Journal of Inclusive Methodology and Technology in Learning and Teaching
ISSN 2785-5104
Anno 6 n. 3 (2026)

EDIZIONI UNIVERSITARIE ROMANE
editoria scientifica dal 1952

AI-Supported Formative Feedback for Campidanese Sardinian Writing

Salvatore Carta [1], Alessandro Giuliani [1], Mirko Marras [1], Marco Manolo Manca [1,*], Alessandro Sebastian Podda [1] and Riccardo Secci [1]

[1] Department of Mathematics and Computer Science, University of Cagliari, Cagliari, Italy; salvatore@unica.it (S.C.); alessandro.giuliani@unica.it (A.G.); mirko.marras@unica.it (M.M.); marcom.manca@unica.it (M.M.M.); sebastianpodda@unica.it (A.S.P.); r.secci53@studenti.unica.it (R.S.)
* Correspondence: marcom.manca@unica.it

Abstract: Digital writing support is unevenly distributed across languages, and minority languages such as Sardinian remain largely excluded. This is an educational issue as well as a technical one, since learners have fewer opportunities to revise and reflect on their writing. This contribution presents an AI-supported writing environment for Campidanese Sardinian, designed as a space for formative feedback rather than automatic correction. The environment flags words that may be misspelled, explains the kind of issue involved, offers alternatives to compare, and summarizes recurrent difficulties, leaving every decision to the writer. It combines two language models with rules derived from the Sardinian graphematic repertoire, used both to simulate spelling errors and to explain them. A preliminary evaluation on 160 sentences containing 409 simulated errors showed that 96.4% of the flagged words were actual errors and that 78.5% of the errors were flagged. When suggestions were shown, the correct form was among the first five in 95.4% of cases, and the feedback category matched the kind of error in 88.5%. These results indicate that the feedback is cautious and reliable enough for educational use, and they prepare the ground for experimentation with learners and teachers.

Keywords: formative feedback; minority languages; AI-supported writing

Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Digital technologies have become part of how people write. Spellcheckers, grammar assistants and, more recently, AI-supported writing environments accompany the production and revision of texts at school, at university and in everyday life. This support, however, is unevenly distributed across languages. Mature tools exist for a small number of widely used languages, while most of the world's languages have little or no presence in the digital sphere (Kornai, 2013). For minority languages, this absence is usually described as a technical gap. It is also an educational one, as without writing support, learners and speakers have fewer opportunities to practise, revise what they have written, and reflect on the written form of their language.

Sardinian is a case in point. It is recognized by Italian law as a historical linguistic minority (Law No. 482/1999), and its varieties are classified as definitely endangered in the UNESCO atlas of the world's languages in danger (Moseley,

EDIZIONI UNIVERSITARIE ROMANE
editoria scientifica dal 1952

2010). Writing in Sardinian, and in the Campidanese variety considered here, involves a degree of uncertainty that writers of fully standardized languages rarely face. This is because local varieties coexist, and spelling conventions are still being consolidated. In this situation, the lack of a dedicated spellchecker means that no immediate feedback is available while writing. For learners, teachers, and the communities engaged in promoting the language, this may affect both linguistic confidence and the perception that Sardinian can actually be used in digital and educational settings (Carta et al., 2025).

Research on formative assessment indicates what kind of support would be educationally meaningful. Feedback supports learning when it helps learners see the gap between their current performance and a desired outcome, and when it leads to action that reduces that gap (Black & Wiliam, 1998; Hattie & Timperley, 2007; Sadler, 1989). Its value therefore lies not only in the accuracy of the information it conveys, but also in what learners do with it. Theories of self-regulated learning make the same point from the learner's side, arguing that learners are active agents who interpret feedback, make decisions and progressively take control of their own processes (Nicol & Macfarlane-Dick, 2006; Zimmerman, 2002), and that using feedback well is itself a capacity that has to be developed (Carless & Boud, 2018). Engagement with feedback, in its cognitive as well as its behavioural and emotional dimensions (Fredricks et al., 2004), is thus a condition for feedback to work. Studies of automated written corrective feedback show that this engagement cannot be taken for granted. Indeed, learners do not always deal critically with what a tool proposes, and the usefulness of the feedback depends on how explicit and interpretable it is (Koltovskaia, 2020; Ranalli, 2018). A spellchecker that silently replaces one form with another delivers a corrected text, but leaves little room for learning. An educational spellchecker should instead help writers recognize possible errors, compare alternatives, and make informed choices during the writing process.

A second perspective comes from media education, which invites us to consider digital technologies not as neutral instruments, but as cultural artefacts that shape practices, representations and forms of participation (Rivoltella, 2017). For minority languages, this dimension is particularly relevant. Whether digital tools exist for a language contributes to defining it as usable, legitimate and teachable in contemporary educational environments. Conversely, digital technologies can support language learning, promotion and revitalization when they are designed around the needs of a specific community (Galla, 2016). An AI-supported writing tool for Sardinian is therefore also a matter of linguistic inclusion and cultural participation.

This contribution presents an AI-supported writing environment for Campidanese Sardinian, a variety for which no dedicated spellchecker was previously available. The environment is designed as a space for formative feedback rather than as a merely corrective tool. While the user writes, it highlights the words that may be misspelled, explains why each of them was flagged, offers alternatives to compare, and leaves the decision to the writer; a revision summary then makes recurrent difficulties visible. The paper describes the design of the environment and the linguistic and AI resources it relies on, reports a preliminary evaluation of how reliable its feedback is, and discusses its implications for formative assessment, self-regulation and inclusive language education. The results indicate that feedback of

this kind can be made cautious enough to be trusted, since correct words are rarely flagged, and they prepare the ground for experimentation with learners and teachers.

2. Materials and Methods

2.1. Design Rationale

The environment was designed starting from the conception of feedback outlined above, which was translated into four principles. First, to ensure the writer remains the decision-maker, suggestions are offered but never applied automatically, so that revising remains an activity of the learner, and feedback is something to interpret rather than a result to accept (Nicol & Macfarlane-Dick, 2006). Second, to provide explanatory feedback, each flagged word is accompanied by the kind of issue involved and by the reason for the suggestion, since elaborated feedback that addresses how a task is carried out supports learning more than a simple indication of what is wrong (Hattie & Timperley, 2007; Shute, 2008). Third, to maintain a cautious approach, a word is flagged only when the system is confident, because false alarms are particularly harmful for writers who are unsure of the norm, as they cannot easily check a suggestion elsewhere, and repeated unjustified flags undermine both their trust in the tool and their confidence in what they already know. Fourth, to make the revision process visible, the issues encountered during a writing session are collected in a summary, so that writers can recognize their recurrent difficulties and assess their own writing.

2.2. Language Resources and Synthetic Errors

The linguistic basis of the environment consists of three resources. The first is a collection of 3,170 sentences in Campidanese Sardinian taken from *Po cantu Biddanoa* by Benvenuto Lobina (1987/2004), a novel written in Sardinian with an Italian translation on the facing page. The sentences were divided into 2,709 used to train the models, 304 used to tune them, and 160 set aside for the evaluation. The second is a word list of about 19,000 forms, which combines the Campidanese lemmas published online by the *Acadèmia de su Sardu* (https://sardu.wiki) with the vocabulary of the training sentences. The third is the experimental graphematic repertoire adopted by the Autonomous Region of Sardinia for the certification of the Sardinian language (Regione Autonoma della Sardegna, 2022), which describes how the sounds of Sardinian are represented in writing.

No collection of spelling errors made by writers of Sardinian is currently available. Errors were therefore simulated by altering correct sentences according to three families of slips. Graphematic confusions were derived from the repertoire, as sets of spellings that are easily mistaken for one another: single and double consonants (*casteddu* written as *castedu*), graphemes that stand for similar sounds (*tz* and *z*), and conventions that differ between Italian and Sardinian (*chi* and *ci*, *xi* and *sci*). Typing errors reproduce common keyboard slips, such as a neighbouring key, two letters in the wrong order, or a missing or an extra letter. Accent errors consist of omitting a written accent (*mesudì* written as *mesudi*). In each sentence, about 15% of the words were altered.

2.3. From Detection to Explained Suggestions

EDIZIONI UNIVERSITARIE ROMANE
editoria scientifica dal 1952

Two language models of the BERT family (Devlin et al., 2019) support the environment. The first was obtained by further training a publicly available Italian model on the Campidanese sentences, so that it learns which words are plausible in a Sardinian sentence; starting from a related Romance language compensates for the small size of the corpus. The second was trained on the altered sentences to recognize, from their context, the words that are likely to be misspelled. When a text is analysed, this second model estimates for each word the probability that it is an error, and the word is flagged only above a threshold, which can be adjusted in the interface. For each flagged word, candidate corrections are searched in the word list among the forms that differ from it by a single edit, and the first model ranks them according to how well they fit the sentence.

Each pair formed by a flagged word and a suggestion is then compared and assigned to one of four feedback categories, as illustrated in Table 1, using the same rules that were used to simulate the errors. The explanation shown to the writer is built from a fixed template for each kind of slip. This choice was deliberate: explanations are not generated freely by an AI model, so the same slip always receives the same explanation and no explanation can be invented. When a word looks doubtful, but no acceptable suggestion is found, it is still pointed out as a possible error, as an invitation to check it.

Table 1. Feedback categories, with an example of flagged word, suggestion and explanation shown to the writer.

Category	When it applies	Example and explanation shown
Accent	The word and the suggestion differ only in the accent	*intendiu* → *inténdiu* The accent is missing: "inténdiu" is written with "é".
Graphematic confusion	Two spellings that stand for similar sounds have been mixed up	*castedu* → *casteddu* Single or double consonant? This word is written with "dd", not "d".
Typing error	A letter was swapped, added, left out or mistyped	*cnu* → *cun* Two letters are in the wrong order: "nu" should be "un".
Contextual suggestion	The word exists, or shows no recognizable slip, but another word fits the sentence better	*cini* → *chini* "cini" is a Sardinian word, but "chini" fits this sentence better ("ci" and "chi" are easily confused).

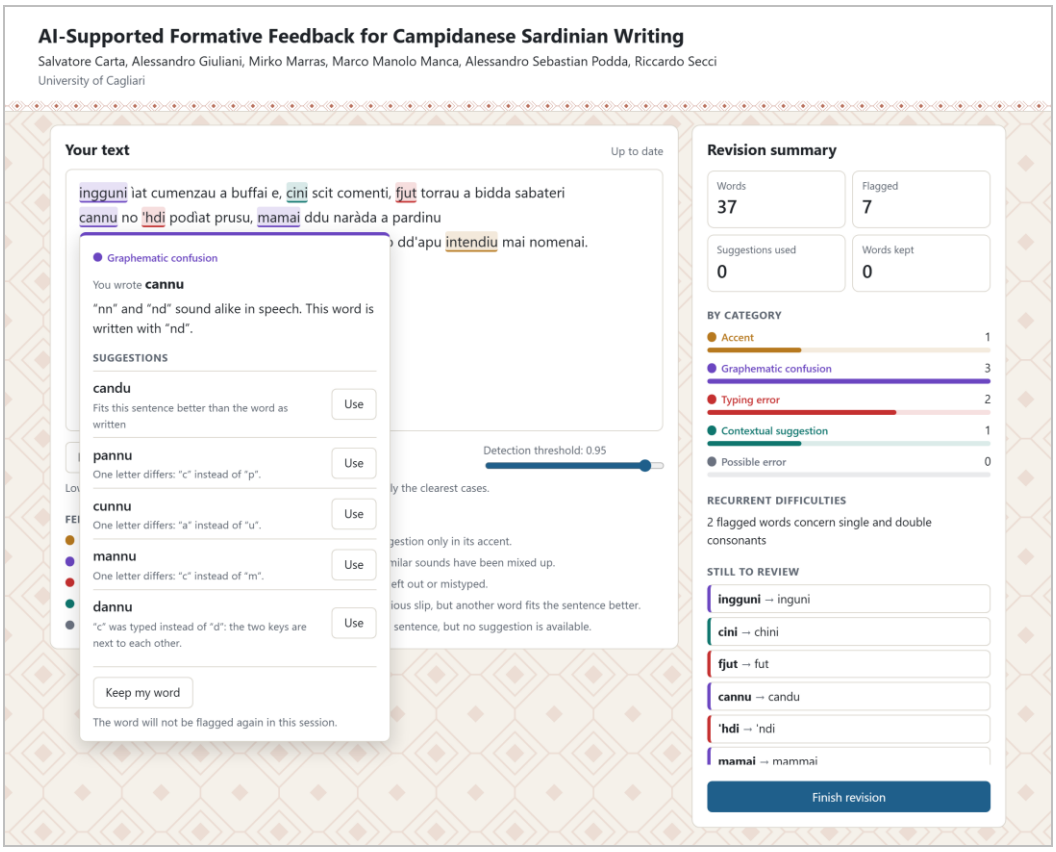
2.4. The Writing Environment

As shown in Figure 1, the environment is a web application that runs on a personal computer. The writer types freely in an editor, and the text is analysed while it is being written. Flagged words are underlined with a different colour for each category; the word that is still being typed is left alone until the writer moves on, so that feedback does not interrupt the flow of writing. Clicking a flagged word opens a feedback card, which shows the category, the explanation, and up to five

alternatives, each with a short note on how it relates to the sentence. The writer can use one of the alternatives or keep the word as written, in which case it is no longer flagged during the session.

A panel beside the editor keeps a running summary of the revision: how many words were flagged, how many suggestions were used, how many words were kept, how the issues are distributed across categories, and which difficulties recur. At the end of the session, a final summary lists the words involved in each difficulty and invites the writer to look for what they have in common. A record of the session (the text, the feedback events, and the summary) is saved on the computer in use, in view of future studies with learners.

Figure 1. The writing environment. Flagged words are underlined in the colour of their feedback category; the card explains why *cannu* was flagged and offers alternatives to compare; the panel on the right shows the revision summary.



2.5. Evaluation Procedure

Before involving learners, we examined whether the feedback is reliable enough to be offered to them. The evaluation used the 160 sentences set aside for this purpose, which the models had never seen, and in which 409 errors were simulated with the procedure described above. The detection threshold and the other settings were chosen on the 304 tuning sentences, not on the evaluation ones.

The evaluation focused on three main aspects, starting with the reliability of the flags. This was assessed through precision (how many flagged words are actually

errors), recall (how many errors are flagged), and their combined F0.5 score. This metric gives precision twice the weight of recall and serves as the reference measure in automatic error correction (Ng et al., 2014), a choice that is consistent with the third design principle. Additionally, the study measured the quality of the suggestions by tracking how often the correct form appeared as the first suggestion, or among the top three to five. Finally, we evaluated the agreement between the feedback category assigned by the system and the kind of error that had been simulated.

The study did not involve human participants, and no ethical approval was required. The code, the data and the instructions to reproduce the models and the results are available at https://github.com/ai4limba/sardinian-formative-feedback.

3. Results

3.1. Reliability of the Feedback

The main results are reported in Table 2. The detection model flagged 321 of the 409 simulated errors (recall 78.5%), and 96.4% of the words it flagged were actual errors (precision), with an F0.5 score of 0.922. Flags on correct words were rare. Among the 3,080 correctly spelled words of the evaluation sentences, 12 were flagged (0.4%), and only 5 of them were accompanied by a suggestion. The cautious behaviour sought in the design was therefore obtained, at the cost of leaving about one error in five unflagged.

Table 2. Main results on the 160 evaluation sentences, which contain 409 simulated errors.

Measure	Result
Flagged words that are actual errors (precision)	96.4%
Errors that are flagged (recall)	78.5%
F0.5 score of the flags	0.922
Correct words flagged	12 of 3,080
Errors shown with suggestions	69.2%
Correct form as the first suggestion	56.2%
Correct form among the first three suggestions	64.1%
Correct form among the first five suggestions	66.0%
Agreement between feedback category and simulated error	88.5%

Of the 321 flagged errors, 283 were shown with at least one suggestion, while the remaining 38 were pointed out as possible errors without a suggestion. For 47 of the 409 errors, the correct form is absent from the word list and therefore could not be suggested in any case. The analysis of a sentence took about 40 milliseconds on a laptop computer, which is fast enough to give feedback while the text is being written.

3.2. Quality of the Suggestions

When an error was shown with suggestions, the correct form was the first suggestion in 81.3% of cases (230 of 283), and it was among the first five in 95.4% (270 of 283). Considering all the simulated errors, the correct form was offered first for 56.2% of them, and among the first five for 66.0%. If the first suggestion were applied automatically, as a conventional corrector would do, 56.2% of the errors would be corrected, with a precision of 0.799 and an F0.5 score of 0.737.

As shown in Table 3, the results are similar for graphematic confusions and typing errors, whereas accent errors, which are few in the evaluation sentences, behave differently. For these errors, the correct form is the first suggestion in only 34.5% of cases, but it is among the first five in 58.6%. The list of alternatives offered to the writer thus contains the correct form considerably more often than the first suggestion alone.

Table 3. Results by kind of simulated error.

Kind of error	Simulated	Flagged	Correct form first	Correct form in the first five
Graphematic confusion	214	78.5%	57.0%	65.0%
Typing error	166	79.5%	59.0%	68.7%
Missing accent	29	72.4%	34.5%	58.6%

3.3. Feedback Categories

The category assigned by the system agreed with the kind of error that had been simulated in 88.5% of cases (362 of 409), namely for 188 of the 214 graphematic confusions, 150 of the 166 typing errors, and 24 of the 29 accent errors. Most disagreements (40 of 47) concern slips that produced another existing word, such as *ci* for *chi*. These cases are deliberately presented as contextual suggestions, since telling the writer that an existing word is a typing error would be misleading. The remaining 7 are ambiguous cases, in which a typing slip coincides with a graphematic confusion or with a missing accent.

4. Discussion

4.1. Educational Implications

Interpreting these results through the lens of the initial design principles highlights a crucial balance between trust and caution. Because false alarms are kept to a mere 0.4%, writers can confidently treat any flag as a genuine signal for attention, an essential feature for learners who lack external reference sources. However, this cautious approach involves a clear trade-off, leaving roughly one in five errors unflagged. Consequently, the environment must not be presented as a foolproof guarantee of correctness, nor should the absence of flags be interpreted as a flawless text. Instead, educators should frame the tool explicitly as a support mechanism for the revision process.

Beyond building trust, the results strongly support the decision to offer a list of alternatives rather than automatically replacing words. While the correct form appears as the top suggestion in approximately 80% of cases, it is present within the

top five options over 90% of the time. Consequently, an automatic replacement feature would introduce errors in a substantial number of instances, whereas providing a list keeps the correct form accessible to the writer. In a formative environment, what might be considered a weakness in a standard spellchecker transforms into an opportunity to compare forms and make informed choices. This process fosters the active engagement with feedback that the literature identifies as essential for learning (Koltovskaia, 2020; Nicol & Macfarlane-Dick, 2006). This dynamic is especially clear with accent errors, where the top suggestion is often inaccurate, yet the correct form remains readily available among the broader alternatives.

Furthermore, in most cases the system's explanations reflect the specific type of spelling slip. By naming the underlying pattern, this feedback addresses the writing process rather than just the isolated word (Hattie & Timperley, 2007). It provides learners with a metalinguistic vocabulary to reflect on their spelling and supports self-assessment through a revision summary that highlights recurrent difficulties. This grouped data can simultaneously help teachers adapt their instruction, aligning with the principle that formative assessment should inform both learning and teaching (Black & Wiliam, 1998).

On a broader cultural level, the environment integrates local linguistic resources into digital practice, presenting Sardinian as a language equipped for contemporary tools (Galla, 2016; Rivoltella, 2017). However, applying a spellchecker to a language with evolving spelling conventions requires caution to avoid penalizing legitimate local varieties. To ensure the tool values linguistic diversity rather than enforcing a rigid norm, it features an adjustable threshold and an option to keep words as written. Ultimately, the system leaves the final decision to the writer and should be introduced in educational settings as just one of several reference tools.

4.2. Limitations and Future Work

The evaluation presented here establishes a first technical baseline, and its limitations indicate where the work should go next. The errors were simulated with the same rules used to train the models, so the results show how the environment handles known error patterns, while its behaviour with the errors that learners actually make, which may be less predictable, remains to be verified. The evaluation set is also small, with only 29 accent errors, and all the sentences come from a single literary work, which offered a controlled vocabulary and register for this initial phase. As a consequence, for about one error in ten the correct form was not in the word list and could not be suggested, which points to the need to expand both the evaluation data and the word list to more varied contexts. Finally, the environment has not yet been used by learners and teachers, so its educational effects are still to be studied.

Moving from technical validation to ecological testing represents the primary focus of upcoming work. Classroom experimentation will be essential to investigate how users interact with the feedback, answering key pedagogical questions regarding how explanations are interpreted and whether learners actively weigh alternatives. To support this participatory refinement alongside teachers and students, the interface will be localized from English into Italian and Sardinian. By analyzing real usage logs, future studies will evaluate usability, student engagement, and long-term metalinguistic awareness. Ultimately, because this approach relies on accessible

elements like a modest text collection, a word list, and basic spelling rules, it can be extended to other Sardinian varieties and low-resource languages, although this extension has not yet been tested.

5. Conclusions

This contribution presented an AI-supported writing environment for Campidanese Sardinian, designed to transform spelling correction into an opportunity for formative feedback. Instead of automatically replacing words, the system flags potential errors, explains the underlying issues, provides alternatives for comparison, and summarizes recurring difficulties, ensuring the writer retains full agency. A preliminary evaluation using simulated data demonstrated that this feedback is sufficiently cautious and reliable for educational use. Correct words are rarely flagged, and when suggestions are provided, the correct form is almost always included among the available options.

While these initial findings validate the technical framework, the pedagogical impact remains to be explored. Establishing whether this type of feedback genuinely enhances self-regulation, metalinguistic awareness, and writing confidence requires direct engagement with learners and teachers, which constitutes the project's next phase. More broadly, this study suggests that AI writing support for minority languages need not replicate the standard autocorrect tools built for majority languages. By centering the writer's autonomy and leveraging local linguistic resources from the outset, technological innovation can effectively support an inclusive, culturally situated approach to language education.

References

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, *5*(1), 7–74. https://doi.org/10.1080/0969595980050102

Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education*, *43*(8), 1315–1325. https://doi.org/10.1080/02602938.2018.1463354

Carta, S. M., Fenu, G., Giuliani, A., Manca, M. M., Marras, M., Podda, A. S., & Pompianu, L. (2025). Enabling Sardinian language learning via word translation and speech synthesis based on artificial intelligence. *Journal of Inclusive Methodology and Technology in Learning and Teaching*, *5*(3).

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)* (pp. 4171–4186). Association for Computational Linguistics. https://doi.org/10.18653/v1/N19-1423

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, *74*(1), 59–109. https://doi.org/10.3102/00346543074001059

Galla, C. K. (2016). Indigenous language revitalization, promotion, and education: Function of digital technology. *Computer Assisted Language Learning*, *29*(7), 1137–1151. https://doi.org/10.1080/09588221.2016.1166137

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81–112. https://doi.org/10.3102/003465430298487

Koltovskaia, S. (2020). Student engagement with automated written corrective feedback (AWCF) provided by Grammarly: A multiple case study. *Assessing Writing*, *44*, Article 100450. https://doi.org/10.1016/j.asw.2020.100450

Kornai, A. (2013). Digital language death. *PLoS ONE*, *8*(10), Article e77056. https://doi.org/10.1371/journal.pone.0077056

Lobina, B. (2004). *Po cantu Biddanoa*. Ilisso. (Original work published 1987)

Moseley, C. (Ed.). (2010). *Atlas of the world's languages in danger* (3rd ed.). UNESCO Publishing.

Ng, H. T., Wu, S. M., Briscoe, T., Hadiwinoto, C., Susanto, R. H., & Bryant, C. (2014). The CoNLL-2014 shared task on grammatical error correction. In *Proceedings of the Eighteenth Conference on Computational Natural Language Learning: Shared Task* (pp. 1–14). Association for Computational Linguistics. https://doi.org/10.3115/v1/W14-1701

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, *31*(2), 199–218. https://doi.org/10.1080/03075070600572090

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

Regione Autonoma della Sardegna. (2022). *Repertorio grafematico sperimentale per la certificazione della lingua sarda* (Allegato alla Deliberazione della Giunta Regionale n. 18/13 del 10 giugno 2022).

Rivoltella, P. C. (2017). La Media Education. In *L'agire didattico* (pp. 247–258). ELS La Scuola.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, *18*(2), 119–144. https://doi.org/10.1007/BF00117714

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, *78*(1), 153–189. https://doi.org/10.3102/0034654307313795

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice*, *41*(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2