ReCALL (2023), 35: 3, 321–338
doi:10.1017/S0958344023000150

RESEARCH ARTICLE

Building a pedagogic spellchecker for L2 learners of Spanish

Miguel Blázquez-Carretero
University of the Philippines Diliman, Philippines (mblazquez@up.edu.ph)

Abstract

In 2016, Lawley proposed an easy-to-build spellchecker specifically designed to help second language (L2) learners in their writing process by facilitating self-correction. The aim was to overcome the disadvantages to L2 learners posed by generic spellcheckers (GSC), such as that embedded in Microsoft Word. Drawbacks include autocorrection, misdiagnoses, and overlooked errors. With the aim of imparting explicit L2 spelling knowledge, this correcting tool does not merely suggest possible alternatives to the detected error but also provides explanations of any relevant spelling patterns. Following Lawley's (2016) recommendations, the present study developed a prototype computer-based pedagogic spellchecker (PSC) to aid L2 learners in self-correcting their written production in Spanish. First, a corpus was used to identify frequent spelling errors of Spanish as a foreign language (SFL) learners. Handcrafted feedback was then designed to tackle the commonest misspellings. To subsequently evaluate this PSC's efficacy in error detection and correction, another learner Spanish corpus was used. Sixty compositions were analysed to determine the PSC's capacity for error recognition and feedback provision in comparison with that of a GSC. Results indicate that the PSC detected over 90% of the misspellings, significantly outperforming the GSC in error detection. Both provided adequate feedback on two out of three detected errors, but the pedagogic nature of the former has the added advantage of facilitating self-learning (Blázquez-Carretero & Woore, 2021). These findings suggest that it is feasible to develop spellcheckers that provide synchronous feedback, allowing SFL learners to confidently self-correct their writing while saving time and effort on the teacher's part.

Keywords: written corrective feedback; automated written evaluation; computer-generated feedback; self-correction; spelling; Spanish as a foreign language

1. Introduction

In the field of second language (L2) acquisition, there is an ongoing shift in perception towards written production. No longer a mere consequence of language learning, writing is now deemed a tool to facilitate the learning process itself. For instance, errors in written production are considered as additional learning opportunities, with research suggesting that written corrective feedback (WCF) not only targets linguistic accuracy but also promotes L2 writing acquisition (Manchón & Vasylets, 2019). Given this shift, WCF emerges as a catalyst for L2 writing acquisition.

Spelling accurately is among the key linguistic competencies that facilitate and indicate L2 writing progress. It influences vocabulary acquisition and retention (Brown & Ellis, 1994), and is important for reading and writing (Graham & Santangelo, 2014). It is also a predominant factor in evaluating L2 writing (Rodríguez, Villoria & Paredes, 1997). Despite its relevance, spelling is often

Cite this article: Blázquez-Carretero, M. (2023). Building a pedagogic spellchecker for L2 learners of Spanish. *ReCALL* 35(3): 321–338. https://doi.org/10.1017/S0958344023000150

© The Author(s), 2023. Published by Cambridge University Press on behalf of EUROCALL, the European Association for Computer-Assisted Language Learning.

https://doi.org/10.1017/S0958344023000150 Published online by Cambridge University Press

neglected in Spanish as a foreign language (SFL) pedagogy (Sánchez Jiménez, 2009), probably because Spanish orthography, unlike opaque orthographies like English, has a relatively high degree of grapheme–phoneme correspondence. Put simply, Spanish spelling can largely be predicted from pronunciation.

A commonplace resource for L2 teachers to explicitly target orthography is WCF, which appears to have a "substantive effect on L2 written accuracy" in immediate post-tests with a small to moderate effect size (Kang & Han, 2015: 10), independently of the WCF's nature and characteristics (see Blázquez-Carretero, 2022, for an extensive review of WCF literature). Kang and Han's (2015) meta-analysis of experimental studies also provided evidence of improvement, even in delayed post-tests. Research therefore recommends that L2 learners write as much as possible to learn from their errors. Yet resource and time constraints limit the teacher's opportunity to provide feedback (Lawley, 2016). With narrow time frames, students are encouraged to self- and peer-correct, usually with the guidance of teachers who provide detailed instructions. Despite a wealth of literature supporting the efficacy of autocorrection (Lázaro Ibarrola, 2013), students often find the process difficult and tedious without professional assistance (Lawley, 2015). Lee (1997) notes that recognition of misspellings is crucial to self-correction but proves challenging for L2 learners if their mistakes are not previously highlighted. Heeding Lee and Lázaro Ibarrola's findings, an aptly programmed software capable of detecting spelling errors would likely help L2 students to independently correct such errors. This is backed by empirical evidence on the positive effects for L2 learning of computer-generated WCF (Heift & Hegelheimer, 2017; Li, Feng & Saricaoglu, 2017; Stevenson & Phakiti, 2014). While frequently used generic spellcheckers (GSCs) like those embedded in Microsoft Word (MW) operate with an error-detecting software, these first language (L1)–oriented tools have been shown to disadvantage L2 learners (Blázquez-Carretero & Fan, 2019; Blázquez-Carretero & Woore, 2021; Heift & Rimrott, 2008).

2. Limitations of existing spellcheckers for SFL learners

Most GSCs are built for native speakers. Given their main objective of facilitating and accelerating the user's writing process, misspellings are automatically corrected. This is beneficial for proficient writers whose errors are mostly just slips (e.g. typos), otherwise known as performance errors (Brown, 1994). For L2 students, whose errors mostly result from actual gaps in their target language knowledge (competence errors [Touchie, 1986]), autocorrection is not as useful. Indeed, the noticing hypothesis posits that "input does not become intake for language learning unless it is noticed, that is, consciously registered" (Schmidt, 2010: 721). Therefore, GSCs arguably impede learners from benefitting from their errors; the software instantly modifies mistakes, hence curtailing learners' conscious processing. Moreover, these mainstream spellcheckers suggest replacement words upon error detection. While a native speaker may effortlessly select the correct alternative, an L2 learner might find the task more challenging and less intuitive. There is sufficient evidence that L2 students are misled by these suggestions because they often presume the correct word to be listed, although this is not necessarily the case (Blázquez-Carretero & Fan, 2019; Heift & Rimrott, 2008; Mitton & Okada, 2007). Apart from the quality of feedback, the capacity for error detection is equally important. Spellcheckers generally use a reference corpus to assess whether words are written correctly. Words absent from the corpus are highlighted, followed by either a list of alternatives or a correction executed by the software (Mitton, 2010). While this method's detection rate is higher than 80% (Bestgen & Granger, 2011; Blázquez-Carretero & Fan, 2019; Rimrott & Heift, 2008), it only pertains to single-word errors, preventing GSCs from identifying context-specific mistakes should the misspelling correspond to an existent word (Blázquez-Carretero & Fan, 2019). As GSCs are L1-oriented, their built-in autocorrect and feedback mechanism is grounded on the notion that spelling errors are performance-based and

https://doi.org/10.1017/S0958344023000150 Published online by Cambridge University Press

typically involve single-letter violations. But L2 misspellings are more complex in nature (Heift & Schulze, 2007), generally requiring two or more operations to correct the error (i.e. edit distance; Ranalli & Yamashita, 2022).

Consequently, researchers have developed systems better attuned to L2 misspellings by drawing from a database of their errors (Rimrott & Heift, 2006) or from contextual information derived from L2 learner corpora (Nagata, Takamura & Neubig, 2017). A host of innovative and accessible web browser–integrated tools (e.g. Grammarly) incorporate advancements in machine learning and natural language processing technology (Ranalli, 2018). However, not many are specifically designed for SFL learners, and those available show similar limitations as GSCs. Indeed, Blázquez-Carretero and Fan's (2019) comparison of three different SFL spellcheckers (LanguageTool, Stilus, and SpanishChecker) suggested global gains in error detection (all detected over 85% of the errors) but observed less success in adequate feedback provision (with MW's 67% accuracy rate outperforming the rest).

Responding to these issues and limitations and drawing from related literature (Blázquez-Carretero, 2019; Blázquez-Carretero & Fan, 2019; Blázquez-Carretero & Woore, 2021; Harvey-Scholes, 2018; Lawley, 2015, 2016), a team of researchers from the *Universidad Nacional de Educación a Distancia* decided to build a prototype spellchecker to help SFL students in self-correcting their written outputs. This tool is specially designed to tackle common SFL misspellings through WCF. Blázquez-Carretero and Woore (2021) empirically demonstrated that SFL learners can benefit from error-specific, metalinguistic WCF targeting an underlying misconception.

The present study deals with the development of an online pedagogic spellchecker (PSC) able to detect and provide WCF to most SFL spelling errors by largely replicating Lawley's (2016) study, adapting it to the SFL context, and improving upon Lawley's methodological gaps. First, the SFL learners' interlanguage was accounted for through real data analysis within the framework of error analysis and corpus linguistics. When all misspellings in the CORANE corpus (Cestero Mancera & Penadés Martínez, 2009) were identified and scrutinised, a selection process was implemented to overcome Lawley's (2016) less meticulous method of identifying errors he deemed worthy of WCF. A pedagogical WCF was carefully designed for each error selected, including potential violations of the same orthographic pattern despite non-occurrence in the corpus. Another corpus, CEDEL2 (Lozano, 2022), was subsequently used to compare the PSC's efficacy in error recognition and feedback provision vis-à-vis MW's spellchecker.[1] This pedagogic tool was tested on a much larger sample for increased reliability. To avoid biased results, one type of filter was exclusively compared with the other (i.e. MW spelling filter vs. PSC spelling filter), unlike Lawley (2016), who drew favourable results on the PSC's error-detection accuracy after comparing its combined bigram and spelling filters with MW's spelling filter alone.

3. Building a spellchecker for SFL learners

Following Lawley's (2016) recommendations, the PSC contains a database of the commonest SFL misspellings, each supplied with its own pedagogic WCF. Equipped with a search-and-match facility, the software highlights the error and then provides corresponding feedback after the erroneous word is clicked on.

3.1 Identifying common spelling errors made by SFL learners

Real data from the CORANE corpus (Cestero Mancera & Penadés Martínez, 2009) was used to identify common SFL spelling errors to build a sound database and design-appropriate feedback. The corpus comprises 957 compositions totalling 219,032 words, written by 290 SFL learners from 22 different L1 backgrounds whose Spanish proficiency ranges between A2 and C1 levels (CEFR;

[1]The 2019 version was used in this study.

https://doi.org/10.1017/S0958344023000150 Published online by Cambridge University Press

Council of Europe, 2001). At this stage, spelling error refers to any string of letters between spaces that is excluded from the *Diccionario de la Lengua Española* (Real Academia Española, 2014). However, grammatical errors like gender and/or number disagreement, lexical errors like the misuse of diacritic marks in homographs, and punctuation errors were not considered as misspellings; such errors can be better identified by analysing their surrounding linguistic context (see Blázquez-Carretero, 2019). This definition was deemed necessary to suit the research objectives as we neither aimed to categorise nor explore the origin of errors, but rather to understand and explain how spellcheckers can effectively detect errors and offer appropriate feedback, thereby paving the way to the wholesale improvement of L2-oriented spellcheckers.

Analysing CORANE (Cestero Mancera & Penadés Martínez, 2009) yielded an initial list of 8,053 misspellings (3.68%).[2] An overwhelming majority (89.8%) contained a single error, that is, requiring only one "edit movement" (omission, substitution, incorporation, or transposition) for correction. The rest (10.2%) contained a combination of at least two errors within a word. Everything considered, the misspellings present in the corpus number 9,110.

3.2 Developing the pedagogic feedback

Given the profusion of spelling errors (9,110) and that a single-occurrence error does not necessarily constitute a recurrent problem, a methodological decision was taken to sift through the data and select only those resulting from violation of Spanish orthographic norms and patterns. Considering the lacuna in applied linguistics of a criterion for inclusion and exclusion of spelling errors, we established our own based on previous experiences in other domains of knowledge, specifically genetics (see Saint Pierre & Génin, 2014). This is a novel methodological addition that was noticeably missing from Lawley (2016). As Figure 1 shows, performance errors were discarded. First, a threshold was established at the individual level: a recurring orthographic deviation was deemed a competence error (i.e. if the same error appeared in half of the instances where the SFL learner intended to write that word, $f1 \geq 0.5$). More challenging was categorising errors within a word used only once ($f1 = 1$) by a given SFL learner, since it is impossible to determine at the individual level whether such deviations result from a performance or a competence error. However, different SFL learners committing the same misspelling prompted us to tag this word as potentially problematic. Thus, a second threshold was established at the corpus level: infrequent errors were discarded (i.e. at least two different individuals should have committed the same misspelling for it to be considered a competence error).[3] However, upon Lawley's (2016) recommendation, all errors violating an orthographic regularity, regardless of the instances of occurrence, were included in the feedback and the database alike.

An initial database of 4,379 spelling errors was compiled. A pedagogic WCF was designed to address each, grounded on Long's (1996) interaction hypothesis (i.e. that L2 acquisition requires exposure to input, meaning negotiation through interaction, feedback reception, and output production), its potential application to computer-based language learning environments (Grgurović, Chapelle & Shelley, 2013), and the empirically proven efficacy of explicit feedback as a correction method in virtual environments (Blázquez-Carretero & Woore, 2021; Heift, 2008; Lin, Liu & Paas, 2017). Following Cotos's (2011) criteria and Lawley's (2016) recommendations, the feedback must (a) be specific for each error, (b) indicate the error type, (c) explain the error clearly

[2]Two reasons account for this lower error-per-word rate versus previous studies (e.g. 10.5% in Sánchez Jiménez, 2010, and 8.64% in Blázquez-Carretero & Fan, 2019); most texts comprising this corpus were written at home with supporting materials, with half of the words (107,017) written by C1-level SFL learners who are supposedly good spellers (Council of Europe, 2001).

[3]In genetics, an allele is considered a rare variant if it appears in less than 1% of the population (Frazer, Murray, Schork & Topol, 2009). In our case, due to the small sample size (CORANE's 290 individual participants), the relative frequency of exclusion was reduced to 0.7% (i.e. two individuals had to have committed the same error).

https://doi.org/10.1017/S0958344023000150 Published online by Cambridge University Press

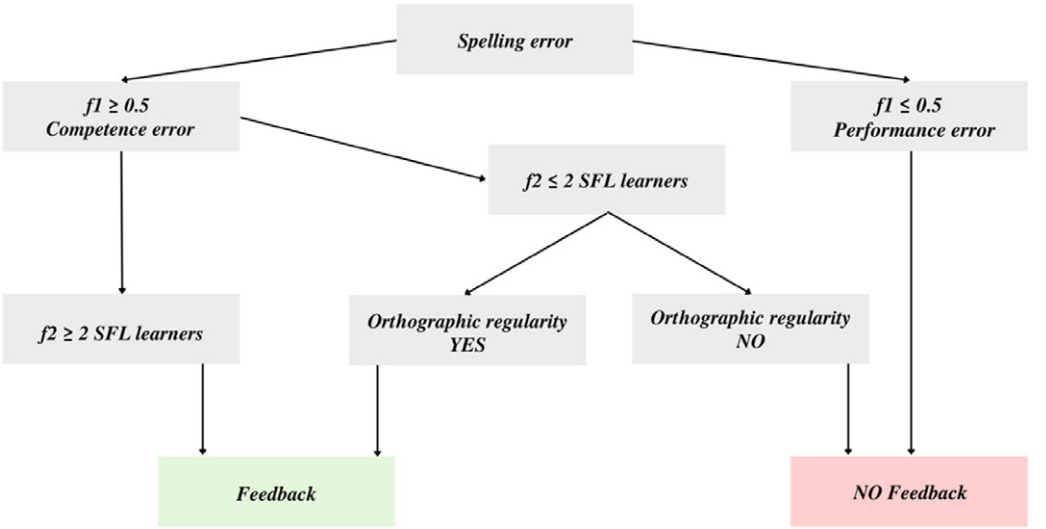


Figure 1. Error selection process

and concisely, and (d) encourage self-correction. Below is the feedback prepared for the error **psicologo* instead of *psicólogo* (psychologist).[4]

> *psicologo*
> The correct spelling is *psic**ó**logo*.
> In Spanish, all words ending in *-ologo have an accent on the first* ***o*** → ***ó****logo*.
> *monólogo (monologue)*
> *sociólogo (sociologist)*
> *arqueólogo (archaeologist)*

The composition of a PSC-generated WCF generally includes

1. the misspelling (**psicologo*), so that SFL learners may discern their interlanguage during the process of interaction and negotiation of meaning (Grgurović *et al*., 2013).
2. the correct spelling with the highlighted error in bold (*psicólogo*) for SFL learners to notice their error (Lee, 1997; Schmidt, 1990) and benefit from a direct WCF (van Beuningen *et al*., 2008).
3. a brief explanation of the likely reason behind the error and its solution. (In Spanish, all words ending in *-ologo* have an accent on the first *o* → *ólogo*.) Kang and Han's (2015) meta-analysis reveals that explicit WCF promotes writing accuracy and L2 learning. Concision is important to avoid saturating the learner's working memory (Kellogg, 2001; Manchón, 2014), and exceptions are excluded as they "make learning the rule more difficult" (Nation, 1990: 48). Exceptions usually refer to deviations from the indicated pattern, obscure words, or low-frequency terms, such as *homologo* ("I homologate"), *prologo* ("I preface"), and *monologo* ("I soliloquize"). Indeed, "exceptions should not be introduced until the rule has been learned" (Nation, 1990: 48).
4. examples of correctly spelled words illustrating the orthographic pattern (*monólogo*, *sociólogo*, *arqueólogo*). Indeed, among the psycholinguistic determinants of vocabulary learning is the recurrence of a letter sequence in a given language (Barcroft & Rott, 2010).

[4]The PSC gives WCF in Spanish. The English translation is provided in this article for linguistic consistency.

https://doi.org/10.1017/S0958344023000150 Published online by Cambridge University Press

Examples were selected based on variables facilitating vocabulary acquisition, like frequency of use (Ellis & Beaton, 1993), orthographic similarity to cognates in other languages (Kelley & Kohnert, 2012), and/or belonging to the same word family (Webb, 2021).

Among the commonest misspellings were accentuation errors (50.8% of errors detected in CORANE). Of these, 155 resulted from the lack of graphic accent in words ending in *-ión* (e.g. **avion* [plane] or **asociacion* [association]). Words ending in *-n*, *-s*, or a vowel require an orthographic accent if the stress is on the last syllable (the common explanation in the SFL classroom). However, this pattern is complex, necessitating further examples. A more visual, succinct WCF was provided to inhibit L2 working memory saturation (Kellogg, 2001; Manchón, 2014): "In Spanish, all words ending in *-ion* have an accent on the *o* → ***ión***". Indeed, it is a learning strategy already employed by SFL students (Blázquez-Carretero & Woore, 2023).

The common misspelling of *diferente* (**differente* appears seven times from five different learners) is likely linked to language interference (Figueredo, 2006). The corresponding WCF ("In Spanish, there are no words with *ff*") not only assists SFL learners on this occasion but also guides them in future instances involving English cognates like coffee (*café*) or office (*oficina*).

Other errors are morphological, as when SFL learners erroneously add affixes to root words (Marcos-Miguel, 2013; Sánchez-Gutiérrez, 2013). This, however, can easily be explained as an orthographic regularity, since "[. . .] there are clear rules indicating how suffixes are appended to words" (Goulandris, 1994: 416).

afortunadomente
The correct spelling is *afortunad**a**mente*.
In Spanish, adverbs ending in *-mente* are formed by adding the suffix to the feminine form of the adjective.

Masculine adjective	Feminine adjective	Adverb
rápido (fast)	*rápida*	*rápidamente*
completo (complete)	*completa*	*completamente*
lento (slow)	*lenta*	*lentamente*

Lexical consistency is another orthographic regularity that helps L2 students to learn and remember patterns (Goulandris, 1994). Hence, for errors like **discotequa* (misspelled *discoteca* "discotheque"), the PSC offers easy-to-remember etymological WCF: "In Spanish, *-teca* is a common suffix indicating some sort of storage, collection, or exhibition". Occasionally, visual clues aid L2 learners in retrieving the spelling of analogous words (Goswani, 1988). Indeed, Blázquez-Carretero and Woore (2021) found that SFL students adopt this learning strategy to correct their spelling. Sometimes, only examples of high-frequency words containing the letter sequence were offered:

dia
The correct spelling is *d**í**a*.
mía (mine)
fría (cold)
guía (guide)

Some misspellings, depending on the grapheme replaced, can lead to various possible replacements:

Major
¿Do you mean *mejor* o *mayor*?
For example:

https://doi.org/10.1017/S0958344023000150 Published online by Cambridge University Press

Juan juega ***mejor*** *al tenis que Luís.* (Juan plays tennis better than Luis.)
Si Juan tiene 20 años y Luís 18, Juan es ***mayor*** *que Luís.* (If Juan is 20 years old and Luis 18, Juan is older than Luis.)

Here, a GSC would have problems deciding which to propose: replacing the *a* with an *e* results in *mejor* (better) and the *j* with a *y* in *mayor* (older). Both are high-frequency terms in Spanish (*mejor* appears 80,404 times and *mayor* 87,561 in the CREA corpus [Real Academia Española, 2008]). Without context, it is impossible to determine the SFL learner's intention. Thus, the PSC's WCF does not merely offer an alternative, but also interacts with the learner and provides examples of context-based usage because "incidental learning via guessing from context is the most important of all sources of vocabulary learning" (Nation, 2001: 232).

The PSC can also detect and provide WCF on certain errors that a GSC might overlook as they are existing, albeit infrequent, words, like *ingles* (groins), *tenia* (tapeworm), or *ultimo* (I finalise). While CREA registers 203, 251, and 124 hits for each word respectively, most SFL learners originally intended the more widely used *inglés* (English), *tenía* (he/she had), and *último* (last) with 10,070, 66,996, and 45,718 occurrences respectively in the same corpus. Hence, a WCF was developed for all these terms, noting that one of either is used much more frequently.

Detecting the most frequent SFL misspellings is less challenging than the near impossible task of providing specific WCF for every single potential error. Therefore, the PSC works with two databases: one containing the most common SFL spelling errors and a reference corpus of 100 million correctly spelled words drawing from recent literature (e.g. stories, essays, articles, and news) (see San Mateo, 2016). A search-and-match detection system will then compare inputted text with both databases. If the misspelled word is a common error or follows a potentially problematic spelling pattern for SFL learners, a WCF is provided (e.g. **officina* with the inexistent double *f*). Otherwise, the PSC uses the reference corpus to verify the spelling and highlight the input should it be absent from the corpus (e.g. **ofizina*). In place of feedback, SFL learners are advised to consult a dictionary, a valuable self-correction strategy (Nesi, 2014).

However, this search-and-match system fails to identify errors where the input is a homophone or an existing lexical item, as in **tu hablas* (*your talk) (*tu* [your] and *tú* [you] are actual words). Evidently, SFL learners and L1 speakers alike commit such errors. For better service, the PSC was further equipped with a detection system able to spot errors in word pairs by analysing the frequency with which words tend to appear together (collocations) (see Blázquez-Carretero, 2019, on the development of this filter).

Overall, 379 WCF were designed, tackling 4,653 misspellings in CORANE and covering 25,609 potential errors. For activation, all feedback was programmed into the PSC alongside the database. Once the search-and-match algorithm was developed, the reference corpus built, and WCFs included, the PSC was ready for trial.

4. Testing the efficacy of the PSC detecting and providing WCF

With the PSC operative, an empirical study was executed to determine its efficacy in both error detection and feedback provision. We lifted 60 compositions comprising 11,475 words written by beginner-level, intermediate, and advanced SFL learners from CEDEL2 (Lozano, 2022). Once the misspellings were identified, the compositions were run through the PSC's and MW's spelling filters to answer these research questions:

RQ1: How effective is the PSC compared to MW in detecting SFL learners' spelling errors?

RQ2: How effective is the PSC vis-à-vis MW in providing adequate feedback to each detected error?

RQ3: Do results vary across different SFL proficiency levels?

https://doi.org/10.1017/S0958344023000150 Published online by Cambridge University Press

4.1 Materials and sampling

To evaluate the PSC's efficacy, we chose the corpus CEDEL2 (Lozano, 2022) containing 802,019 words, involving 2,578 participants from two different L1 backgrounds (as of February 2019).[5] CEDEL2 was piloted before conducting the study to explore other possible variables relevant to sample selection. Fifteen compositions comprising 2,732 words were randomly chosen. Unsurprisingly, a positive Pearson's correlation between proficiency level[6] and essay length was found ($r_{(13)}$ = .545, p = .036) versus a significant negative Pearson's correlation between proficiency level and number of misspellings ($r_{(13)}$ = .642, p < .001). Simply put, more advanced SFL learners tend to write longer texts and commit fewer errors. Additionally, take-home essays contain a considerably lower percentage of errors ($M = 6.56\%$, $SD = 8.99\%$) versus in-class tasks with no support for students ($M = 9.25\%$, $SD = 4.8\%$); an independent sample t-test revealed a statistically significant difference ($t_5 = 0.526$, $p = 0.31$). Finally, to account for the wide-ranging corpus content (with 14 different topic/prompts), the final sample was selected using blocked stratified randomisation considering these variables: proficiency level, use of supporting materials, and topic/prompt. We also adapted Sinclair's (1991) grammatical analysis technique for this study's objectives to determine a sound sample size, ensuring reliable results. An initial sample of 30 compositions (5,521 words) by 30 different SFL learners was selected. Two university-level SFL professors identified, first individually and then together, 577 spelling errors. They categorised these based on the edit distance required to correct the misspelling. A second sample of 30 compositions (5,954 words) by 30 different SFL learners was then selected, revealing 642 errors. As the second sample has a similar error-per-word rate with no new categories of error, "the research would have covered a critical mass of data" (Harvey-Scholes, 2018: 149).

Final sampling (11,475 words) settled on 60 compositions (20 for each proficiency level, representing all topic/prompts, and executed without supporting materials). This quadruples Lawley's (2016) sample size (2,648 words) and falls within the range of successful ones from similar studies: Hernández García's (2017) contained 7,184 words, San Mateo's (2016) 9,500, Chacón-Beltrán's (2017) 12,063, and Harvey-Scholes's (2018) 13,644.

4.2 Method and analysis

The selected 60 compositions were inputted one by one on the PSC. We first tallied the detected misspellings, followed by the number of errors that received adequate WCF. The same compositions were inputted on MW with the spellchecker activated but the grammar checker and autocorrection feature deactivated. As with the PSC, we first counted the errors detected by MW. Then, its proposed alternatives (feedback) were analysed to verify (1) whether the correct word figured in the drop-down list and (2) in which position the correct replacement word appeared on the list (see Blázquez-Carretero & Fan, 2019). If the correct word failed to appear in first position, MW's feedback was deemed inadequate. There is widespread conviction among L2 learners that the first suggestion is always the correct replacement (Blázquez-Carretero & Woore, 2023; Heift & Rimrott, 2008; Rimrott & Heift, 2006), which is not always the case (Blázquez-Carretero, 2019; Rimrott & Heift, 2008). Indeed, L2 learners often choose an inappropriate replacement when the first alternative in the list is incorrect (Rimrott & Heift, 2008). A statistical comparison of the data gathered from PSC and MW followed. The frequency of misspellings was standardised and defined as number of errors/100 words, since the number of errors/composition was influenced by the total number of words (i.e. the longer the text, the more probable that an SFL learner commits misspellings).

[5]As of March 2022, CEDEL2 consists of 1,105,936 words written by 4,399 participants from 11 different L1 backgrounds. Visit http://cedel2.learnercorpora.com/ for further details.

[6]CEDEL2 uses two proficiency-level measurements: the University of Wisconsin's standardised placement test and a self-assessment measure.

https://doi.org/10.1017/S0958344023000150 Published online by Cambridge University Press

Table 1. Summary statistics of measures of spelling errors and feedback adequacy

Variable	*M*	*SD*	Min.	Max.
Errors per 100 words	11.5	6.22	2.18	28
PSC				
Error detection rate	92.5%	10.0%	50.0%	100.0%
Feedback-to-error ratio	71.8%	19.8%	0.0%	100.0%
MW				
Error detection rate	83.2%	16.2%	20.0%	100.0%
Feedback-to-error ratio	72.9%	20.5%	0.0%	100.0%

Note. PSC = pedagogic spellchecker; MW = Microsoft Word.

4.3 Results

4.3.1 Evaluating the detection capacity

Overall, 1,218 spelling errors were detected (i.e. 10.6% of 11,475 words were misspelled). The PSC detected 1,109 among these, while MW highlighted 1,011. Table 1 shows that SFL learners misspelled an average of one per 10 words ($M = 11.5$, $SD = 6.22$). The PSC's average error detection rate was 92.5% – that is, 9.3% higher than MW's (83.2%). Another important finding is PSC's lowest error detection rate at 50%. At its worst, the PSC detected one of two errors, effectively outperforming MW, sometimes detecting as few as 20% of misspellings.

The primary reason behind this is that MW overlooks misspellings corresponding with words documented in its reference corpus regardless of frequency, such as *intimo* (I become close), *arboles* (you hoist [present subjunctive]), or *frio* (I fry). Furthermore, MW failed to detect errors written in upper case, which are treated as proper nouns not requiring correction. Unlike MW, the PSC managed to spot these thanks to its secondary database of common SFL errors. However, the PSC failed to identify L1 words (e.g. **for*, **fashion*). A closer look at the surrounding linguistic context of these words revealed that the PSC's reference corpus (San Mateo, 2016) contained some passages written directly in English. A thorough revision of said corpus may address these gaps for future studies.

Finally, both spellcheckers increased their error detection efficacy when jointly using spelling and grammar filters. Combining both, the PSC identified 1,168 errors (95.9%), while MW identified 1,106 (90.8%). The PSC's bigram filter (see Blázquez-Carretero, 2019) failed to identify words inputted in English (e.g. *the*, *bell*). The same applies to MW's grammar filter, but with the further limitation of overlooking incorrect capitalisation (e.g. **Julio* [July]) and lack of diacritic accents in instances of infrequent collocation (e.g. MW detected the misspelling in *hablamos* **ingles* [we speak English] but not *en* **ingles* [in English]).

4.3.2 Evaluating the adequacy of the WCF

The PSC's mean correct feedback-to-error ratio (i.e. the number of adequate WCF/the number of spelling errors) was 71.8%, slightly lower than MW's 72.9% (see Table 1). Globally (see Table 2), the PSC provided pedagogical WCF to 838 errors (75.6% of the total detected), while MW suggested the correct alternative to 815 misspellings (80.6% of the total detected) thanks to the edit-distance algorithm's reliability in solving single-letter violations. The correct alternative was placed second in 4.75% of cases (48 errors) and third in 1.58% (16 errors).

Figure 2 shows that the correct alternative to **attendo* (i.e. *atiendo* [I'm paying attention]) appears second to *atando* (I'm lacing). As MW's algorithm prioritises the spelling that requires a

https://doi.org/10.1017/S0958344023000150 Published online by Cambridge University Press

Table 2. Number of errors detected and written corrective feedback (WCF) provided

	PSC				MW					
						WCF				
	Words	Errors	Det.	WCF	Det.	1st	2nd	3rd	W	No
Beg.	2,605	345	312	223	288	207	12	7	57	2
Inter.	3,585	497	447	326	403	322	18	4	50	8
Adv.	5,285	352	352	289	324	286	18	5	14	1
Total	11,475	1,218	1,109	838	1,011	815	48	16	121	11

Note. PSC = pedagogic spellchecker; Beg. = beginner; Inter. = intermediate; Adv. = advanced; Det.= detected; MW = Microsoft Word; W = wrong WCF; No = no WCF.

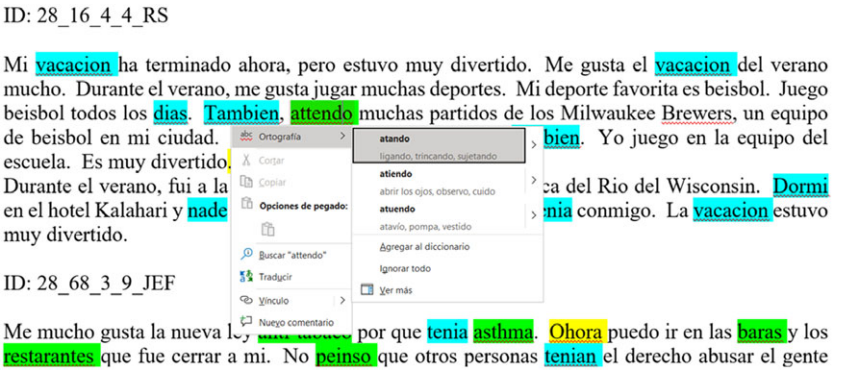


Figure 2. Microsoft Word's feedback for **attendo*

shorter edit distance, substituting *e* for *a* gains preference over adding an *i* because the second requires a longer edit and is hence deemed less precise.

Conversely, the PSC draws on a database of common SFL misspellings. Thanks to its search-and-match facility, each error appearing in the database or following an orthographic pattern is provided a tailored pedagogic WCF after the error is clicked on. For example, the digraph <tt> does not exist in Spanish, prompting the PSC to offer this WCF to the error **attendo* (see Figure 3).

To avoid misleading the L2 learner with inaccurate WCF, the PSC does not address errors absent from its database. In contrast, MW failed to provide the correct alternative in 12% of cases (121 errors), which could be highly confusing for SFL learners. In Figure 4, we see that no suggestion matches with either *por ciento* (percent) or *invierno* (winter), misspelled as **percento* and **vierno*.

In exceptional cases (11 errors or 1.08% of the total detected), MW provided absolutely no feedback (see Figure 5).

Overall, MW failed to list the correct alternative in the first position to 64 errors and provided wrong feedback (121) or no feedback (11) to 132 errors, amounting to 196 incorrect and potentially misleading feedback.

https://doi.org/10.1017/S0958344023000150 Published online by Cambridge University Press

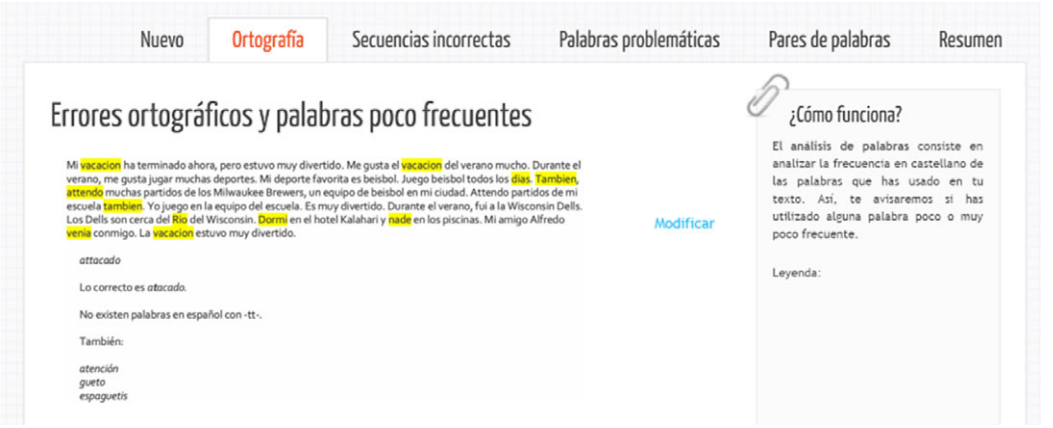


Figure 3. Pedagogic spellchecker's feedback for *attendo

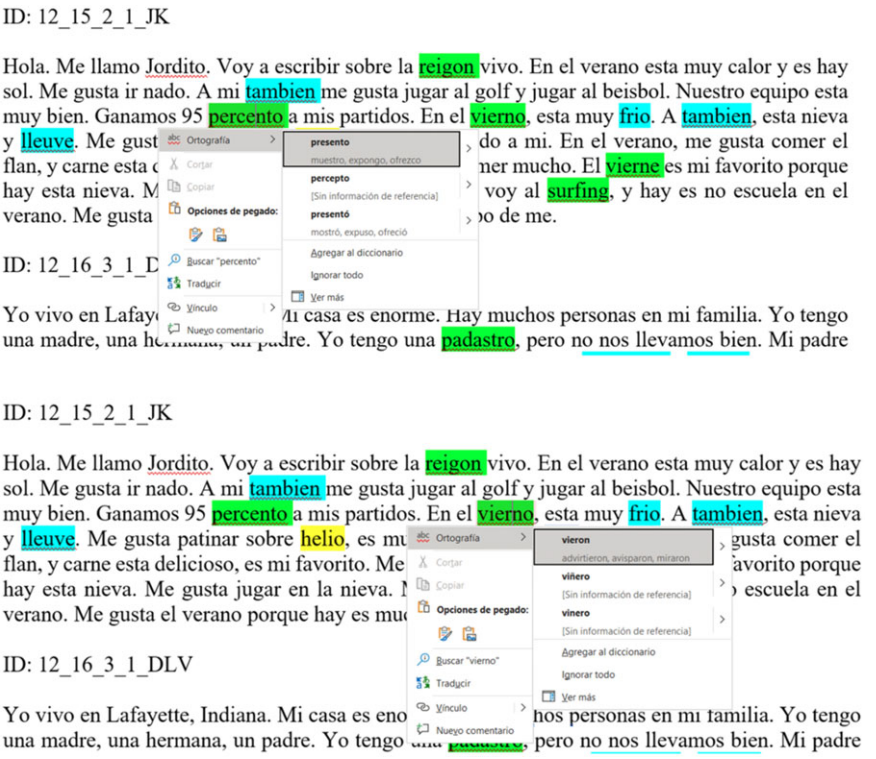


Figure 4. Microsoft Word's feedback for **percento* and **vierno*

4.3.3 Evaluating efficacy across proficiency levels

Finally, we analysed the correlation between learner proficiency and each spellchecker's capacity for error detection and feedback provision. As shown in Table 2, beginner-level SFL learners wrote fewer words (2,605) while committing 345 misspellings, intermediate learners 3,585 words with 497 misspellings, and advanced learners 5,285 words with 352 misspellings. Regarding the average number of errors per 100 words, advanced SFL learners made approximately 50% fewer misspellings (7.87) than their beginner-level (13.2) and intermediate (13.5) peers (see Table 3).

https://doi.org/10.1017/S0958344023000150 Published online by Cambridge University Press

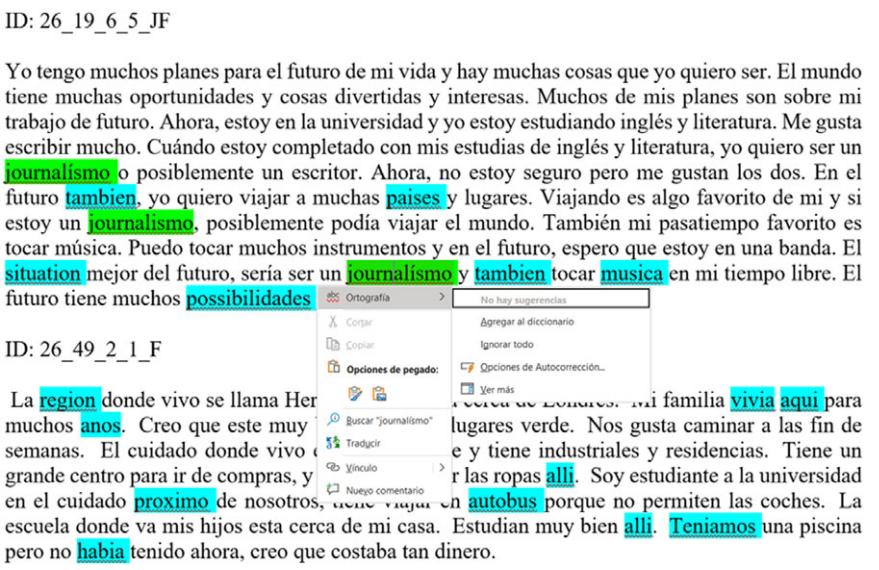


Figure 5. Microsoft Word's feedback for **journalísmo*

Table 3 displays the PSC's detection rate according to proficiency level, demonstrating a 90.8% effectivity for beginners, 93.6% for intermediates, and 93.1% for advanced, each scoring higher than MW by 8.13%, 11.3%, and 8.50% respectively. A paired sample *t*-test (see Table 4) revealed that this difference is significant across all proficiency levels: beginners ($t_{(19)} = 2.125$, $p < .05$), intermediates ($t_{(19)} = 2.59$, $p < .05$), and advanced ($t_{(19)} = 3.171$, $p < .01$).

Meanwhile, both spellcheckers yielded different feedback-to-error ratios (see Table 3). The PSC's ratio starts at 65.7% for beginners and steadily increases with the proficiency level until it reaches 76.4% for advanced learners. The same upwards slope contingent on SFL competency is observed with MW, from 66.8% (beginner) to 80.1% (advanced). However, although MW scored a higher feedback-to-error ratio for beginners and advanced learners, the PSC yielded the highest ratio for intermediate learners. As illustrated in Table 5, a paired sample *t*-test discloses no significant difference between the correct feedback-to-detected-error ratio of both spellcheckers. For beginners, the PSC scored 1.13% lower on average than MW ($t_{(19)} = -0.201$, $p > .05$). For intermediate learners, the PSC posed a higher, albeit statistically, insignificant ratio of approximately 1.58% than MW ($t_{(19)} = 0.305$, $p > .05$). Lastly, MW's correct feedback-to-detected-error ratio of 3.71% is higher than PSC's, although still statistically insignificant ($t_{(19)} = -0.811$, $p > .05$).

5. Discussion

5.1 The PSC detects more than 90% of the spelling errors

The spellcheckers under scrutiny (the PSC and MW) display a high error detection rate, spotting over 80% of the misspellings present in 60 essays written by SFL students. The present findings indicate a sound method behind the creation of a competitive PSC with a well-designed detection engine adapted to L2 contexts. Furthermore, they corroborate the results of previous studies suggesting a near 90% capacity for error recognition (Blázquez-Carretero & Fan, 2019; Rimrott & Heift, 2008).

The PSC significantly identified more SFL spelling errors than its generic counterpart, demonstrating that the former is a more effective detection software than the latter in L2 learning

https://doi.org/10.1017/S0958344023000150 Published online by Cambridge University Press

Table 3. Mean scores of spelling errors and feedback adequacy per proficiency level

Variable	Beginner	Intermediate	Advanced
Errors per 100 words	13.2	13.5	7.87
PSC			
Error detection rate	90.8%	93.6%	93.1%
Feedback-to-error ratio	65.7%	73.4%	76.4%
MW			
Error detection rate	82.7%	82.3%	84.6%
Feedback-to-error ratio	66.8%	71.8%	80.1%

Note. PSC = pedagogic spellchecker; MW = Microsoft Word.

Table 4. Spelling error detection rates per SFL learners' proficiency level

		95% Confidence interval				
Proficiency level	Mean difference	Lower	Upper	*t* stat	*df*	*p* value
Beginner	8.13%	0.12%	16.13%	2.125	19	0.0469*
Intermediate	11.29%	2.17%	20.41%	2.59	19	0.0180*
Advanced	8.50%	2.89%	14.13%	3.171	19	0.0050**

Note. SFL = Spanish as a foreign language.*Denotes significance at 5% alpha.
**Denotes significance at 1% alpha.

Table 5. Correct feedback-to-error ratio per SFL learners' proficiency level

		95% Confidence interval				
Proficiency level	Mean difference	Lower	Upper	*t* stat	*df*	*p* value
Beginner	−1.13%	−12.95%	10.68%	−0.201	19	0.8428
Intermediate	1.58%	−9.29%	12.46%	0.305	19	0.7639
Advanced	−3.71%	−13.28%	5.86%	−0.811	19	0.4275

Note. SFL = Spanish as a foreign language.

environments. MW's main drawback is its failure to detect errors among words present in their reference corpus. Another disadvantage is MW's incapability to spot capitalised misspellings, automatically tagging them as acceptable proper nouns. In contrast, the PSC highlighted more misspellings thanks to its second database, developed from real data and containing the commonest SFL orthographic errors. This validated Lawley's (2016) suggestion that the detection capacity of L2-oriented spellcheckers can be strengthened by enriching the database of correct words (dictionary or corpus), conducting detailed studies of frequently occurring errors among L2 learners, and including potential errors based on orthographic patterns.

5.2 The PSC provides feedback to 69% of the spelling errors

Both spellcheckers successfully provided WCF to two of three misspellings detected in the 60 compositions. However, feedback quality differed. Although our findings confirm Rimrott and

https://doi.org/10.1017/S0958344023000150 Published online by Cambridge University Press

Heift's (2006, 2008) argument that MW's edit-distance algorithm effectively fulfils its primary purpose of correcting performance-based single-letter violations, relying on L1-oriented spellcheckers arguably poses certain disadvantages for SFL learners. Besides limitations in error detection, MW's autocorrection feature limits learner awareness, impeding them from recognising the gap between their interlanguage and proper usage in the target language. As the noticing hypothesis posits, this curtails the possibility of learning from one's errors (Schmidt, 1990, 2010). Secondly, in approximately one of three misspellings, MW offered an incorrect replacement word, validating Mitton and Okada's (2007) conclusions. This poses a problem as L2 learners tend to select an inappropriate substitute for their error (Blázquez-Carretero & Woore, 2023; Heift & Rimrott, 2008). In one of 10 instances, the correct word did not even figure among the suggested alternatives, so that the feedback provided may ultimately be counterproductive not only for self-correction but also for L2 acquisition. Finally, and most evidently, MW does not explain orthographic regularities even if teaching spelling rules explicitly proves more effective than implicit learning (Blázquez-Carretero & Woore, 2021; Graham & Santangelo, 2014). Despite MW's usefulness for competent writers who accidentally commit misspellings, it was never designed as a didactic tool for non-native writers, most of whose errors are competence-based. Hence, MW proves far less satisfactory in addressing L2 needs.

In contrast, not only does the PSC reveal the learning gap (how SFL learners write versus how they ought to write), its pedagogic feedback detailing orthographic patterns also imparts understanding of why certain errors are committed and how these can be avoided in future writing. This encourages deep learning (Lee, 2013), since profound cognitive processes are involved when L2 learners are stimulated to notice the gap, understand its root cause, and rewrite their misspelling. Furthermore, the short-term storage required by these actions may be a viable catalyst for L2 learning (Lázaro Ibarrola, 2013). This is precisely why the PSC was not designed to autocorrect but to provide specific WCF explaining common errors, and to require the learner to rewrite the word rather than offer alternatives that can be selected with a simple mouse click.

WCF has been proved to help L2 learners in immediately correcting their writing (Kang & Han, 2015), demonstrating further that such improvements were maintained in delayed post-tests (Bitchener & Knoch, 2010). There is also "modest evidence" that computer-generated WCF has a positive effect on the produced output (Stevenson & Phakiti, 2014: 62) and can facilitate error correction (Li *et al.*, 2017). This was confirmed by Blázquez-Carretero and Woore (2021), who empirically evaluated the PSC's feedback provision. Also noteworthy is the synchronous nature of WCF in contrast with the common pedagogic habit of providing feedback after the writing process. As such, L2 learners can self-correct while committing an error, thus immediately incorporating the feedback into their output. This encourages autonomous learning and revision while narrowing the unproductive temporal gap between error and correction (Guichon, Bétrancourt & Prié, 2012; Warschauer, 2010).

5.3 The PSC provides WCF to two out of three words independently of the proficiency level of the learners

Results suggest that, irrespective of L2 proficiency level, the PSC was constantly and significantly more effective in detecting errors than MW. Both spellcheckers fared equally in the number of misspellings found to which appropriate WCF was provided.

As expected, higher-proficiency SFL learners made fewer mistakes compared to their lower-proficiency peers. However, spelling accuracy improvement seems to only occur once the learner has reached advanced proficiency (further investigation is necessary to understand why and when this happens). Their errors also differ in nature. Beginners tended to write more words directly in their L1 and committed more randomised misspellings (skewing more than one letter), which sometimes hindered the spellcheckers from providing appropriate WCF. This explains why both spellcheckers displayed a lower accuracy rate in feedback provision at this level (around 60%) than

https://doi.org/10.1017/S0958344023000150 Published online by Cambridge University Press

at higher-proficiency levels. Contrarily, single-letter violations are frequently attributed to advanced learners. On many occasions, intermediate and advanced learners were guilty of hypercorrection to the point of misgeneralising spelling patterns (e.g. placing more accents than necessary instead of omitting uncertain ones). This proved the reliability of MW's edit-distance algorithm in suggesting the correct alternative for an ortho-typographical error. However, MW autocorrects most errors, inhibiting students' active error recognition and subsequent learning (Schmidt, 1990, 2010). The PSC, in contrast, never autocorrects. Despite the impossibility of providing feedback to every single error, especially in cases where it does not result from misunderstanding or ignorance of a spelling pattern, the PSC's accurate detection of around 90% of the errors offers a huge advantage, considering that error recognition is a necessary first step to addressing the problem (Schmidt, 1990, 2010).

6. Conclusions and implications

Research suggests that it is desirable for L2 learners to self-correct their writing (Quinn, 2015; Vickers & Ene, 2006) not only to compensate for resource and time constraints but also to encourage "deep processing" (Craik & Tulving, 1975). Since learners need to draw on external resources and their L2 knowledge to self-correct, the process stimulates and promotes independent learning (Park & Kim, 2019). However, L2 students might fail to identify their errors (Lee, 1997) and/or properly address them. Spellcheckers thus emerge as a sustainable learning resource to overcome these limitations.

Considering GSCs to be inadequate didactic tools in an L2 environment (Blázquez-Carretero & Fan, 2019; Heift & Rimrott, 2008), there remains a compelling reason to build a reliable pedagogic software to improve L2 spelling competency. There are strong arguments in favour of a spellchecker especially crafted for SFL learners. An L2-oriented software capable of targeting SFL errors that vastly differ from L1 mistakes and of giving WCF illustrating patterns and rules governing Spanish orthography would ultimately facilitate spelling accuracy (Blázquez-Carretero & Woore, 2021). Furthermore, Lawley (2016) and Blázquez-Carretero and Woore (2023) evidenced that L2 students prefer a spellchecker specifically tailored to their needs.

This study presents immediate implications to the development of spellcheck packages and L2 pedagogy alike as it illustrates how complex spelling rules can be easily adapted and rendered accessible to SFL learners. Consequently, providing SFL teachers and students with an online error detection tool adept at offering feedback with an explanation of why the error was committed and how it can be corrected avoids recourse to time-consuming tasks like editing and self-correction.

The student is enabled to autocorrect confidently and autonomously, which not only serves as a learning opportunity during the writing process but also saves considerable time and effort on the teacher's part. It is demanding for the latter to give feedback as precise as that provided by the PSC for every single error committed by their students, and to further ensure sufficient awareness and future application. Given orthography's importance in L2 acquisition and the observable difficulties in adequately addressing orthographic issues in diverse L2 learning set-ups, Lee (2013) persuasively argues for teachers to be relieved of this burden by delegating the task to a pedagogic tool specifically crafted for L2 peer- and self-correction. As such, teachers would significantly save time that could potentially be allotted to more pressing aspects of L2 writing acquisition requiring human attention, like content, argumentation, organisational structure, or register (Chen & Cheng, 2008).

New technologies have been observed to influence L2 writing processes as editing is rendered increasingly feasible and accessible to a diverse set of learners. These advancements are redrawing the landscape of L2 writing acquisition, effectively altering classroom practices and learning timetables. In sum, technologically aided delivery of WCF, subject to constant updates and improvements, would doubtless enable diligent L2 students to learn autonomously and, perhaps, empower them to progress at a much faster rate.

https://doi.org/10.1017/S0958344023000150 Published online by Cambridge University Press

Acknowledgements. This study would not have been possible without the seminal work of Dr. James Lawley. I would like to express my gratitude to the Departamento de Lengua y Literatura Españolas, Teoría de la Literatura y Lingüística General of the Facultad de Filología of the University of Santiago de Compostela (Spain) for their support, with special thanks to Prof. Mª Victoria Vázquez Rozas and Mª José Rodríguez Espiñeira. I am also grateful to Prof. Jillian Melchor, Prof. Cecilia Blázquez, Dr. Alberto Cifuentes Negrete, and the anonymous reviewers for their invaluable feedback on earlier versions of this paper.

Ethical statement and competing interests. This research is based on publicly available anonymised corpus data. This research is original, and I declare that I have no conflict of interest.

References

Barcroft, J. & Rott, S. (2010) Partial word form learning in the written mode in L2 German and Spanish. *Applied linguistics*, 31(5): 623–650. https://doi.org/10.1093/applin/amq017

Bestgen, Y. & Granger, S. (2011) Categorising spelling errors to assess L2 writing. *International Journal of Continuing Engineering Education and Life Long Learning*, 21(2–3): 235–252. https://doi.org/10.1504/IJCEELL.2011.040201

Bitchener, J. & Knoch, U. (2010) Raising the linguistic accuracy level of advanced L2 writers with written corrective feedback. *Journal of Second Language Writing*, 19(4): 207–217. https://doi.org/10.1016/j.jslw.2010.10.002

Blázquez-Carretero, M. (2019) Using bigrams to detect written errors made by learners of Spanish as a foreign language. *CALL-EJ*, 20(2): 55–69. http://callej.org/journal/20-2/Carretero2019.pdf

Blázquez-Carretero, M. (2022) Is it worth correcting L2 learners' written work? In García Orta, M. J. & Martín Santos, R. (eds.), *El poder de la comunicación: periodismo, educación y feminismo*. Madrid: Dykinson S.L., 829–862. 'https://www.researchgate.net/publication/366544503_Is_it_worth_correcting_L2_learners'_written_work

Blázquez-Carretero, M. & Fan, C. (2019) The efficacy of spell check packages specifically designed for second language learners of Spanish. *Pertanika Journal of Social Sciences & Humanities*, 27(2): 847–861. http://www.pertanika.upm.edu.my/pjssh/browse/regular-issue?article=JSSH-3837-2018

Blázquez-Carretero, M. & Woore, R. (2021) Can a 'pedagogical' spellchecker improve spelling accuracy in L2 Spanish? *Language Learning & Technology*, 25(2): 135–157. http://hdl.handle.net/10125/73437

Blázquez-Carretero, M. & Woore, R. (2023) *Spelling Spanish: L2 learners' interaction with spellcheckers* [Manuscript in preparation].

Brown, G. D. A. & Ellis, N. C. (1994) *Handbook of spelling: Theory, process and intervention*. Chichester: Wiley.

Brown, H. D. (1994) *Principles of language learning and teaching*. Englewood Cliffs, NJ: Prentice Hall

Cestero Mancera, A. M. & Penadés Martínez, I. (2009) *Corpus de textos escritos para el análisis de errores de aprendices de E/LE (CORANE)* [CD-ROM]. Alcalá de Henares: Universidad de Alcalá.

Chacón-Beltrán, R. (2017) Free-form writing: Computerized feedback for self-correction. *ELT Journal*, 71(2): 141–149. https://doi.org/10.1093/elt/ccw064

Chen, C.-F. E. & Cheng, W.-Y. E. (2008) Beyond the design of automated writing evaluation: Pedagogical practices and perceived learning effectiveness in EFL writing classes. *Language Learning & Technology*, 12(2): 94–112. https://doi.org/10125/44145

Cotos, E. (2011) Potential of automated writing evaluation feedback. *CALICO Journal*, 28(2): 420–459. https://doi.org/10.11139/cj.28.2.420-459

Council of Europe (2001) *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.

Craik, F. I. M. & Tulving, E. (1975) Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, 104(3): 268–294. https://doi.org/10.1037/0096-3445.104.3.268

Ellis, N. C. & Beaton, A. (1993) Psycholinguistic determinants of foreign language vocabulary learning. *Language Learning*, 43(4): 559–617. https://doi.org/10.1111/j.1467-1770.1993.tb00627.x

Figueredo, L. (2006) Using the known to chart the unknown: A review of first-language influence on the development of English-as-a-second-language spelling skill. *Reading and Writing*, 19(8): 873–905. https://doi.org/10.1007/s11145-006-9014-1

Frazer, K. A., Murray, S. S., Schork, N. J. & Topol, E. J. (2009) Human genetic variation and its contribution to complex traits. *Nature Reviews Genetics*, 10(4): 241–251. https://doi.org/10.1038/nrg2554

Goswami, U. (1988) Children's use of analogy in learning to spell. *British Journal of Developmental Psychology*, 6(1): 21–33. https://doi.org/10.1111/j.2044-835X.1988.tb01077.x

Goulandris, N. K. (1994) Teaching spelling: Bridging theory and practice. In Brown, G. D. A. & Ellis, N. C. (eds.), *Handbook of spelling: Theory, process and intervention*. Chichester: Wiley, 407–423.

Graham, S. & Santangelo, T. (2014) Does spelling instruction make students better spellers, readers, and writers? A meta-analytic review. *Reading and Writing*, 27(9): 1703–1743. https://doi.org/10.1007/s11145-014-9517-0

Grgurović, M., Chapelle, C. A. & Shelley, M. C. (2013) A meta-analysis of effectiveness studies on computer technology-supported language learning. *ReCALL*, 25(2): 165–198. https://doi.org/10.1017/S0958344013000013

https://doi.org/10.1017/S0958344023000150 Published online by Cambridge University Press

Guichon, N., Bétrancourt, M. & Prié, Y. (2012) Managing written and oral negative feedback in a synchronous online teaching situation. *Computer Assisted Language Learning*, 25(2): 181–197. https://doi.org/10.1080/09588221.2011.636054

Harvey-Scholes, C. (2018) Computer-assisted detection of 90% of EFL student errors. *Computer Assisted Language Learning*, 31(1–2): 144–156. https://doi.org/10.1080/09588221.2017.1392322

Heift, T. (2008) *Corrective feedback in CALL: A longitudinal study of learner uptake*. Paper presented at the annual meeting of the American Association of Applied Linguistics, Washington, DC.

Heift, T. & Hegelheimer, V. (2017) Computer-assisted corrective feedback and language learning. In Nassaji, H. & Kartchava, E. (eds.), *Corrective feedback in second language teaching and learning: Research, theory, applications, implications*. New York: Routledge, 51–65. https://doi.org/10.4324/9781315621432-5

Heift, T. & Rimrott, A. (2008) Learner responses to corrective feedback for spelling errors in CALL. *System*, 36(2): 196–213. https://doi.org/10.1016/j.system.2007.09.007

Heift, T. & Schulze, M. (2007) *Errors and intelligence in computer-assisted language learning: Parsers and pedagogues*. New York: Routledge. https://doi.org/10.4324/9780203012215

Hernández García, F. (2017) La detección y corrección de errores en la deixis temporal del verbo en redacciones escritas en inglés como lengua extranjera. *Estudios de lingüística inglesa aplicada*, 17: 183–207. https://doi.org/10.12795/elia.2017.i17.08

Kang, E. & Han, Z. (2015) The efficacy of written corrective feedback in improving L2 written accuracy: A meta-analysis. *The Modern Language Journal*, 99(1): 1–18. https://doi.org/10.1111/modl.12189

Kelley, A. & Kohnert, K. (2012) Is there a cognate advantage for typically developing Spanish-speaking English-language learners? *Language, Speech, and Hearing Services in Schools*, 43(2): 191–204. https://doi.org/10.1044/0161-1461(2011/10-0022)

Kellogg, R. T. (2001) Competition for working memory among writing processes. *The American Journal of Psychology*, 114(2): 175–191. https://doi.org/10.2307/1423513

Lawley, J. (2015) New software to help EFL students self-correct their writing. *Language Learning & Technology*, 19(1): 23–33. https://doi.org/10125/44396

Lawley, J. (2016) Spelling: Computerised feedback for self-correction. *Computer Assisted Language Learning*, 29(5): 868–880. https://doi.org/10.1080/09588221.2015.1069746

Lázaro Ibarrola, A. (2013) Reformulation and self-correction: Insights into correction strategies for EFL writing in a school context. *Vigo International Journal of Applied Linguistics*, 10: 29–49. https://revistas.uvigo.es/index.php/vial/article/view/55/55

Lee, I. (1997) ESL learners' performance in error correction in writing: Some implications for teaching. *System*, 25(4): 465–477. https://doi.org/10.1016/S0346-251X(97)00045-6

Lee, I. (2013) Research into practice: Written corrective feedback. *Language Teaching*, 46(1): 108–119. https://doi.org/10.1017/S0261444812000390

Li, Z., Feng, H.-H. & Saricaoglu, A. (2017) The short-term and long-term effects of AWE feedback on ESL students' development of grammatical accuracy. *CALICO Journal*, 34(3): 355–375. https://doi.org/10.1558/cj.26382

Lin, P.-H., Liu, T.-C. & Paas, F. (2017) Effects of spell checkers on English as a second language students' incidental spelling learning: A cognitive load perspective. *Reading and Writing*, 30(7): 1501–1525. https://doi.org/10.1007/s11145-017-9734-4

Long, M. H. (1996) The role of the linguistic environment in second language acquisition. In Ritchie, W. C. & Bhatia, T. K. (eds.), *Handbook of second language acquisition*. San Diego: Academic Press, 413–468. https://doi.org/10.1016/B978-012589042-7/50015-3

Lozano, C. (2022) CEDEL2: Design, compilation and web interface of an online corpus for L2 Spanish acquisition research. *Second Language Research*, 38(4): 965–983. https://doi.org/10.1177/02676583211050522

Manchón, R. M. (2014) The internal dimension of tasks: The interaction between task factors and learner factors in bringing about learning through writing. In Byrnes, H. & Manchón, R. M. (eds.), *Task-based language learning: Insights from and to L2 writing*. Amsterdam: John Benjamins, 27–52. https://doi.org/10.1075/tblt.7.02man

Manchón, R. M. & Vasylets, O. (2019) Language learning through writing: Theoretical perspectives and empirical evidence. In Schwieter, J. W. & Benati, A. (eds.), *The Cambridge handbook of language learning*. Cambridge: Cambridge University Press, 341–362. https://doi.org/10.1017/9781108333603.015

Marcos-Miguel, N. (2013) *Analyzing instruction and learning of derivational morphology in the Spanish foreign language classroom*. University of Pittsburgh, PhD dissertation.

Mitton, R. (2010) Fifty years of spellchecking. *Writing Systems Research*, 2(1): 1–7. https://doi.org/10.1093/wsr/wsq004

Mitton, R. & Okada, T. (2007) *The adaptation of an English spellchecker for Japanese writers*. London: Birkbeck ePrints. https://eprints.bbk.ac.uk/id/eprint/592/

Nagata, R., Takamura, H. & Neubig, G. (2017) Adaptive spelling error correction models for learner English. *Procedia Computer Science*, 112: 474–483. https://doi.org/10.1016/j.procs.2017.08.065

Nation, I. S. P. (1990) *Teaching and learning vocabulary: Teaching methods*. Boston: Heinle & Heinle.

Nation, I. S. P. (2001) *Learning vocabulary in another language*. Stuttgart: Ernst Klett Sprachen. https://doi.org/10.1017/CBO9781139524759

https://doi.org/10.1017/S0958344023000150 Published online by Cambridge University Press

Nesi, H. (2014) Dictionary use by English language learners. *Language Teaching*, 47(1): 38–55. https://doi.org/10.1017/S0261444813000402

Park, E. S. & Kim, O. Y. (2019) Learners' engagement with indirect written corrective feedback: Depth of processing and self-correction. In Leow, R. P. (ed.), *The Routledge handbook of second language research in classroom learning*. New York: Routledge, 214–228.

Quinn, C. (2015) Training L2 writers to reference corpora as a self-correction tool. *ELT Journal*, 69(2): 165–177. https://doi.org/10.1093/elt/ccu062

Ranalli, J. (2018) Automated written corrective feedback: How well can students make use of it? *Computer Assisted Language Learning*, 31(7): 653–674. https://doi.org/10.1080/09588221.2018.1428994

Ranalli, J. & Yamashita, T. (2022) Automated written corrective feedback: Error-correction performance and timing of delivery. *Language Learning & Technology*, 26(1): 1–25. http://hdl.handle.net/10125/73465

Real Academia Española (2008) Banco de datos (CREA). *Corpus de Referencia del español Actual*. http://www.rae.es

Real Academia Española (2014) *Diccionario de la lengua española* (23 ed.). Madrid: Espasa. http://www.rae.es

Rimrott, A. & Heift, T. (2006) Language learners and generic spell checkers in CALL. *CALICO Journal*, 23(1): 17–48. https://doi.org/10.1558/cj.v23i1.17-48

Rimrott, A. & Heift, T. (2008) Evaluating automatic detection of misspellings in German. *Language Learning & Technology*, 12(3): 73–92. https://doi.org/10125/44156

Rodríguez, A. J., Villoria, D. B. & Paredes, A. F. (1997) *Las redacciones en lengua extranjera en bachillerato y COU: Investigación pedagógica*. Valladolid: Universidad de Valladolid.

Saint Pierre, A. & Génin, E. (2014) How important are rare variants in common disease? *Briefings in Functional Genomics*, 13(5), 353–361

San Mateo, A. (2016) Un corpus de bigramas utilizado como corrector ortográfico y gramatical destinado a hablantes nativos de español. *Revista signos*, 49(90): 94–118. https://doi.org/10.4067/S0718-09342016000100005

Sánchez Jiménez, D. (2009) Una aproximación a la didáctica de la ortografía en la clase de ELE. *MarcoELE. Revista de didáctica ELE*, 9: 1–22. https://marcoele.com/descargas/9/sanchez_ortografia.pdf

Sánchez Jiménez, D. (2010) El análisis de errores ortográficos de estudiantes filipinos en el aprendizaje de español como LE y su aplicación didáctica. En actas del *I Congreso de Español como Lengua Extranjera en Asia-Pacífico*. Manila: Instituto Cervantes, 106–126.

Sánchez-Gutiérrez, C. H. (2013) *Priming morfológico y conciencia morfológica. Una investigación con estudiantes norteamericanos de E/LE*. Salamanca University, Spain, PhD dissertation.

Schmidt, R. (2010) Attention, awareness, and individual differences in language learning. In Chan, W. M., Chi, S., Cin, K. N., Istanto, J., Nagami, M., Sew, J. W., Suthiwan, T. & Walker, I. (eds.), *Proceedings of CLaSIC 2010, Singapore, December 2-4*. Singapore: National University of Singapore, Centre for Language Studies, 721–737.

Schmidt, R. W. (1990) The role of consciousness in second language learning. *Applied Linguistics*, 11(2): 129–158. https://doi.org/10.1093/applin/11.2.129

Sinclair, J. (1991) *Corpus, concordance, collocation*. Oxford: Oxford University Press.

Stevenson, M. & Phakiti, A. (2014) The effects of computer-generated feedback on the quality of writing. *Assessing Writing*, 19: 51–65. https://doi.org/10.1016/j.asw.2013.11.007

Touchie, H. Y. (1986) Second language learning errors: Their types, causes, and treatment. *JALT Journal*, 8(1), 75–80

van Beuningen, C. G., de Jong, N. H. & Kuiken, F. (2008) The effect of direct and indirect corrective feedback on L2 learners' written accuracy. *ITL: International Journal of Applied Linguistics*, 156(1): 279–296. https://doi.org/10.2143/ITL.156.0.2034439

Vickers, C. H. & Ene, E. (2006) Grammatical accuracy and learner autonomy in advanced writing. *ELT Journal*, 60(2): 109–116. https://doi.org/10.1093/elt/cci097

Warschauer, M. (2010) New tools for teaching writing. *Language Learning & Technology*, 14(1): 3–8. https://doi.org/10125/44196

Webb, S. (2021) Word families and lemmas, not a real dilemma: Investigating lexical units. *Studies in Second Language Acquisition*, 43(5): 973–984. https://doi.org/10.1017/S0272263121000760

About the author

Miguel Blázquez-Carretero is Associate Professor and Spanish Section Coordinator of the Department of European Languages, University of the Philippines Diliman. He is the leading consultant of the Philippines' Department of Education (DepEd) on Spanish curricular design and founded the first Association of Spanish Teachers (AFELE) in the country.

Author ORCiD. Miguel Blázquez-Carretero, https://orcid.org/0000-0003-4001-328X

https://doi.org/10.1017/S0958344023000150 Published online by Cambridge University Press