

Scramble to Understand: A Conceptual Framework for Gamified, Ai-Explained Word-Reordering Practice in English Language Teaching

¹Rajashree K, ²R. Shobia

¹Ph.D. Research Scholar, Department of English, School of Engineering and Technology, Dhanalakshmi Srinivasan University, Tiruchirappalli, Tamil Nadu, India

²Associate Professor, Department of English, School of Engineering and Technology, Dhanalakshmi Srinivasan University, Tiruchirappalli, Tamil Nadu, India;

DOI: <https://doi.org/10.51244/IJRSI.2026.1308000298>

Received: 05 September 2026; Accepted: 10 September 2026; Published: 22 September 2026

ABSTRACT

In this paper, a conceptual plan for an English language learning activity is presented in which the familiar jumbled-word format is maintained as the main activity and what happens to the learner who mistakenly answers a jumbled word is altered. The suggested design involves providing a group of jumbled words to a learner and allowing them 1-5 attempts to correct and type the words into a grammatically correct sentence. Correct answer within the allotted number of attempts is accepted and the learner proceeds to the next problem, as is done with conventional scramble games. The difference is that when there is an incorrect order, an open-source large language model (LLM) is called that analyzes the learner's order, pinpoints the grammatical principle violated, explains the violation in plain language, and demonstrates how the sentence should be corrected. This paper does not present an implementation, prototype or experimental trial, but is a conceptual proposal to encourage discussion among ELT practitioners, curriculum designers and educational technologists of the possibility that explanation-driven feedback could be overlaid upon an already popular form of gamified feedback. It places the idea in the context of the existing literature on gamification in language teaching (L2L), corrective feedback (L2CFF), and the new large language model (LLM) technology for grammar explanation, and examines its potential for mapping onto the four core language skills: Listening, Speaking, Reading and Writing (LSRW).

Keywords: Gamification and jumbled-word games (JWGs), scrambled sentences; corrective feedback; large language models; open-source LLM; English language teaching; LSRW skills

INTRODUCTION

For decades, exercises that rearrange words to form sentences have been part of the English language classroom. The simplest form of a learner is to give a group of words in a random order and ask them to reseed the words into a meaningful and grammatically correct sentence. Teachers prefer the format because it is easy to set up, can be modified to almost any grammatical concept and lends itself well to features like timers, points, and levels that are common in games. When these exercises are presented in the context of game elements, then they are often referred to as gamification, which is the use of game design elements (points, levels, badges, timers, competitive and cooperative types) in a non-game context like language learning.

Popular in spite of its well-known drawback, the traditional jumbled-word game provides limited feedback to the learner when they provide an incorrect ordering. Very few implementations will tell you that the response is incorrect, and they will often disclose the correct response, but not explain why the learner's response was ungrammatical, nor what the rule of English syntax, morphology or usage was violated. A learner is therefore able to keep on failing the same type of item over and over without ever understanding this pattern, as the game tells them that they are wrong, without explaining to them why they are wrong.

In this paper, I suggest an approach that can fill that gap, but does not make any changes to what already works well with jumbled-word gamification. The format of scrambled word and the limited number of attempts are preserved, and the game-like setting is maintained. What is added is a layer of automated individualised grammar explanation, created by an open-source large language model when a learner's attempt is incorrect. The learner still has up to 5 attempts per item and successfully submitting at any time during those 5 attempts is considered successful. However, after each incorrect attempt (and, of course, by the fifth attempt) the learner is given a brief, non-technical description of the particular grammatical error in his/her sentence, and a specimen of the correct sentence is given. The goal is, of course, that this one scrambled-sentence scenario should be able to teach the learner a transferable grammar principle, not just whether he or she more or less successfully guessed a set grammar principle with one of the words.

The boundaries of this paper need to be precisely drawn. No data collection, piloting or evaluation with learners has occurred and no built system, working prototype or experimental study with learners is reported. It's an idea paper: a formal statement of the pedagogical and technical idea that will be read, commented on and developed by ELT practitioners and education technologists. The rest of this paper summarizes the literature relevant to the proposed mechanic, provides a detailed description of it, elaborates its pedagogical foundations in terms of the LSRW framework, and outlines potential benefits and open challenges.

REVIEW OF RELATED WORK

The present proposal is based on the following strands of research: research on gamified scrambled word and word order exercises in the teaching of English; research on corrective feedback in L2 acquisition and more recent work on using LLM to provide explanatory feedback to grammatical errors.

Gamified Scrambled- Word Activities

Scrambled word or jumbled word games have already been proven to be the widely used method for gamification in the English language teaching field. In guidance to practising teachers, other word-challenging exercises such as word scramble (where words are scrambled, scrambled words are put in order in a certain amount of time) and other word games such as matching flashcards and role-play simulations are often given. In addition, the technique has been directly tested in studies carried out in the classroom, such as Ramadhanti and Pratiwi's study (2025) which was conducted in grade XI EFL students in junior high school, and found that the technique was able to increase the mean score of students' grammar skills, which went from 61 to 73, and the students showed more involvement, participation, and grammatical accuracy in writing arguments. There are other general studies on gamification in HELS which have also demonstrated that word-ordering and sorting tasks (such as sorting words into grammatical categories) can improve students' grammar competence and that the improvement is only marginally greater than that of the conventional grammar teaching using textbooks.

Overall, this literature suggests that scrambled-word is one of the known and popular class activities used in ELT classrooms and that people do learn from using this type of class activity, although not necessarily in a very different manner from that of other non-gamified types of class activities. This second is another sign that it's not only the game that is crucial to maximising learning, but the feedback provided upon taking an incorrect action.

Corrective Feedback in Second Language Acquisition

In L2 research, it is found that L2 corrective feedback can be implicit (e.g. recast by presenting the correct L2 form without any metalinguistic comments) and explicit (e.g. a metalinguistic explanation indicating that the L2 form is wrong and the L2 rule that the learner has violated). Explicit metalinguistic feedback has been found to be more effective than implicit feedback in promoting learners' control of grammatical structures, and to yield more lasting effects (Ellis, Loewen, & Erlam, 2006). This is relevant to the present proposal because it uses a jumbled-word-game that gives only the correct answer as a weak implicit form of feedback, whereas a system that explains the grammatical principle that underlies the error is used as the more powerful,

explicit, form of feedback that the second language acquisition literature associates with better learning outcomes.

In recent years, attempts have been made to encode such explicit metalinguistic feedback explicitly in computation. One way is to enhance the grammatical error correction system by providing the learners with an explanation of the error in natural language, since it is believed that an explanation will facilitate the internalization of one corrected sentence, instead of merely memorization (Fei et al., 2023). In this paper a feedback mechanism is proposed, directly inspired by this concept of correction and explanation.

Large Language Models as Grammar Tutors

As a direct result of the development of large language models (LLMs) which are now readily available, automated personalized feedback for language learners has seen a resurgence of interest. In the aforementioned review, LLM-generated feedback has been shown to be on the rise with more than just error feedback and is now being used to correct issues of organisation and content, suggesting a typology of the roles it can play in writing feedback, including error-correcting agent, formative evaluator, and revision scaffold. But feedback that is part of an explicit and coherent instructional context, and not just an isolated correction, is generally related to greater learning, something that favors embedding LLM explanations in a bounded-attempt game instead of providing them as a stand-alone resource, the same review noted.

An aspect of the feedback-comment generation task that has been studied explicitly is feedback comments explaining a grammatical concept behind an error, such as: "You need to change x (direct object) to y (direct object) because x takes a direct object," and "You cannot use x as a direct object, use y instead. Feedback-comment generation has been studied explicitly as a particular type of feedback, in which a grammatical concept is explained behind an error, for example: "You need to change x (direct object) to y (direct object) because x takes a direct object" or "You cannot use x as a direct object, you must use y instead. The results of studies involving adult learners of L2 grammar with the use of chatbots have also shown that the timing of feedback is important, that is, immediate feedback after each incorrect response is more likely to be related to more positive L2 learning outcomes and learner experience, while delayed feedback is more likely to result in negative outcomes.

Open-Source Models as a Deployment Choice

It is not only based on an open-source LLM, but also on the literature. It should also be noted that open-source model families can be hosted locally (which would mean no per-use licensing fees) and that this might also ensure that information about learners stays on the institution's own infrastructure, rather than being passed to a third-party API (especially when users are minors, or when an institution has very strict data-governance requirements). The work on small, locally-trained open-source models for tasks that confront educators has demonstrated the potential of models at this scale, along with other techniques such as retrieval-augmented generation, to be useful in the context of a narrow task, such as reasoning about a limited set of grammatical structures, where open-ended generation is not required for the task, and is encouraging in the context of the grammar-explanation task.

The LSRW Framework

Finally, the proposal is introduced in the context of LSRW, the model in which the development of language skills is organized in listening, speaking, reading and writing in the English language. While there is not a total consistency in the foundation and newer writings on LSRW, all agree that grammar is a connective tissue which runs through all four skills and that all four skills are interconnected and interdependent when taught together are not taught separately (as is expressed in the discussion of LSRW in open and distance education and the general discussion of LSRW's role in language pedagogy). The proposed activity is text-based because it involves Reading and Writing directly, and the correction-explanation cycle offers a natural way of connecting Listening or Speaking to Reading and Writing in the extensions. These are discussed in more detail in Section 5.

The Proposed Idea

The core concept behind this paper is to maintain the same structure of the jumbled word game as learners are already familiar with and instead of a failure moment, have an automatic generation of a grammar moment, based on an open source LLM that automatically generates a grammar moment according to the error that was made.

Core Mechanic

The main part of the activity consists of a target sentence that is chosen by a teacher, or from a graded item bank, like e.g. third person singular verb agreement, adverbs of frequency, sequencing of auxiliary verbs in a question. The target sentence is scrambled and presented to the learner who is asked to type in the word sentence in the correct order, or drag and drop word tiles in to the correct word sentence. They are made available in both modes since learners may have different preferences for adopting the mode, and they may be different types of devices available for use. Typing may be more appropriate on a device that includes a keyboard, while rearranging the tiles with drag & drop may be more approachable for younger learners or those using touch-screen devices.

The Five-Attempt Structure

A learner will have no more than 5 attempts to do each scrambled item. It does two things from the above gamification literature: it eliminates the possibility of the learner who is having trouble with the learning task making endless guesses, but it also keeps the familiar and motivating experience of the learner having only a few "lives" (attempts) to use before he or she is game over. The 5 attempts of the flow are:

- If the learner guesses an attempt is correct, the learner is congratulated, is awarded points according to the number of attempts he or she has previously made (first attempt is worth more points than fifth attempt is, etc.) and then moves to the next scrambled item.
- It doesn't just penalize the attempt; it has a number of attempts left to go for each incorrect attempt. There is a countdown that decreases the number of attempts for every incorrect attempt. It passes the learner's submitted word order, as well as the target sentence and the grammar point surrounding the item to the open-sourced LLM, and it gets back an explanation of what exactly is wrong with the learner's ordering, in a very brief statement. The learner reads this explanation and is asked to repeat the action again, but this time, he or she is asked to make an effort to apply the rule he or she just read, not to look things up!
- Fifth attempt is wrong, the learner doesn't get the 5th attempt. Instead, the system shows him the entire explanation, as well as the right sentence, and the student leaves the item with the explanation of why he has been wrong in his reasoning, and an example of what a correct answer should look like; then the game moves on to the next item.

The Role of the Open-Source LLM

The task is quite limited and focused for the open-source LLM: given a target sentence, a grammar points in this sentence, and a learner's incorrect order, the LLM is supposed to identify the targeted grammar point that the learner's order is incorrect in, state the grammar point in a learner-friendly way so that it is not in technical jargon, and present the corrected sentence. It resembles the following task of feedback-comment generation that is investigated in the computational linguistics literature in Section 2, in which a system explains that a certain verb does not agree with a direct object, but requires a preposition to be followed by a noun or noun phrase. With an open-source system (as opposed to a closed commercial API), an institution can deploy the model on its own infrastructure, extract learner interaction data and can avoid any recurrent expenses that could be substantial in large classrooms or in the event of a per-query charge.

Since the grammar point of each item can be specified in advance by the teacher who is creating the item bank, the task for the LLM does not ask it to diagnose the error, but simply to provide an explanation that the teacher has specified is the grammar point of the item, the scope of what the LLM has to reason about is greatly reduced, and the risk of it coming up with an out-of-scope or incorrect explanation is also greatly

reduced. The literature review has informed this design because it has indicated there was more accuracy in the feedback for LLM when the feedback was presented within a structured lesson plan rather than without.

Illustrative Scenario

Let's consider a scrambled word that contains an adverb of frequency, like the following scrambled tiles: "drinks / she / in / the / always / coffee / morning /". If a learner doesn't recognize the position of the adverbs of frequency, such as always, never, etc., in front of the main verb in the sentence, the system would display their sentence and ask them to input a response to the sentence, to the given grammar point (adverb placement), and to the student's input. The explanation returned by the LLM can be returned as natural language, such as "The adverbs of frequency, such as "always", are usually placed in front of the main verb, so we have to reorder "drinks always" to "always drinks". The student would then give it a shot again, this time with an explanation and may succeed on the second or third attempt, or after 5 unsuccessful attempts may be given the full corrected statement and explanation. The use of the same general explanation of how adverbs work should help sentences with the same sentence pattern not only with the one that is "correctly" explained, but with all other sentences with the same pattern; this is the primary pedagogical assumption of this concept: that a single scrambled sentence scenario, which is "correctly" explained, will make it less likely that the same type of error will be repeated later in the classroom.

Pedagogical Rationale

Mapping the Activity to LSRW Skills

The LSRW framework does not consider Listening, Speaking, Reading, and Writing as discrete skills but rather as interconnected and sees grammar as the common thread that weaves through each skill. In a simple version the activity described here is a Reading and Writing activity: the learners read a set of scrambled words and the LLM's explanatory feedback and they write, or reconstruct a grammatically correct sentence. This is important because Reading and Writing are frequently referred to in the LSRW literature as reinforcement skills which support grammatical accuracy and structural awareness skills alongside the primary communicative skills of Listening and Speaking. The design can also be broken down into the specific LSRW skill that is emphasized at each stage of the activity so that a teacher can determine at a quick glance what the item is training:

Reading (R): Learner reads the jumbled word tiles, and most importantly after an incorrect attempt, the learner reads the grammar explanation generated by the LLM, which is itself a short reading-comprehension activity that the learner must do to understand the explanation.

Writing (W): learner writes (or reconstructs) a sentence, using the productive skill of putting words together to form a grammatically correct unit, either by typing or arranging tiles.

Listen (L): In an extended version of the concept, the scrambled words could be presented as an extended audio clip of individual words rather than as printed text, and the learner would have to listen to and retain each of the words in his/her mind before he/she could reorder them, thereby moving the emphasis of the activity from Reading to Listening.

Speaking (S): Likewise, an extended version may request the learner to both type and say the reconstructed sentence; in this case, it will also involve pronunciation and fluency, in addition to grammar, while the learner speaks the sentence out loud. If each item – or each variant of an item – could have the LSRW skill or skills it highlights attached, it would be possible for a teacher to deliberately balance a lesson or a course over time across all four skills – instead of it being a default setting by the activity of Reading and Writing. This is included in the tagging here as a design consideration of the idea and not built or tested.

From a Single Error to a Transferable Rule

The pedagogical logic for the five attempt, explanation-based approach is directly taken from the difference between implicit and explicit corrective feedback that was discussed in Section 2. The explanation of the error

is not just a corrected word order, but also identifies the grammatical principle at work, so that the learner is introduced to a rule that should, in principle, apply to the sentence he is working on as well as to a wider range of sentences. In this sense, as described in the original idea for this paper, a learner will not be given flaw for similar scenarios: If a learner is told how to use an adverb in a particular sentence, he will be better able to use it in a similar sentence later in the same session or in a subsequent session. It is an empirical question whether such a transfer effect does exist and to what extent; this paper only argues that the design is set up in a manner that would facilitate transfer, according to the literature on corrective feedback.

Motivation through Bounded Challenger and Visible Process

The idea of retained game elements for higher engagement and more active participation among learners who have been exposed to scrambled-word games, as well as points or badges being linked to the rate of a learner's item solving, and a sense of progression through an item bank, are motivated by maintaining the advantages already documented for gamified grammar practice. Explanations are supposed to be very brief, and actionable, eg, for an error that a learner made, rather than just a general explanation of the error, and immediately to the learner's point, allowing a learner to read the explanation in a few seconds and try the activity again without slowing the game down or putting too much burden on the learner to try to understand the explanation; the goal of the explanations is not to slow down the game or to make it feel like a lecture.

Foreseen benefits and open challenges

Anticipated Benefits

- Each wrong answer does not have to be a dead end, but a mini-lesson, which may further enrich the teaching benefit gained from each scrambled word.
- Unlike static, decontextualised correction, explanations are tailored to the specific error actually made by the learner, rather than being a grammar note that is more or less suitable for all learners (as found in existing LLM-feedback research to be associated with better learner engagement).
- A concern often discussed in the literature on using LLM in schools is the continued costs associated with large-scale classroom deployment; with an open-source model, the costs of hosting it are not recurring, and the institution maintains control over learner interactions data.
- The limited number of tries (5) and bounded format provide a clear set of rules and game like structure to the activity, as opposed to a potentially discouraging drill.
- Tagging by grammar point and by LSRW skill provides teachers a clear and transparent method for planning for balanced practice across all the Reading, Writing, and, in extended versions, Listening and Speaking skills.

Open Challenges and Questions

- The reliability of automatically generated explanations: in some cases, an automatic explanation generated by an open-source LLM might be imprecise or overly technical, or even linguistically incorrect in rare instances, so a review or moderation process, or restricting the LLM with a pre-approved template of explanations for each grammar point, would likely be necessary before deployment.
- Some scrambled sentences can be scrambled in more than one way to be a grammatically correct sentence, meaning that item design and the logic that would be used prior to sending the item to the LLM would have to consider that acceptable alternative orderings exist.
- Computational resourcing: even a small open source model needs useful hardware to be served by many concurrent users with low latency; a point made in recent comparisons of various size language-models locally hosted in school environments.
- Evaluating actual learning transfer: This paper has drawn from the existing research on corrective feedback that explanation-based feedback should help the transfer to similar sentences, but this needs to be tested directly, for instance, by comparing the performance with novel sentences containing the same grammar point as the previously explained ones.

- Learner fatigue of authorial explanations: Explanations may become too lengthy or too frequent, risk undermining the pace which makes gamified practice attractive in the first place, and therefore should be carefully designed for instructional purposes and, finally, tested with real learners.

CONCLUSION AND FUTURE DIRECTIONS

This paper is presented as an idea, not a developed system that requires testing but is instead an approach to adding automated, LLM-generated grammatical explanation to the tried-and-tested format of gamified, jumbled-word English practice. The proposed design features the same scrambled word concept and the same limited 5 attempts concept while changing a right or wrong verdict to a targeted explanation of why the given word is wrong and how to make it correct, with the goal of making every incorrect answer a learning opportunity for a transferable concept of grammar instead of a wasted life in a game. The deployment option of the explanatory feedback, based on an open source large language model, instead of a proprietary cloud service is proposed as a more realistic and sustainable solution for education institutions, like schools and language institutes, that are cost conscious and have concerns about the privacy of student data.

The concept is also compatible with the LSRW framework, which is prevalent in English language teaching in its basic form, and extends the activity with spoken word tiles and/or by having learners pronounce the answers aloud could integrate Listening and Speaking into the same framework. There is no new empirical data for this paper, and its conclusions about engagement, motivation, and learning transfer are based upon current research on gamification, corrective feedback, and explanations generated by LLM. This paper leaves a number of challenges as future areas of research: using a small prototype of the five-attempt, LLM-explained approach, pilot it with a group of learners, and test the approach for a small number of grammar points to see if the explanation-driven approach yields the transfer and engagement benefits that this paper has argued it should.

REFERENCES

1. "Bigger Isn't Always Better: Choosing the Right Size Large Language Model for Locally Hosted School Settings." *Applied Sciences*, vol. 16, no. 11, 2026, <https://doi.org/10.3390/app16115268>.
2. "Enhancing Listening, Speaking, Reading and Writing (LSRW) Skills through Open and Distance Learning (ODL) in English Language Teaching (ELT)." ResearchGate, 2020, <https://www.researchgate.net/publication/343319800>.
3. Ellis, Rod, Shawn Loewen, and Rosemary Erlam. "Implicit and Explicit Corrective Feedback and the Acquisition of L2 Grammar." *Studies in Second Language Acquisition*, vol. 28, no. 2, 2006, pp. 339–368.
4. Fei, Yunchang, et al. "Enhancing Grammatical Error Correction Systems with Explanations." *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Association for Computational Linguistics, 2023, pp. 7489–7501.
5. "Gamification in TEFL." EBC TEFL, 2025, <https://www.ebcteflcourse.com/gamification-in-tefl/>.
6. "GEE! Grammar Error Explanation with Large Language Models." arXiv, <https://arxiv.org/pdf/2311.09517>.
7. "Language Models in the Classroom: Bridging the Gap between Technology and Teaching." Stanford HAI, 2025, <https://hai.stanford.edu/news/language-models-in-the-classroom-bridging-the-gap-between-technology-and-teaching>.
8. "Language Skills (LSRW) and Their Role in Language Teaching." Exam Affinity, 2026, <https://examaffinity.in/language-skillslsrw/>.
9. "LLM-Generated Feedback in L2 Writing: A Scoping Review." *Education Sciences*, vol. 16, no. 8, 2026. <https://doi.org/10.3390/educsci16081196>.
10. "On the Application of Large Language Models for Language Teaching and Assessment Technology." arXiv, 2023, <https://arxiv.org/pdf/2307.08393>.
11. Ramadhanti, Siti F., and Dwi Pratiwi. "Enhancing English Grammar Mastery through Gamification: The Impact of Scrambled Games on Secondary School Students' Writing Skills." *JOLIT Journal of*

-
- Languages and Language Teaching, vol. 13, no. 2, 2025, pp. 900–911. <https://doi.org/10.33394/joltt.v13i2.14440>.
12. “Small Models, Big Support: A Local LLM Framework for Educator-Centric Content Creation and Assessment with RAG and CAG.” arXiv, 2025, <https://arxiv.org/pdf/2506.05925>.
 13. “Understanding LSRW Is Language Learning.” Teacher Tribe, 2025, <https://www.teachertribe.world/post/understanding-lsrw-is-language-learning>.
 14. “Use of Gamification in English Learning in Higher Education: A Systematic Review.” Journal of Technology and Science Education, 2023, <https://www.jotse.org/index.php/jotse/article/view/1740/723>.
 15. “Why Open Source Large Language Models Are Popular.” Open Source For You, 2026, <https://www.opensourceforu.com/2026/04/why-open-source-large-language-models-are-popular/>.