypical environments for *-able* and invite the non-standard spellings <laudible> and <available>.

The predictions of the two hypotheses diverge in the case of *-ible*-words with strong boundaries. For example, the word *accessible* is semantically transparent and is less frequent than its base *access* (for example based on the SUBTLEXus database, Brysbaert and New, 2009). The segmentability hypothesis predicts *accessible* to be a relatively easy spelling target – the word is highly segmentable – whereas the typicality hypothesis predicts it to be a difficult spelling target – the word contains an atypical instance of *-ible*, making <accessable> the expected spelling. More generally, the segmentability hypothesis predicts affixes in highly segmentable words to be easy spelling targets, regardless of the identity of the affix. The predictions of the typicality hypothesis depend on the identity of the affix, more specifically on the match or mismatch between properties of affixes and words.

We tested these hypotheses by examining three pairs of competing spellings, representing three constellations of morphological boundary strength: <able>/<ible>, <ence>/<ance>, and <ment>/<mint>. The suffixes *-able*/*-ible* differ in boundary strength, while *-ance* and *-ence* do not. The suffix *ment* tends to form salient boundaries. While <ment> lacks a suffixal twin in standard orthography, it competes with a non-standard affixal spelling <mint>. The pairs *-ance*/*-ence* and the singleton *-ment* thus function as a control condition. If spelling behavior reflects the boundary strength associated with each affix (as opposed to the salience of the boundary in a given word, i.e. a stem+suffix combination), then boundary strength should affect the spelling of *-ible*/*-able*, i.e. the suffixes that differ in boundary strength, but not *-ence*/*-ance* or *-ment*, i.e. the suffixes that did not differ in the boundary strength, or the suffix without a competitor.

Our data come from a large unmoderated source of written productions: Tweets, i.e. short messages posted on the internet by means of a messaging

and social networking service, Twitter (Twitter, 2006, <http://www.twitter.com>). Related research also using Tweets is reported in Schmitz et al. (2018). We analyze the distribution of spelling variants by means of logistic regression models taking into account a range of morphological, orthographic and lexical predictors of spelling errors.

Morphological boundary strength and word segmentability have been found to be reflected in a number of behavioral and distributional phenomena. Accordingly, a range of variables have served as measures of boundary strength. We therefore begin our discussion by summarizing prior research on measures of morphological boundary strength generally and the boundary-relevant properties of our target suffixes in particular. We then turn to other factors besides boundary strength that may affect the spelling of our target words. A general assumption underlying the current study is that one and the same affix may occur in words that are more or less segmentable. In considering the role of boundary strength in spelling behavior, it is important, therefore, to consider the properties of bases and whole word forms, as well as those of affixes. We do so in the opening section of the Results.

Background

Morphological boundary strength

There are at least three types of measures of boundary strength: semantic transparency, base type (a categorical measure), and gradient distributional properties of morphemes and phonological segments. In this study, we concentrate on base type and distributional measures.

Undoubtedly the most widely discussed measure of boundary strength in English is a binary distinction between 'weak' boundaries (often represented with a plus sign) and 'strong' ones (often represented with a hash mark) (see e.g. Chomsky & Halle, 1968; Dressler, 1985; Kiparsky, 1982; Siegel, 1979). In this categorical scheme, boundaries between bound roots and affixes are held to be 'weak', while boundaries at the edges of free bases are held to be 'strong'. Similarly, English derivational affixes have been analyzed as falling into classes differing in boundary strength, on the basis of phonological (e.g. stress-shifting) and semantic criteria. Also related to this binary distinction is the fact that compounds, as combinations of morphologically free constituents, are held to contain strong internal boundaries.

More recently, Hay (2003); Hay and Baayen (2005) proposed gradient measures of boundary strength. One such measure, sometimes referred to as

'base-to-derived' or 'relative' frequency, is estimated by calculating the ratio of the frequencies of the base and the whole-word frequency, i.e. the derived form. The higher the ratio of base frequency and whole-word frequency, the stronger the morphological boundary and the more segmentable the word. For instance, *government* is far more frequent than its base *govern* and is therefore less easily segmented than, for example, *enjoyment*, whose base is far more frequent than its base. Two additional gradient lexical properties correlating with boundary strength are semantic transparency and morphological productivity (Hay & Baayen, 2002): More semantically transparent formations such as *shoeless* are argued to contain stronger boundaries than semantically more opaque ones such as *regardless*, and highly productive affixes, such as *-ness*, tend to be associated with stronger boundaries than less productive ones such as *-th*. Finally, boundary strength also manifests itself in affix ordering (Hay & Plag, 2004; Plag & Baayen, 2009; Zirkel, 2010): Weak morphological boundaries tend to occur 'inside' of strong boundaries, rather than 'outside', a pattern termed 'complexity-based ordering' (see Hay and Baayen, 2002 for discussion and illustration).

Morphological boundaries and typing speed

Temporal properties of written language production have already been shown to be affected by the presence and strength of morphological boundaries. Gagné and Spalding (2014); Libben, Weber, and Miwa (2012) and Libben and Weber (2014), for example, investigated typing latencies for English compounds without spaces (e.g. *strawberry*) and found that inter-keystroke intervals were significantly elevated at the boundary between the stems. Similarly, Gagné and Spalding (2016a) found differences in typing speed between monomorphemic words and compounds. Similar results were obtained by Libben, Jarema, Luke, and Bork (2018) for other kinds of complex words in English and French, i.e. stem-stem combinations, as in *xylo-phone*, prefix-stem combinations, as in *im-plant*, and stem-suffix combinations as in *form-ation*). In all conditions apart from French prefix-stem words the letter transition across the morpheme boundary showed longer keystroke intervals than the preceding and following letter transitions. Not only the presence of a morphological boundary, but also its strength has been shown to affect typing behavior: Gagné and Spalding (2014, 2016a, 2016b); Libben and Weber (2014) found that compounds with semantically transparent constituents showed different inter-keystroke intervals compared to those with non-transparent constituents. Testing compound frequency and head frequency in addition to semantic transparency, Sahel, Nottbusch, Grimm, and Weingarten (2008) found that keystroke intervals varied with the strength of the compound-internal boundary. These findings lend general support to the idea that boundary strength can affect written language production.

The variables relating to boundary strength in our statistical models were informed by three measures: a binary distinction between free bases and bound roots, and two gradient measures (relative frequency and bigram probabilities). We describe these measures in detail in our Methods section. The variables are grounded in previous treatments of our target suffixes, to which we turn next, before discussing additional factors that we had reason to believe would affect the spelling of our target words.

Target suffixes

To understand the effects of boundary strength on spelling behavior, we examined three pairs of spellings, each of which appeared to be tricky spelling targets, based on informal observations: <able>/<ible>, <ence>/<ance>, and <ment>/<mint>. While <ment> lacks a suffixal twin in standard orthography, it competes with a non-standard affixal spelling <mint>. We are not aware of any systematic studies of the non-standard spelling <mint>. The segmentability of words with the affixes *-able/-ible*, *-ence/-ance*, and *-ment*, however, has been studied in a fair amount of detail.

able/ible

Many sources treat <ible> and <able> as two orthographic variants or two allomorphs of one and the same suffix (see e.g. Plag, 2003, 95). Bauer, Lieber, and Plag (2013, 307) explicitly label the two items ‘allomorphs’, on the grounds that they do not appear to differ in meaning. And yet, previous research also provides evidence for differences between *-ible* and *-able*, summarized in Table 1, which is partly based on Table 14.1 in Bauer et al. (2013, 290ff.). It will be observed that *-able* tends to be associated with stronger morphological boundaries than *-ible*, in that (1) it attaches to a wider range of bases than *-ible*, (2) is considered highly productive, and (3) does not typically induce stress shifts or other morphophonological alternations in the bases it combines with (with few exceptions, such as *admirable*).

Table 1. Characteristics of *-able* and *-ible*, based on Adams (2001); Bauer et al. (2013); Marchand (1969)

Characteristic	<i>-able</i>	<i>-ible</i>
Base category	verbs, phrasal verbs, nouns, bound roots, compounds	(non-phrasal) verbs, bound roots
Stratum of bases	native, non-native	non-native
Stress shifting	rare	rare
Base allomorphy	rare	frequent
Productivity	high	limited

-ance/-ence

Like *-able* and *-ible*, *-ance* and *-ence* are treated as allomorphs of one and the same affix by Bauer et al. (2013, Section 10.2.1), on the grounds that we find the same semantics for a whole set of phonologically related formatives: *-ance*, *-ence*, *-ce*, *-cy*. As Bauer et al. point out, *-ance*, *-ence*, *-ce*, *-cy* have several puzzling properties, two of which may give rise to spelling uncertainty. First, almost every derivative in *-ance* or *-ence* has a corresponding adjectival derivative ending in *-ant* or *-ent*, justifying two morpheme-based parses: $Xent + ce$ or $X + ence$ and thus making the location of the boundary unclear. Second, the distribution of the <a> vs. the <e> is not easily predictable from the base, which might further increase the chances of misspelling. Bauer et al. (2013) also mention several facts suggesting that *-ance* may be associated with slightly stronger boundaries than *-ence*: There are a few cases of *-ance* attaching to non-latinates bases (*believance*, *coming-outtance*), but none for *-ence*, suggesting a wider range of bases for *-ance* compared to *-ence*. There also appear to be more word types with *-ance* than with *-ence*. In all other respects, the two endings appear to behave similarly.

-ment

The suffix *-ment* derives event nominalizations with a wide range of possible readings, depending on the base and the context (see e.g. Kawaletz & Plag, 2015). *-ment* is most often found with verbal bases, though other bases are also found. Many words with *-ment* contain bound roots, but many others contain free bases. The suffix was highly productive in the 19th century, and it is still moderately productive today. Given its productivity, and given that *-ment* is the only consonant-initial suffix in the current study, a property associated with perceptually salient boundaries (Hay, 2003), we expect it to be associated with strong morphological boundaries, compared to the other target suffixes.

Other factors likely to affect spelling difficulty

Among other factors likely to affect spelling behavior, perhaps the most prominent one is lexical frequency. Other things being equal, one might expect highly frequent (and therefore familiar) words to be easier to spell than rare ones (see e.g. Assink, 1985; Fayol et al., 1994; Largy, 1996, for discussion). On the other hand, high usage frequency also entails frequent opportunities for misspelling a word, and for finding it in a corpus. We return to this point in the description of our data and sampling methods.

We also considered the segmentability of the base. We suspected that morphological complexity of bases might affect word segmentability and hence spelling

behavior: Words in which our target suffixes follow morphologically complex material (e.g. *indescribable*, *uncombable* or *imperturbable*) may themselves be more readily segmented than words with morphologically simple bases. Although we are not aware of studies testing this intuition directly, we reasoned that salient boundaries within (complex) bases might promote segmentability of the form as a whole we therefore included the presence of a boundary within bases as a predictor in our model. We return to this issue in the General Discussion.

Base segmentability may also shape the behavior of another variable that has been extensively studied in research on lexical processing, which is word length. The expected effects of word length in letters on suffix spelling may not be obvious: Although long words offer more opportunities for error, that fact need not result in greater numbers of errors on a particular letter in an affix, such as the <a> or <i> in <able> or <ible>. Long words need not be difficult to spell, especially if they saliently contain components that are relatively easy spelling targets. Indeed, long words are more likely to be morphologically complex than short ones. If it is indeed the case, as we hypothesized, that high segmentability promotes correct spelling, then word length might be positively associated with spelling accuracy.

We are aware that word length (in letters) has been found to be negatively associated with accuracy, both in individuals with dysgraphia (Caramazza, Miceli, Villa, and Romani, 1987) and in healthy individuals (Bloomer, 1956; Cahen, Craun, and Johnson, 1971; Carlisle, 1988; Spencer, 2007). However, as noted in Spencer (2007), among others, apparent effects of word length may reflect other factors, such as the number of letters per grapheme, as well as lexical frequency.

Several additional visual and phonological factors may affect spelling behavior. For example, misspellings resulting in the letter sequence <ii>, such as *amiible* or *remediible* 'look wrong', presumably partly because the sequence <ii> is extremely rare in English (although it does occur, e.g. in the spelling <Hawaii>, i.e. <Hawai'i> without the okina). In our analyses, we take effects of the rarity of specific letter combinations into account by considering bigram probabilities.

Another fact related to specific letter combinations that may well affect spelling concerns pronunciation. In most of our target pairs pronunciation offers little help in deciding between the two spellings: For example, <available> and <available> are both plausible spellings of [ə'veɪləbl̩]. However, in one class of target words, pronunciation does provide strong spelling clues: The pronunciation of <g> and <c> after [i] vs. [a] may make misspellings like <allocible> for allocable or <legable> for legible far easier to avoid and detect than forms like invisible. Therefore, we expect such misspellings to be relatively rare. We return to this issue in the description of our variables and in the discussion of the results.

Methods

Target words

The target words for the analysis were all words ending in <able>, <ible>, <ance>, <ence>, and <ment> in the CELEX lexical database (Baayen, Piepenbrock, & Gulikers, 1995), with the following exceptions:

Doublet Pairs. Doublet pairs such as *discernable/discernible*, *indispensable/indispensible*, *collectable/collectible*, *valance/valence*, where both spellings were listed in the English Lexicon Project (Balota et al., 2007).

Pseudoaffixes. Words ending in the strings <ance>, <ence>, <ment>, <able>, and <ible> where these endings represented stems, such as *(un)able*, *dement* or parts of stems, such as *freelance*, *bible*, *table*, *foible*, *advance*, *askance*.

Pseudobases. Words like *cement*, whose base is not attested in other words in CELEX (in transparently related meanings). Such items warrant future investigation, particularly as informal web searches revealed that spellings such as *cemint* are indeed attested.

Hyphens and whitespaces. Items containing hyphens or white spaces, such as *enfant terrible*.

We excluded words like *crucible* and *thurible*, in which <ible> is historically derived from Latin *-(i)bulum* rather than *-a/ibilis*. We considered excluding words in which the target endings optionally bore main stress, e.g. category-ambiguous words like *torment* and certain unambiguous forms, such as *refinance* and *dalliance*. In the case of *dalliance*, some speakers not only stress the final syllable in the words, but also nasalize the vowel in the final syllable. We decided to include this word on the grounds that (unlike *thurible* or *crucible*) it contains the target suffix *-ance*. Our list of spelling variants initially included the form <ambiance>, which we later identified as an officially recognized orthographic doublet. The pair <ambiance, ambiance> was therefore also excluded.

Data collection

The data were collected using the searchTwitter function in the twitterR library (Gentry, 2015) in R (R Development Core Team, 2008) in five separate data collection sessions between December 2016 and October 2017. The searchTwitter function returns tweets posted during the previous seven days (Gentry, 2015). Each target spelling was included in two data collection sessions. At the time when these searches were carried out, the number of hits for a given target in a given session was capped at 2000. The data collection sessions were spaced several months apart. Our analyses are based on the total number of hits, i.e. the sum of

the number of hits in the two sessions for each target. The search targets were the misspelled versions ('spelling variants') of our target words, which were created by changing <able>, <ible>, <ance>, <ence>, and <ment> in the standard spelling to <ible>, <able>, <ence>, <ance>, and <mint>, respectively. The resulting corpus consisted of 22857 tweets. The counts of word types and misspelled tokens for each target suffix are listed in Table 2.

Table 2. Number of distinct word types and total count of misspelled tokens for each of five target *suffixes*

Suffix	Types	Misspelled tokens
-able	500	4055
-ance	41	1868
-ence	77	5753
-ible	96	10022
-ment	254	1159

Statistical modeling strategies

We used binary logistic regression to fit two sets of models. In one set, the outcome variable coded whether a given spelling variant (e.g. <edable>) was attested at all in our corpus. In the second set, we set a higher threshold, with the outcome variable indicating whether a given spelling variant was attested at least six times. As Table 3 shows, the proportion of variant word types attested at least six times ranged from 0.09 for *-ment* to 0.51 for *-ence*, *-ance*. While the higher threshold reduces the size of the data set, it also potentially reduces the amount of noise and the possibility of overestimating the occurrences of misspellings. A similar strategy, of using binary logistic regression models with varying thresholds for modeling the probability of spelling errors was implemented in Bar-On and Kuperman (2019). An alternative strategy would have been to include both correct and incorrect spellings in our corpus, and then to use a variable coding whether a form was spelled correctly as the outcome variable. That strategy was inadvisable here, due to the fact that, as already mentioned, the number of hits per session was capped at the time the corpus was put together. This cap would have inevitably led to a distorted picture of the relative frequency of correct and incorrect spellings, particularly in the case of high-frequency words: Most of the target words would have produced exactly 2000 hits – and any word whose spelling variant was also attested 2000 times would have come to look as though it was misspelled 50% of the time.

We used backwards elimination, i.e. initially including the full set of predictors and subsequently eliminating non-significant predictors, beginning with

Table 3. Number and proportion of misspellings found at least once or 6 times for each of five target suffixes

	n > 0	n > 5	Prop above 0	Prop above 5
<i>-ible, -able</i>	194	111	.33	.19
<i>-ment</i>	62	22	.24	.09
<i>-ence, -ance</i>	90	60	.76	.51

non-significant interactions. Model improvement was determined based on change in the AIC. The order of removal of non-significant effects was based on the degree to which a given effect was established in previous literature, with well-established effects being retained longer than effects for which, to our knowledge, little or no previous empirical support was available. All continuous variables were centered and log-transformed. Separate models were fitted for each (pair of) target suffixes. For each pair, we report the model using the outcome variable coding whether a given spelling variant (e.g. <edable>) was attested at all in our corpus. Where the results differed for the models using the more stringent threshold, i.e. modeling whether a given spelling variant was attested at least six times, we note that fact.

In order to test the predictions of the 'typicality hypothesis' introduced above, we included the interactions of Suffix with each of the other variables, except in cases where too few target words were available for a given combination of Suffix with the factor in question (for factorial variables), or where target words for a given suffix were too sparsely distributed (in the case of continuous variables). We also included interactions of Target frequency with the variables intended to capture segmentability effects in our initial models, to explore the effects of the segmentability variables at different frequency bands.

We considered alternative outcome measures, such as the ratio of the frequency of the spelling variants and the standard spellings, or the entropy when selecting one of the available variants given their probability distribution (see, for example, Bar-On and Kuperman (2019) for discussion). However, for our data set we were not able to compute these measures meaningfully because, as mentioned before, at the time when data collection was performed, the searchTwriteR capped the frequency counts for each target at 2000. This constraint would have resulted in identical frequency counts (of 2000) for many of our target words, including words that clearly differ in frequency, based on SUBTLX or CELEX. None of the spelling variants occurred frequently enough for the cap to be a problem.

As an alternative to logistic regression we also explored models of count data (negative binomial regression and hurdle regression), but the distribution of the counts of misspellings violated the pertinent model assumptions. We therefore used logistic regression with two different thresholds, as described above.

Descriptions of variables

Our aim was to take into account factors known to affect the processing of derivationally complex words, including morphological complexity and boundary strength, base segmentability, lexical frequency, and word length. In addition, we wished to take into account visual factors. In exploratory analyses, we determined a set of variables indexing these factors:

Base type

A binary factor distinguishing two types of morphological bases: Roots (e.g. applic-*able*) vs. word (e.g. govern-*able*).

Base complexity

A binary factor distinguishing words with morphologically simple bases (e.g. washable) vs. words in which the material preceding the target suffix contains internal morphological boundaries, such as complex bases (e.g. *act-ionable*, *exchangeable*), prefixes before the morphological base the target suffix combines with (e.g. *non-flammable*), and morphologically ambiguous forms (e.g. *resolvable*). The term 'base' is used somewhat loosely here, to refer to the string preceding the target suffix; we do not intend any claim as to the morphological base of the word or the combinatory properties of the suffix.

Length

The target word's length in letters.

Target bigram

The transitional probability of the initial letter of the target, given the base-final letter (i.e. the forward bigram probability). We estimated this probability based on the number of distinct words the bigram occurs in (termed the 'non-positional versatility') as reported in Solso and Juel (1980, 298). The Target bigram variable represents the number of word types in which a given base-final letter is followed by a given suffix-initial letter, divided by the total number of word types with a given base-final letter.

Variant bigram

The transitional probability of the initial letter of the spelling variant (e.g. <available>), given the final letter of the base, estimated in the same way as for the Target bigram variable. The intuition behind this variable is that spelling variants 'look odd' to varying degrees: For example, consecutive tokens of <i> (as in <insatiable>) seem highly noticeable because of the rarity of <ii> in correct

spellings of English words (as mentioned above). More generally, low-probability letter sequences may make misspellings easier to spot and hence more likely to be corrected. We discuss this variable further in the Discussion below.

Target frequency

As estimates of the frequency of the target word, we used the lemma frequencies in the Corpus of Contemporary American English (COCA, Davies, 2013).

Base frequency

As a gradient measure of boundary strength, we used the frequency of the orthographic base to which the target suffix is attached. The orthographic base was the letter string preceding the target suffix, i.e. the word minus the suffix (e.g. <avail>, <comprehens>). Higher values of this measure indicate greater segmentability.

Our measure of base frequency calls for further explanation. Measures of base frequency, and relative (base-to-derived) were originally developed in investigations of words containing free bases, i.e. forms that have lexical frequencies of their own. Estimating the frequency of bound roots raises methodological difficulties, as bound roots by definition do not occur independently. Two main strategies have been employed to circumvent this problem. Some studies have excluded words with bound roots from consideration altogether (e.g. Hay, 2001), others have assigned bound roots a frequency of zero and added a small constant to all frequencies, to enable logarithmic transformations of items with a raw frequency of zero (e.g. Baayen, Feldman, and Schreuder, 2006, 294). However, this solution is somewhat unsatisfactory, in that many bound roots appear in multiple words, whose frequency may well affect the target word's decomposability.

For the current study, we developed a string-based measure of base frequency usable for bound roots and free bases equally. The string-based base frequency is the cumulative frequency of all words containing the sequence of letters formed by removing the target suffix from a target word, e.g. in the case of *possible* or *readable* the frequency of all words beginning with <poss> or <read>, respectively. The resulting estimates of base frequency are by necessity noisy, due to accidental overlap in letter strings. For example, the base frequency of *possible* includes the frequency of *possum* by this metric. We accept the noisiness of the string-based measure of relative frequency, considering it an empirical issue whether it would converge with other variables targeting boundary strength. To see whether this was the case, we checked the relationship between base frequency and base type: If the measure succeeded in recovering effects of boundary strength, then estimated base frequencies should be higher for bases that are words (e.g. *read*) than for bases that are bound roots (e.g. *poss-*), other things being equal. That expectation was partially supported: For the words with *-ible*, *-able*, base frequency was signifi-

cantly higher for free bases than bound roots ($W = 44218, p < .0001$). For words with *-ment*, however, base frequencies of words with free bases vs. bound roots did not differ significantly. We did not perform a similar check for the words with *-ance* and *-ence* because free bases were almost entirely absent from that set.

We entered Base frequency and Target frequency as separate predictors in our models, instead of a variable coding the relative (base-to-derive) frequencies. One reason for this decision was that we did not expect an additional effect of relative frequency on top of base frequency and target frequency (Plag & Baayen, 2009). A second reason is that relative frequency is computed as the quotient of base frequency and whole-word frequency and is therefore not independent of either, which renders using relative frequency together with base frequency or word frequency (or both) in one regression model problematic. A third reason is conceptual: Whereas we were interested in Base frequency as a measure of boundary strength, we expected that Target frequency would play a role not only because of its relationship to boundary strength, but also because of its role in determining sample size: Misspellings of words with high frequency are *ipso facto* more likely to be attested in our sample, regardless of boundary strength or any other morphological property. We therefore wished to treat Base frequency and Target frequency as two separate variables in our models, so as to examine the sampling effect and any effects of morphological boundary strength separately.

Results

Properties of the target suffixes and target words

We selected the affixes discussed here in the expectation that *-ible* vs. *-able* differed in boundary strength, that *-ence* vs. *-ance* did not, and that *-ment* was associated with fairly strong boundaries. We therefore begin our analyses by determining whether that expectation was borne out in our sample.

With respect to the distribution of base types, i.e. words whose bases are roots vs. free bases, our target suffixes patterned as expected, as shown in Table 4:

Table 4. Categorical properties of target *suffixes*

Suffix	Word types	Bound roots	Complex bases	Misspelled tokens
<i>-able</i>	500	111	189	4055
<i>-ible</i>	96	46	38	10022
<i>-ance</i>	41	39	7	1868
<i>-ence</i>	77	74	17	5753
<i>-ment</i>	254	44	83	1159

Bound roots account for only about one fifth (111 out of 500) of the words with *-able*, but about one half (46 out of 96) of the words with *-ible*, in line with the expectation that *-able* words should be more likely to contain strong boundaries than *-ible* words. Also in line with our expectations, *-ance* and *-ence* did not differ with respect to base type: both attach almost exclusively to bound roots. Finally, only one seventh (44 out of 254) of the words *-ment* contained bound roots, confirming that this suffix tends to be associated with strong morphological boundaries.

We also compared the relative (i.e. base-to-derived) frequencies associated with these suffixes: If our string-based measure of base frequency is valid, one would expect the relative (i.e. base-to-derived) frequency for *-able* to be higher than for *-ible*. That was indeed the case ($t(140) = 2.928, p = 0.002$), adding support to the assumption that *-able* was indeed associated with stronger morphological boundaries than *-ible*. The suffixes *-ance* and *ence*, on the other hand, did not differ in base-to-derived frequency ($(t(66.04) = -0.316, p = .624)$), consistent with our assumption.

Descriptive statistics for the continuous variables in our models are shown in Table 5. Tables 6 through 8 show the pairwise (Spearman) correlations among the numeric variables in our models.

Table 5. Median values of numerical properties of target words by *suffix*

Suffix	Length	Target frequency	Base frequency	Target bigram	Variant bigram
<i>-able</i>	10.00	5.23	8.82	.09	.14
<i>-ance</i>	10.00	6.18	8.04	.09	.20
<i>-ence</i>	10.00	6.15	7.96	.20	.07
<i>-ible</i>	11.00	5.89	8.42	.11	.09
<i>-ment</i>	10.00	6.50	8.73	.03	.03

Table 6. Pairwise (Spearman) correlations of gradient variables for *-ible/-able*

	Length	Target frequency	Base frequency	Variant bigram
Length	1.00	-.08	-.02	-.03
Target frequency	-.08	1.00	.02	-.01
Base frequency	-.02	.02	1.00	-.03
Variant bigram	-.03	-.01	-.03	1.00

Table 7. Pairwise (Spearman) correlations of gradient variables for *-ence/-ance*

	Length	Target frequency	Base frequency	Variant bigram
Length	1.00	-.29	-.11	.16
Target frequency	-.29	1.00	.14	-.14
Base frequency	-.11	.14	1.00	.05
Variant bigram	.16	-.14	.05	1.00

Table 8. Pairwise (Spearman) correlations of gradient variables for *-ment*

	Length	Target frequency	Base frequency	Variant bigram
Length	1.00	.00	-.19	-.03
Target frequency	.00	1.00	.13	-.04
Base frequency	-.19	.13	1.00	.05
Variant bigram	-.03	-.04	.05	1.00

Modeling results

-ible/-able

The model for *-able/-ible* after stepwise elimination of non-significant predictors is summarized in Table 9. There was a significant effect of Suffix, reflecting the fact that words with *-ible* were more liable to be misspelled than words with *-able* ($\beta = 3.892$, $p < .0001$). There was also a significant effect of Base complexity, indicating that words with complex bases were less likely to be misspelled ($\beta = 0.003$, $p < .0001$). The effect of Variant bigram indicates that misspellings were increasingly likely to be found with increasing probability of the bigram at the suffix boundary ($\beta = 11.133$, $p < .0001$). Increasing target word frequency was associated with greater likelihood of the spelling variant being attested ($\beta = 0.664$, $p < .0001$), and so was increasing base frequency ($\beta = -0.151$, $p = 0.0017$). Finally, there was a significant interaction of Base frequency with Target frequency ($\beta = 0.081$, $p = 0.0013$). None of the interactions of Suffix with any other variable reached significance.

Table 9. Logistic regression model of *-ible/-able* misspellings

	Estimate	Std. Error	<i>z</i> value	Pr(> <i>z</i>)
(Intercept)	-1.5371	0.1834	-8.38	.0000
Suffix	3.8922	0.3906	9.96	.0000
Base complexity	-0.7497	0.2528	-2.97	.0030
Variant bigram	11.1330	1.8481	6.02	.0000
Target frequency	0.6640	0.0758	8.76	.0000
Base frequency	-0.1514	0.0483	-3.13	.0017
Target frequency:Base frequency	0.0812	0.0253	3.21	.0013

There were 111 word types with *-ible/-able* with at least 6 spelling variants in our sample. The pattern of results in the model using that higher threshold as the outcome variable (summarized in Table 10) was similar to the previous model, with two differences: First, in the model with the higher threshold, there was a marginally significant interaction of Suffix with Base frequency, indicating that words with *-ible* were less likely to be misspelled with increasing base frequency. Secondly, the interaction of Base frequency with Target frequency was marginally

significant ($\beta=0.054$, $p=0.0643$), and the simple effect of Base frequency was non-significant.

Table 10. Logistic regression model of *-ible/-able* misspellings attested 6 times or more

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-2.6852	0.2563	-10.48	.0000
Suffix	3.7902	0.4077	9.30	.0000
Base complexity	-0.8909	0.3058	-2.91	.0036
Variant bigram	11.7904	2.5304	4.66	.0000
Target frequency	0.5770	0.0809	7.13	.0000
Base frequency	0.0115	0.0729	0.16	.8747
Target frequency:Base frequency	0.0539	0.0291	1.85	.0643
Suffix:Base frequency	-0.2248	0.1189	-1.89	.0588

The interaction of Base frequency with Target frequency is plotted in Figure 1, which shows predictions and confidence bands of Base frequency for ten percentile ranges of Target frequency. The effect of Base frequency (plotted along the x-axis) and the outcome variable (i.e. presence of at least one misspelling) varied across frequency bands: For target words of low to medium frequency, increasing base frequency was associated with fewer errors, consistent with the idea that higher segmentability was associated with fewer misspellings. The effect of Base frequency was attenuated with increasing Target frequency and was reversed in the two highest percentile ranges of Target frequency, in which increasing Base frequency was associated with greater numbers of misspellings.

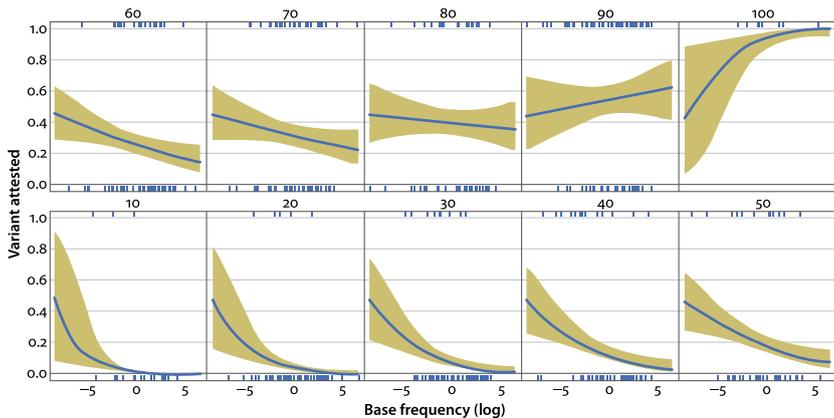


Figure 1. The interaction of Base frequency and Target frequency: Effect of Base frequency on <ible>/<able> variation for ten quantiles of target frequency

Finally, as a test of the validity of Base frequency, we compared the behavior of this variable to that of a traditional estimate of base frequency in models of the subset of target words with free bases. Recall that traditional estimates only apply to free bases, not to bound roots. If the two variables tap into the same underlying property, they should have similar effects when applied to words with free bases. This was the case. We interpret this as an indication that our string-based measure of base frequency taps into the same underlying property as traditional measures applicable only to free bases.

-ance/-ence

The set of words with *-ance/-ence* included very few words with complex bases (only 7 for the suffix *-ence*), so we refrained from entering Base complexity in the model. Including Length also turned out to be problematic because of the presence of five very long words (of 13 or more letters). In a preliminary model, there was a significant effect of Length, which disappeared after exclusion of the outliers. Here, we document the models without Length as a predictor, but including the five long words.

Table 11. Logistic regression model of *-ence/-ance* misspellings

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	1.76	0.334	5.26	.000
Variant bigram	12.020	5.250	2.29	.022
Target frequency	0.69	0.157	4.40	.000

The model for *-ance/-ence*, after backward elimination, is summarized in Table 11. As was the case for *-able/-ible*, there was a significant effect of Variant bigram ($\beta=12.02$, $p=0.022$), indicating that misspellings were increasingly likely to be found with increasing probability of the bigram straddling the suffix boundary. There was also a significant effect of Target frequency ($\beta=0.69$, $p<.0001$), indicating that misspellings were increasingly likely to be found with increasing target word frequency. Suffix (*-ance* vs. *-ence*) did not yield a significant main effect, nor did it participate in any significant interactions. In the model using the higher threshold, i.e. modeling the probability of a spelling variant being found more than five times, Target frequency was the only significant predictor; as in all other models, increasing target frequency was associated with an increasing probability of variants being attested the required number of times.

-ment

The model for *-ment* is summarized in Table 12. After backwards elimination, there was a significant effect of Target frequency ($\beta = 0.646$, $p < .0001$), indicating that variants were more likely to be attested with increasing target word frequency. There was also a marginally significant effect of Base type ($\beta = -0.761$, $p = 0.064$), suggesting that target words with free bases were less likely to be misspelled. However, the effect of Base type was non-significant in the model using the higher threshold (not shown here), where only the effect of Target frequency was significant ($\beta = 0.462$, $p < .001$). At neither threshold was there an effect of Variant bigram or Base frequency or an interaction of Target frequency with Base frequency.

Table 12. Logistic regression model of *-ment* misspellings

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-0.95	0.367	-2.59	.010
Base type	-0.76	0.411	-1.85	.064
Target frequency	0.65	0.101	6.39	.000

Summary of results

The pattern of results is summarized in Table 13. For all three pairs we considered, higher lexical frequency of the target word was associated with an increased probability of the spelling variant being attested, as one would expect for any form: The more frequently an item occurs, the more likely it is to be attested in any given corpus.

Table 13. Summary of results (threshold > 1). ‘yes’ represents a significant effect ($p < 0.05$, or smaller), ‘(yes)’ represents a marginally significant effect

Type of effect	Variable	-able/-ible	-ance/-ence	-ment/-mint
morphology	Suffix	yes		
	Base type			(yes)
	Base complexity	yes		
	Base frequency	yes		
sampling	Target frequency	yes	yes	yes
orthography	Variant bigram	yes	yes	

The effects of the other predictors differed for the three pairs we considered. For the pairs *-able*, *-ible* and *-ance*, *-ence*, i.e. the suffixes in which target and variant differed in the initial segment, there was a positive effect of Variant bigram, such that the higher the variant bigram probability, the higher the probability of the

spelling variant being attested. No other variables besides Target frequency and Variant bigram were significant for *-ence/-ance*. In particular, there was no significant effect of Suffix, meaning that *-ence* vs. *-ance* appeared to be about equally likely to be misspelled.

Spelling variants for *-ible/-able* were less likely to be attested for words with complex bases and higher base frequency. There was an interaction of Base frequency and Target frequency, such that increasing Base frequency was associated with decreasing probability of misspelling for words up to about the 80th percentile for Target frequency. For targets with high lexical frequency, there was no effect of Base frequency. The suffixes *-ible* and *-able* differed from one another in that misspellings were more likely to be attested for *-ible* compared to *-able*, even after controlling for lexical frequency. There were no significant interactions with Suffix, meaning that the effects of the predictors did not appear to differ for *-ible* vs. *-able*.

For *-ment*, Target frequency was the only variable that was predictive of whether the spelling variant with <mint> would be attested at least six times. When the outcome variable reflected whether a <mint>-variant was attested at all, there was a marginally significant effect of Base type, i.e. of whether the word contained a bound root vs. a free base, with free bases being associated with a lower incidence of <mint>-variants.

Discussion

We explored predictors of misspellings such as <comprehensible>, and <avoidance>, in which the standard spelling of a target suffix is replaced by that of a similar-sounding suffix. Our general hypothesis was that these spelling variants would show systematic patterns, rather than occurring at random, and that they would reflect morphological boundary strength, among other factors. To explore that hypothesis, we examined spelling errors in one pair of suffixes differing in boundary strength, *-ible/able*, and one pair of roughly equal boundary strength, *-ance/ence*. We also included the spelling <mint> of the suffix *-ment* (as in <statemint>) in our analysis, to ask whether occurrences of <mint> likewise reflected boundary strength. We evaluated two specific hypotheses, which we termed 'segmentability' and 'typicality', about the potential role of morphological boundary strength in spelling variation. In order to be able to include words with bound roots in the scope of our investigation, we developed a new, string-based, measure of base frequency. We validated the measure by comparing its behavior for words with free bases to more conventional estimates. There were several clear patterns, indicating that these spellings indeed did not occur randomly, as well as

some evidence consistent with the hypothesis that they reflected morphological boundary strength, among other factors.

Crucial to evaluating our hypothesis were the presence of an interaction between Target Frequency and Base Frequency for *-ible* /*able*, and the absence of such an interaction for *-ance*/*-ence*, which are not thought to differ in boundary strength, and *-ment*, which lacks a competitor in standard spelling. That interaction is only one of several variables one might wish to consider in an investigation of morphological boundary strength. We refrained from attempting to include additional variables reflecting boundary strength: Such additional variables can be expected to correlate with the base-to-derived frequency, precisely to the extent that they all correlate with – or reflect, or themselves determine – morphological boundary strength.

Importantly, the overall pattern of (morphological and other) effects does not appear to be the result of an across-the-board, 'morphology-blind' default to the most frequent spelling in case of uncertainty. In the case of *-able* / *-ible*, defaulting to <able> would be a reasonable strategy, as word types with *-able* far outnumber word types with *-ible*. Consistent with the 'reasonable default' strategy, *-ible* was more likely to be spelled <able> than the other way around when lexical frequency was controlled. However, no such default strategy seemed to be at work in the case of *-ance* / *-ence*: Even though there were twice as many word types with *-ence* as with *-ance* in our data set, words with *-ance* were no more likely to be misspelled than those with *-ence* when lexical frequency was controlled. As for *-ment*, there are of course words ending in <mint> in standard orthography (e.g. <mint>, <spearmint>, and <varmint> (as a regional variant of <vermin>)), but <mint> in these words does not represent the suffix *-ment*. Therefore, defaulting to <ment> offers a perfectly safe strategy – provided writers recognize the ending in question as representing a suffix at all. The pattern we observed with *-ment* is consistent with <ment> being a default spelling – but in a manner that reflects the recognition of morphological structure. Taken together, our models suggest that writers do not simply default to whichever ending they have encountered more commonly.

Interestingly, spellings like <spearmint> and <peppermint> are by no means rare in print, judging by an informal search on Google Books (<https://books.google.com/>). Such spellings may reflect a kind of 'folk morphology', with speakers treating spearmint as though it contained a suffix. Conversely, forms like <government>, <adjoinmint>, <ailmint>, and <settlemint> might reflect a 'folk suffix' *-mint* competing with the suffix spelled <ment> in standard orthography. On that reading, *-mint* might be considered a 'weak boundary counterpart' of *-ment*, analogous to the relationship between *-ible* and *-able*. In any case, the case of *-ment*

strongly suggests that suffixal spellings are not due to a simple surface default strategy, but reflect morphological structure.

Typicality vs. Segmentability

The modeling results allow us to evaluate two specific hypotheses about the role of morphological boundaries in spelling variation. On the first, the Typicality hypothesis, spelling variants should be more likely whenever the variant is expected, given the properties of the target word and suffix. For example, Typicality would favor spellings like <availible> and <suggestable>: The word *available* is more frequent than its base, which is typical for words with *-ible*, and *suggestible* is less frequent than its base, which is typical for words with *-able*. The standard spellings <available> and <suggestible> are therefore somewhat unexpected. On the Typicality hypothesis, one would expect that words with free bases should be less likely to be misspelled if the standard spelling is <able> vs. <ible>, but more likely to be misspelled if the standard spelling is <ible>. More generally, one would expect interactions of Suffix with other predictors in our model of *-ible/-able*. That was not the case, meaning that there was no support for the Typicality hypothesis in our models.

On the Segmentability hypothesis, on the other hand, there should be fewer misspellings with increasing strength of morphological boundaries, regardless of the target suffix. There was partial support for this hypothesis: Higher Base frequency and Base type (free as against bound roots) were associated with decreased probability of attested spelling variants for *-ible/-able* and *-ment*, respectively. In the case of *-ance/-ence*, we did not observe effects of Base frequency or Base type. The fact that the dataset for *-ance/-ence* was smaller than for *-ible/-able*, may explain the absence of significant effects, but there are also several other complicating factors.

Before considering these additional factors more closely, we note that the Segmentability hypothesis is consistent with evidence from several strands of previous research: For example, the ability to identify derivational morphemes is associated with better spelling performance in children (see e.g. Carlisle, 1988; Singson, Mahony, & Mann, 2000), as well as high school and college-age students (Mahony, 1994). In addition, complex words have sometimes been found to be better preserved than morphologically simple ones in individuals with neuropsychological impairments (Rapp & Fischer-Baum, 2014). Badecker, Hillis, and Caramazza (1990), for example, describe the case of an individual with dysgraphia who was more successful at producing word-final letters immediately preceded by a morpheme boundary than those not immediately preceded by a morpheme boundary. These observations suggest that, sometimes, morphologically complex

words have a processing advantage over morphologically simple ones. That processing advantage may in turn help explain the 'spelling advantage' of complex words, i.e. the segmentability effect.

On the other hand, there is also evidence that the presence of morpheme boundaries may make spelling errors more, not less, likely in some cases: In a study of a frequent spelling error in Hebrew, involving the insertion of a character representing a vowel in certain types of nouns, Bar-On and Kuperman (2019) found the locus of the insertion to be sensitive to morphological structure. Among other patterns, Bar-On and Kuperman (2019) found insertions to have a strong tendency to occur immediately preceding a suffix. While this result, like the previous ones, shows that morphology influences spelling variation, the presence of a morphological boundary in the Hebrew case was associated with an increased chance of error, unlike what we saw in the present study. It is difficult to know whether this difference is due to the non-concatenative nature of Hebrew morphology, different phoneme-grapheme mapping for vowels vs. consonants, or some other difference either in the structure of Hebrew vs. German and English, or in the tasks and methods employed.

Our results might also appear to run counter to another published finding. Schmitz et al. (2018, p. 111) report that there were "fewer errors for more frequent word forms" in a corpus of 17,432 tweets containing 1,185 misspelled forms. However, the frequency in question, according to the discussion of the regression models in Schmitz et al. (2018), pertains to the relative frequency of two homophonous forms, not of the absolute frequency of either form.

The observed direction of the effects of segmentability on spelling is far from inevitable, even in a language like English. We also considered the alternative possibility that high segmentability should be associated with increased spelling difficulty, due to a paradigmatic consequence of segmentability: Recognizing the morpheme boundary in a word like *available* makes that word both easier and more difficult to spell. It makes it easier in that it privileges two options (<available> and <availible>) from the much larger set of possibilities that includes <availabble>, <avaible>, and <availbel>. On the other hand, writers must now make a choice between <able> and <ible>, both of which represent common affixes. Paradigmatic competition has been demonstrated to affect pronunciation variation (Cohen, 2014; Kuperman, Pluymaekers, Ernestus, & Baayen, 2007), and it seems plausible that it might also affect spelling variation. Competition might be particularly strong in highly segmentable words and might make such words difficult spelling targets. That is the opposite of what we observed here. However, competition as discussed in (Cohen, 2014; Kuperman et al., 2007) depends on several other factors (such as morphological family size). We consider the relationship between segmentability and competition to be an avenue worth exploring in future research.

Several patterns in our models, and several other variables that may have affected our results, merit closer inspection. In particular, we discuss here the effects of Base complexity, and Target bigram, i.e. the probability of the initial letter of the target suffix, given the final letter of the base, a variable that we believe reflects several distinct properties of letters, words, and sounds.

Base complexity

We suspected that morphological complexity of the base might affect spelling behavior, such that forms in which the target suffix follows a morphologically complex form might be easier to spell than forms in which the suffix attaches to a monomorphemic base. It will be observed that we are using the term 'base' somewhat loosely here: We suspected that complexity would play a role even in words like *un-seasonable* or *un-sinkable*, i.e. words in which the material preceding the suffix would not be considered the morphological base of the target suffix. There was partial support for this idea. There was no effect of base complexity for *-ence*, *-ance* or *-ment*, possibly because Base complexity is not independent of other variables (a fact that informed variable selection for each set of models): For *-ment*, for example, only six target words with bound roots contained complex bases. In the model of *-ible* and *-able*, however, complex bases were associated with fewer errors. Recall that *-ible* vs. *-able* did not differ from one another with respect to base complexity. Therefore, it appears unlikely that the effect of Base complexity was actually an effect of Suffix in disguise. Instead, we believe that base complexity promotes segmentability.

Why should base complexity be associated with higher segmentability? Recall that in multiply affixed words weak morphological boundaries tend to occur inside of strong boundaries (Hay & Baayen, 2002; Hay & Plag, 2004; Plag & Baayen, 2009; Zirkel, 2010). Our target words fall into two classes: Those in which another suffix precedes the target suffix (e.g. *real-ize-able*, *diagonal-ize-able*, *classify-able*) and those in which the target suffix is the only suffix, but which contain prefixes. In the former case, our target suffix has a stronger boundary than the preceding suffix. We are not aware of any studies that have tested the segmentability of prefix-suffix combinations. A priori, however, we note that there are two possible bracketings [Prefix-Base]-Suffix, and Prefix-[Base-Suffix] (though in many words, the bracketing is ambiguous). Given the parse [Prefix-Base]-Suffix, the same reasoning applies as with doubly suffixed bases: the target suffix has the strongest boundary that is present in the word. Only with the parse Prefix-[Base-Suffix] would the target suffix have a weaker boundary than the other affix present in the word. Thus, based on considerations of complexity-based ordering, in two out of three affix-configurations we would expect an enhanced tendency for segmentation for the target suffix.

To our knowledge, previous literature has been silent on effects of base complexity on behavioral measures such as lexical decision or reading times: There is a copious literature on affix ordering and other combinatorial properties of affixes, but far less information seems to be available on the effects of multiple affixation on recognition, reading, or writing. Processing effects of morphological boundaries have so far primarily been studied in words containing only one derivational affix. We believe that the effect of Base complexity underscores the need to study how multiply affixed words are processed.

Variant bigram probability

Turning now to the variant bigram probability, we take the effects of this variable to reflect at least three sets of factors: The first is that high-bigram-probability errors may ‘look right’, making errors harder to detect. The second is that the process of typing (or thumbing) may be routinized to a higher degree for high-probability bigrams than low-probability ones, making errors harder to avoid. The third set of factors concerns pronunciation: Some spelling variants, e.g. <legable>, <revocible>, and <diligance>, invite pronunciations that differ from those of their intended targets, e.g. <legible>, <revocable>, and <diligence>. In fact, we believe that many misspellings represent what one might term ‘pronunciation spellings’, a converse of ‘spelling pronunciations’. While the latter is an accepted term for non-standard pronunciations based on standard spellings (e.g. [marzld] for <misled>), ‘pronunciation spellings’ are non-standard spellings based on (standard or non-standard) pronunciation. We avoid the more familiar term ‘phonetic spelling’ here, as that term is typically applied in discussion of learning and development. The Tweets analyzed here do not generally give the impression that their authors were in the process of learning to write – or to spell, for that matter, despite the occasional non-standard spelling. The properties of letter combinations like <gi>, <ge> or <ci> vs. <ga> or <ca>, and their different pronunciations, serve as a reminder that the bigram probabilities in our data are not truly gradient – particularly because the set of base-final letters that occur with a given suffix is quite restricted in some cases. Before considering this point further, we wish to draw attention to a related issue, concerning the bigrams present in the standard spelling.

Target bigram probability

It is tempting to think that bigram probability of the target may index boundary strength associated with our target suffixes, consistent with previous work on transitional probabilities in speech perception (Hay, 2002, and references

therein), but we do not believe that the effects we observed should be attributed directly to these distributional properties. Instead, effects of target bigram probability likely arise for reasons that are analogous to those mentioned in connection with variant bigram probabilities. An additional complicating factor in the interpretation of target bigram probability is that the range of letters that may precede the standard spelling of a given suffix may be quite restricted: For example, <ible> only follows 9 distinct letters in our dataset, whereas <able> follows 22 distinct letters. We leave it to future research to determine the extent to which such regularities affect spelling in cases where a writer is uncertain of the standard spelling.

Setting aside the underlying mechanisms, the effect of Variant bigram may informally be described as reflecting whether spelling variants ‘look wrong.’ The question then arises how the variants compare to their orthographically correct cousins in this regard: If standard and variant both ‘look right’ (or wrong), that fact might increase spelling uncertainty. Put differently, the question is whether using a non-standard spelling, makes things better or worse. To address that question, we added Target bigram as a predictor to the final model of *-ible*, *-able* spelling, along with an interaction of Target bigram and Variant bigram. The strong correlation between the two bigram variables (Spearman’s $\rho = 0.31$, $p < .001$) means that the model estimates should be taken with some caution. We nevertheless included this model, as a preliminary check of the possibility just mentioned, of the effect of variant bigram being modulated by that of the target bigram.

Table 14. Logistic regression model of *-ible*/*-able* misspellings, taking into account target and variant bigram probabilities

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-1.3510	0.1941	-6.96	.0000
Suffix	3.3789	0.4462	7.57	.0000
Base complexity	-0.7832	0.2571	-3.05	.0023
Target bigram	-3.0493	2.8305	-1.08	.2813
Variant bigram	6.6203	2.3937	2.77	.0057
Target frequency	0.6806	0.0781	8.71	.0000
Base frequency	-0.1767	0.0502	-3.52	.0004
Target bigram:Variant bigram	-184.5247	56.8764	-3.24	.0012
Target frequency:Base frequency	0.0819	0.0263	3.12	.018

We fitted the model following the same procedure as before, i.e. starting with a model containing all predictors and using stepwise backward elimination. The resulting model is summarized in Table 14. Collinearity within the model appeared to be acceptably low: The highest (generalized) variance inflation factor,

of 1.96, was associated with the variant bigram probability. There was a significant interaction of the two bigram probabilities ($\beta = -184.525$, $p = 0.001$). The interaction is visualized in Figure 2 for three sections of target bigram probability. As can be seen in the plot, the positive effect of variant bigram probability was strongest for words with low target bigram probabilities, attenuated for words with medium-range target bigram probabilities, and possibly reversed for high target bigram probabilities; the large confidence interval in the highest frequency range renders that last point inconclusive. This pattern suggests that high variant bigram probability can interfere with correct spelling unless high target bigram probability is also high. Stated informally, when standard spellings ‘look right’ to begin with, writers are less likely to deviate from the standard spelling. We further asked whether low (target) bigram probabilities tended to occur in infrequent words; if so, then the vulnerable state of low-probability target spellings might be a word frequency effect in disguise. That was not the case (Spearman’s $\rho = -0.01$, n.s.). We refrained from entering both bigram probabilities into a model of *-ance/-ence*: Given the small number of word types for each bigram, such a model would almost certainly be overfitted.

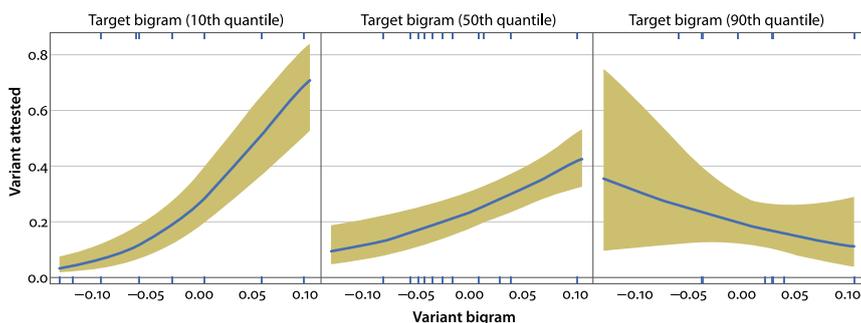


Figure 2. Effect of variant bigram probability (see text) on spelling variation in *-ible/-able* words for three target bigram probability bands, from lowest (leftmost panel) to highest (rightmost)

The role of boundary strength in spelling variation

Our general hypothesis was that not only the presence, but also the strength of morphological boundaries would affect misspellings. One piece of evidence for this that we have not yet discussed concerns the differences among the spelling targets considered here: Not only are the suffixes *-ible* and *-able* difficult spelling targets, they also differ from one another in difficulty, unlike *-ence/-ance*. The direction of difference is consistent with the notion that stronger morphological boundaries facilitate standard spellings. By contrast, we did not find any evidence

for suffix-specific effects in words with *-ence* vs. *-ance*, which are also difficult spelling targets, but do not differ in boundary strength. We interpret this pattern as an indication that the differences in the behavior of *-able/ible* vs. *-ence/-ance* are indeed related to the difference in boundary strength in *-able/ible*, and the absence of such a difference in *-ence/-ance*. It remains to be seen whether this interpretation is correct: Our set of words with *-ance/-ence* was far smaller than the set of words with *ible/able* ($n = 118$ vs. $n = 596$, respectively), which may explain the absence of significant effects. The marginally significant effect of Base Type on the spelling of *-ment* may be another instance of an effect of boundary strength.

Our findings tie in with several strands of previous research on boundary strength. Nottbusch, Grimm, Weingarten, and Will (2005); Sahel et al. (2008) and Weingarten, Nottbusch, and Will (2004), for example, found that inter-keystroke intervals in typed productions of German noun compounds reflected lexical frequency, head frequency, semantic transparency, and relative (base-to-derived) frequency, i.e. variables that are associated with gradient morphological boundaries. The more general finding of misspellings reflecting morphological properties of words meshes well with research on inflectional affixes, specifically the work of Sandra (2010); Sandra and Fayol (2003); Sandra, Frisson, and Daems (1999, 2004); Schmitz et al. (2018) on patterns of misspellings in homophonous Dutch inflectional suffixes.

Limitations

We wished to focus specifically on effects of segmentability on legitimate suffixes, i.e. spellings representing suffixes in standard orthography. A corollary of our hypothesis, which we were unable to explore here, was that the lower segmentability of words in *-ible* vs. *-able* should make non-suffixal spellings of *-ible* words (e.g. <legibbel> or <plauzebble>) more likely than those of *-able* words (e.g. <washebble> or <portebell>). Broadening the current line of investigation to other spelling variants may also help clarify the extent to which misspellings of suffixed words are due to morphological structure vs. factors like keyboard layout or screen responsiveness. The current study presents a case study, comparing a single pair of suffixes differing in boundary strength to a single other pair that does not, and to a single other suffix without a competitor. To properly evaluate the hypotheses we considered, the investigation has to extend to more affixes.

Another set of limitations has to do with the way Tweets are produced. The main advantage of Twitter as a source of data lies in the diversity of topics covered and in the fact that tweets are not generally subject to editorial review. However, Twitter data have several drawbacks: For example, some Twitter users

may be using auto-complete editors or spell checkers, both of which may filter out many patterns one might observe in uncorrected spelling behavior. Secondly, the mechanics of typing or swiping words differ for different input devices, meaning that different letters are adjacent to one another or conveniently reached on keyboards. Thirdly, Twitter users include native speakers of languages such as French or German, where cognates of English *-ible/-able* are phonetically clearly distinct from one another, considerably reducing the risk of substituting their written forms for one another.

Conclusion

Misspellings might not seem to be of particular interest to research on the mental lexicon because many orthographic errors are no doubt attributable to the physical environments of typing and hand-writing, such as keyboard layout, touch-screen responsiveness, tactile properties of keyboards, touchscreens, and pens (Crump & Logan, 2010; Deorowicz & Ciura, 2005). However, there is a substantial body of research demonstrating effects of typing and handwriting on stages of language production that precede planning and execution of motor movements (Delattre, Bonin, & Barry, 2006; Lambert, Kandel, Fayol, & Espéret, 2008; Roux, McKeef, Grosjacques, Afonso, & Kandel, 2013; Scaltritti, Arfé, Torrance, & Peressotti, 2016). Morphological boundaries and syllable boundaries have been the focus of several studies demonstrating that such boundaries affect hand writing and typing (Baus, Strijkers, & Costa, 2013; Bertram, Tønnessen, Strömqvist, Hyönä, & Niemi, 2015; Nottbusch et al., 2005; Roux et al., 2013; Sahel et al., 2008; Weingarten et al., 2004). There is also previous research arguing for misspellings of inflected forms as reflecting morphological representations (e.g. Sandra & Fayol, 2003; Sandra et al., 2004). We have argued that misspellings in English derivational suffixes similarly reflect lexical structure, in addition to the mechanics of keyboards or pens, in much the same way that pronunciation variation reflects lexical processing along with articulatory and acoustic aspects of speech production. In light of previous findings on the effects of spelling variability (both orthographically licensed and nonstandard variation) on reading (Falkauskas & Kuperman, 2015; Rahmanian & Kuperman, 2019), the case studies we presented here underscore the value of misspellings as a tool for understanding the processes underlying both writing and reading.

Acknowledgements

We are grateful to Victor Kuperman and Sandra Dominiek for their insightful, stimulating, and constructive comments, as well as to audiences at Berkeley, and at the Spoken Morphology workshops 2016 and 2017 (DFG Research Unit FOR2373). We are very grateful to the Deutsche Forschungsgemeinschaft for funding parts of this research (Grants PL151/8-1 and PL151/8-2 'Morpho-phonetic Variation in English' and PL151/7-1 and PL151/7-2 'FOR 2737 Spoken Morphology: Central Project' awarded to Ingo Plag).

References

- Adams, V. (2001). *Complex words in English*. Longman.
- Assink, E. M. (1985). Assessing spelling strategies for the orthography of Dutch verbs. *British Journal of Psychology*, 76(3), 353–363. <https://doi.org/10.1111/j.2044-8295.1985.tb01958.x>
- Baayen, R. H. (2014). Experimental and psycholinguistic approaches to studying derivation. *Handbook of derivational morphology*, 95–117.
- Baayen, R. H., Feldman, L. B., & Schreuder, R. (2006). Morphological influences on the recognition of monosyllabic monomorphemic words. *Journal of Memory and Language*, 55(2), 290–313. <https://doi.org/10.1016/j.jml.2006.03.008>
- Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). The CELEX lexical database (release 2). *Distributed by the Linguistic Data Consortium, University of Pennsylvania*.
- Badecker, W., Hillis, A., & Caramazza, A. (1990). Lexical morphology and its role in the writing process: Evidence from a case of acquired dysgraphia. *Cognition*, 35(3), 205–243. [https://doi.org/10.1016/0010-0277\(90\)90023-D](https://doi.org/10.1016/0010-0277(90)90023-D)
- Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., ... Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods*, 39(3), 445–459. <https://doi.org/10.3758/BF03193014>
- Bar-On, A., & Kuperman, V. (2019). Spelling errors respect morphology: a corpus study of Hebrew orthography. *Reading and Writing*, 32.5, 1107–1128.
- Bauer, L., Lieber, R., & Plag, I. (2013). *The Oxford reference guide to English morphology*. Oxford University Press. Retrieved from <https://doi.org/10.1093/acprof:oso/9780198747062.001.0001>
- Baus, C., Strijkers, K., & Costa, A. (2013). When does word frequency influence written production? *Frontiers in Psychology*, 4. Retrieved from <https://doi.org/10.3389/fpsyg.2013.00963>
- Bertram, R., Tønnessen, F. E., Strömquist, S., Hyönä, J., & Niemi, P. (2015). Cascaded processing in written compound word production. *Frontiers in Human Neuroscience*, 9, 207.
- Bloomer, R. H. (1956). Word length and complexity variables in spelling difficulty. *The Journal of Educational Research*, 49 (7), 531–536. <https://doi.org/10.1080/00220671.1956.10882314>
- Blumenthal-Dramé, A., Glauche, V., Bormann, T., Weiller, C., Musso, M., & Kortmann, B. (2017). Frequency and chunking in derived words: a parametric fMRI study. *Journal of Cognitive Neuroscience*. https://doi.org/10.1162/jocn_a_01120

- Brysaert, M., & New, B. (2009). Moving beyond kučera and francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, 41 (4), 977–990. <https://doi.org/10.3758/BRM.41.4.977>
- Cahen, L. S., Craun, M. J., & Johnson, S. K. (1971). Spelling difficulty: A survey of the research. *Review of Educational Research*, 41 (4), 281–301.
- Caramazza, A., Miceli, G., Villa, G., & Romani, C. (1987). The role of the graphemic buffer in spelling: Evidence from a case of acquired dysgraphia. *Cognition*, 26 (1), 59–85. [https://doi.org/10.1016/0010-0277\(87\)90014-X](https://doi.org/10.1016/0010-0277(87)90014-X)
- Carlisle, J. F. (1988). Knowledge of derivational morphology and spelling ability in fourth, sixth, and eighth graders. *Applied Psycholinguistics*, 9 (3), 247–266. <https://doi.org/10.1017/S0142716400007839>
- Chomsky, N., & Halle, M. (1968). The sound pattern of English.
- Cohen, C. (2014). Probabilistic reduction and probabilistic enhancement. *Morphology*, 24 (4), 291–323. <https://doi.org/10.1007/s11525-014-9243-y>
- Crump, M. J. C., & Logan, G. D. (2010). Warning: This keyboard will deconstruct – the role of the keyboard in skilled typewriting. *Psychonomic Bulletin & Review*, 17 (3), 394–399. Retrieved from <https://doi.org/10.3758/PBR.17.3.394>
- Cutler, A. (2011). *Slips of the tongue and language production*. Walter de Gruyter.
- Davies, M. (2013). *The Corpus of Contemporary American English (full text on CD): 440 million words, 1990–2012*.
- Delattre, M., Bonin, P., & Barry, C. (2006). Written spelling to dictation: Sound-to-spelling regularity affects both writing latencies and durations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(6), 1330.
- Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological review*, 93(3), 283. <https://doi.org/10.1037/0033-295X.93.3.283>
- Deorowicz, S., & Ciura, M. G. (2005). Correcting spelling errors by modelling their causes. *International Journal of Applied Mathematics and Computer Science*, 15, 275–285.
- Dressler, W. (1985). *Morphonology*. Ann Arbor: Karoma.
- Falkauskas, K., & Kuperman, V. (2015). When experience meets language statistics: Individual variability in processing English compound words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 41(6), 1607.
- 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*, 47(2), 437–464. <https://doi.org/10.1080/14640749408401119>
- Gagné, C. L., & Spalding, T. L. (2014). Typing time as an index of morphological and semantic effects during English compound processing. *Lingue e linguaggio*, 13(2), 241–262.
- Gagné, C. L., & Spalding, T. L. (2016a). Effects of morphology and semantic transparency on typing latencies in English compound and pseudocompound words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 42(9), 1489.
- Gagné, C. L., & Spalding, T. L. (2016b). Written production of English compounds: effects of morphology and semantic transparency. *Morphology*, 26(2), 133–155. <https://doi.org/10.1007/s11525-015-9265-0>
- Gentry, J. (2015). twitter: R based twitter client [Computer software manual]. Retrieved from <https://CRAN.R-project.org/package=twitter> (R package version 1.1.9)

- Hay, J. (2001). Lexical frequency in morphology: is everything relative? *Linguistics*, 39 (6), 1041–1070. <https://doi.org/10.1515/ling.2001.041>
- Hay, J. (2002). From speech perception to morphology: Affix ordering revisited. *Language*, 78 (3), 527–555. <https://doi.org/10.1353/lan.2002.0159>
- Hay, J. (2003). *Causes and Consequences of Word Structure*. New York: Routledge. Retrieved from <https://doi.org/10.4324/9780203495131>
- Hay, J. (2007). The phonetics of ‘un’. *Lexical creativity, texts and contexts*, 39–57. <https://doi.org/10.1075/sfsl.58.09hay>
- Hay, J., & Baayen, H. (2002). Parsing and productivity. In *Yearbook of Morphology* (pp. 203–235). Springer Netherlands. https://doi.org/10.1007/978-94-017-3726-5_8
- Hay, J., & Baayen, H. (2005). Shifting paradigms: gradient structure in morphology. *Trends in Cognitive Sciences*, 9 (7), 342–348. <https://doi.org/10.1016/j.tics.2005.04.002>
- Hay, J., & Plag, I. (2004). What constrains possible suffix combinations? On the interaction of grammatical and processing restrictions in derivational morphology. *Natural Language & Linguistic Theory*, 22 (3), 565–596. <https://doi.org/10.1023/B:NALA.0000027679.63308.89>
- Kawaletz, L., & Plag, I. (2015). Predicting the semantics of English nominalizations: a frame-based analysis of *-ment* suffixation. In L. Bauer, P. Stekauer, & L. Kortvelyessy (Eds.), *Semantics of Complex Words* (pp. 289–319). Dordrecht: Springer.
- Kemps, R. J. J. K., Ernestus, M., Schreuder, R., & Baayen, R. H. (2005). Prosodic cues for morphological complexity: The case of Dutch plural nouns. *Memory & Cognition*, 33 (3), 430–446. <https://doi.org/10.3758/BF03193061>
- Kiparsky, P. (1982). Lexical morphology and phonology. In I.-S. Yang (Ed.), *Linguistics in the Morning Calm: Selected Papers from SICOL* (pp. 3–91). Seoul: Hanshin.
- Kuperman, V., & Bertram, R. (2013). Moving spaces: Spelling alternation in English noun-noun compounds. *Language and Cognitive Processes*, 28 (7), 939–966. Retrieved from <https://doi.org/10.1080/01690965.2012.701757>
- Kuperman, V., Pluymaekers, M., Ernestus, M., & Baayen, H. (2007). Morphological predictability and acoustic duration of interfixes in Dutch compounds. *The Journal of the Acoustical Society of America*, 121(4), 2261–2271. <https://doi.org/10.1121/1.2537393>
- Lambert, E., Kandel, S., Fayol, M., & Espéret, E. (2008). The effect of the number of syllables on handwriting production. *Reading and Writing*, 21(9), 859–883. <https://doi.org/10.1007/s11145-007-9095-5>
- Largy, P. (1996). The homophone effect in written French: The case of verb-noun inflection errors. *Language and cognitive processes*, 11(3), 217–256. <https://doi.org/10.1080/016909696387178>
- Lee-Kim, S.-I., Davidson, L., & Hwang, S. (2013). Morphological effects on the darkness of English intervocalic /l/. *Laboratory Phonology*, 4(2), 475–511. <https://doi.org/10.1515/lp-2013-0015>
- Libben, G., Jarema, G., & Luke, J. (May, 2018). *Same words, different languages: Examining English-French written word recognition and production*. *Annual Meeting of the Canadian Linguistics Association, Regina, Canada*. <https://brocku.ca/social-sciences/applied-linguistics/people/gary-libben/#publications072e-2dd8>
- Libben, G., & Weber, S. (2014). Semantic transparency, compounding, and the nature of independent variables. In F. Rainer, F. Gardani, H. C. Luschützky, & W. U. Dressler (Eds.), *Morphology and Meaning* (pp. 205–221). Amsterdam / Philadelphia: Benjamins. <https://doi.org/10.1075/cilt.327.14lib>

- Libben, G., Weber, S., & Miwa, K. (2012). P3: A technique for the study of perception, production, and participant properties. *The Mental Lexicon*, 7(2), 237–248. <https://doi.org/10.1075/ml.7.2.05lib>
- Mahony, D. L. (1994). Using sensitivity to word structure to explain variance in high school and college level reading ability. *Reading and Writing*, 6(1), 19–44. <https://doi.org/10.1007/BF01027276>
- Marchand, H. (1969). *The categories and types of present-day English word-formation* (2nd ed.). München: Verlag C. H. Beck.
- Nottbusch, G., Grimm, A., Weingarten, R., & Will, U. (2005). Syllabic structures in typing: Evidence from deaf writers. *Reading and Writing*, 18(6), 497–526. <https://doi.org/10.1007/s11145-005-3178-y>
- Plag, I. (2003). *Word-formation in English*. Cambridge University Press. Retrieved from <https://doi.org/10.1017/CBO9780511841323>
- Plag, I. (2014). Phonological and phonetic variability in complex words: An uncharted territory. *Italian Journal of Linguistics/Rivista di Linguistica*, 26(2), 209–228.
- Plag, I., & Baayen, R. H. (2009). Suffix ordering and morphological processing. *Language*, 85, 106–149.
- Plag, I., & Ben Hedia, S. (2018). The phonetics of newly derived words: Testing the effect of morphological segmentability on affix duration. In S. Arndt-Lappe, A. Braun, C. Moulin, & E. Winter-Froemel (Eds.), *Expanding the Lexicon: Linguistic Innovation, Morphological Productivity, and the Role of Discourse-related Factors* (pp. 93–116). Berlin, New York: de Gruyter Mouton. <https://doi.org/10.1515/9783110501933-095>
- R Development Core Team. (2008). *R: A language and environment for statistical computing [Computer software manual]*. Vienna, Austria. Retrieved from <https://www.R-project.org/> (ISBN 3-900051-07-0)
- Rahmanian, S., & Kuperman, V. (2019). Spelling errors impede recognition of correctly spelled word forms. *Scientific Studies of Reading*, 23 (1), 24–36.
- Rapp, B., & Fischer-Baum, S. (2014). Representation of orthographic knowledge. *The Oxford handbook of language production*, 338.
- Roux, S., McKeeff, T. J., Grosjacques, G., Afonso, O., & Kandel, S. (2013). The interaction between central and peripheral processes in handwriting production. *Cognition*, 127 (2), 235–241. <https://doi.org/10.1016/j.cognition.2012.12.009>
- Sahel, S., Nottbusch, G., Grimm, A., & Weingarten, R. (2008). Written production of German compounds: Effects of lexical frequency and semantic transparency. *Written Language & Literacy*, 11 (2), 211–227.
- 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., & Fayol, M. (2003). Spelling errors with a view on the mental lexicon: Frequency and proximity effects in misspelling homophonous regular verb forms in Dutch and French. *Trends in Linguistics Studies and Monographs*, 151, 485–514.
- 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...: Limited attentional resources and homophone frequency account for spelling errors on silent verb suffixes in Dutch. *Written Language & Literacy*, 7 (1), 61–77. <https://doi.org/10.1075/wll.7.1.07san>
- Scaltritti, M., Arfé, B., Torrance, M., & Peressotti, F. (2016). Typing pictures: Linguistic processing cascades into finger movements. *Cognition*, 156, 16–29. Retrieved from <https://doi.org/10.1016/j.cognition.2016.07.006>
- Schmitz, T., Chamalaun, R., & Ernestus, M. (2018). The Dutch verb-spelling paradox in social media. *Linguistics in the Netherlands*, 35 (1), 111–124.
- Seyfarth, S., Garellek, M., Gillingham, G., Ackerman, F., & Malouf, R. (2017). Acoustic differences in morphologically-distinct homophones. *Language, Cognition and Neuroscience*, 1–18.
- Siegel, D. (1979). *Topics in English morphology*. Garland.
- Singson, M., Mahony, D., & Mann, V. (2000). The relation between reading ability and morphological skills: Evidence from derivational suffixes. *Reading and Writing*, 12 (3), 219–252. <https://doi.org/10.1023/A:1008196330239>
- Smith, R., Baker, R., & Hawkins, S. (2012). Phonetic detail that distinguishes prefixed from pseudo-prefixed words. *Journal of Phonetics*, 40 (5), 689–705. Retrieved from {<http://www.sciencedirect.com/science/article/pii/S0095447012000356>}. <https://doi.org/10.1016/j.wocn.2012.04.002>
- Solso, R. L., & Juel, C. L. (1980). Positional frequency and versatility of bigrams for two-through nine-letter English words. *Behavior Research Methods*, 12 (3), 297–343. <https://doi.org/10.3758/BF03201669>
- Spencer, K. (2007). Predicting children's word-spelling difficulty for common English words from measures of orthographic transparency, phonemic and graphemic length and word frequency. *British Journal of Psychology*, 98(2), 305–338. <https://doi.org/10.1348/000712606X123002>
- Sproat, R., & Fujimura, O. (1993). Allophonic variation in English /l/ and its implications for phonetic implementation. *Journal of Phonetics*, 21, 291–311.
- Twitter. (2006). *Twitter*. Retrieved from {<https://twitter.com>}
- Vannest, J., Newport, E. L., Newman, A. J., & Bavelier, D. (2011). Interplay between morphology and frequency in lexical access: The case of the base frequency effect. *Brain Research*, 1373, 144–159. <https://doi.org/10.1016/j.brainres.2010.12.022>
- Weingarten, R., Nottbusch, G., & Will, U. (2004). Morphemes, syllables and graphemes in written word production. In T. Pechmann & C. Habel (Eds.), *Multidisciplinary approaches to language production* (pp. 529–572). Mouton de Gruyter. <https://doi.org/10.1515/9783110894028.529>
- Wikipedia. (2017). *Wikipedia:lists of common misspellings — Wikipedia, the free encyclopedia*. Retrieved from {https://en.wikipedia.org/wiki/Wikipedia:Lists_of_common_misspellings} ([Online; accessed 04 September 2017])
- Zirkel, L. (2010). Prefix combinations in English: Structural and processing factors. *Morphology*, 20(1), 239–266. <https://doi.org/10.1007/s11525-010-9151-8>

Address for correspondence

Susanne Gahl
Department of Linguistics
University of California at Berkeley
1203 Dwinelle Hall
Berkeley, CA 94720-2650
USA
gahl@berkeley.edu

Co-author information

Ingo Plag
Department of English and American Studies
Heinrich-Heine Universität Düsseldorf
ingo.plag@uni-duesseldorf.de