Analysis of University STEM Students’ Mathematical, Linguistic, Rhetorical–Organizational Assignment Errors

María-José Arévalo 1, María Asun Cantera 2,*, Vanessa García-Marina 2 and Marian Alves-Castro 3

Abstract: Although Error Analysis (EA) has been broadly used in Foreign Language and Mother Tongue learning contexts, it has not been applied in the field of engineering and by STEM (Science, Technology, Engineering, and Mathematics) students in a systematic way. In this interdisciplinary pilot study, we applied the EA methodology to a wide corpus of exercises and essays written by third-year students of mechanical engineering, with the main purpose of achieving a precise diagnosis of the students’ strengths and weaknesses in writing skills. For the analysis to be as exhaustive as possible, the errors were typologized into three main categories (linguistic, mathematical, and rhetorical–organizational), each of which is, in turn, subdivided into 15 items. The results show that the predominant errors are rhetorical–organizational (39%) and linguistic (38%). The application of EA permits the precise identification of the areas of improvement and the subsequent implementation of an educational design that allows STEM students to improve their communicative strategies, especially those related to the writing skills and, more precisely, those having to do with the optimal use of syntax, punctuation, rhetorical structure of the text, and mathematical coherence.

Keywords: Error Analysis; integrated STEM education; writing skills; technical and linguistic competences

1. Introduction

This study explored the application of the methodology of Error Analysis, most commonly used in the context of second-language acquisition (SLA), to the written production of engineering students. We have observed that students in technical branches are prone to believe that oral and written communication are subsidiary to the accuracy and correctness of the technical content of their subjects. This leads to carelessness in their written (and oral) production. In order to tackle this problem, it is first necessary to make a precise diagnosis of the students’ needs. EA may be a useful tool to identify and analyze their errors in order to eventually propose didactic solutions. Before proceeding, some theoretical background on EA is in order.

1.1. Error Analysis: Theoretical Background

Theoretically based on generative linguistics and cognitive psychology, EA (Error Analysis) emerged in the late 1960s and became fully established in the 1970s [1]. EA is, according to the scholar who contributed to its establishment and development, “the study of erroneous utterances produced by groups of learners” [2]. Years later, Cook [3] defined EA as “a methodology for dealing with data, rather than a theory of acquisition”. EA is therefore not conceived as a theory of language, but rather as a methodology, a technique. Crystal´s [4] definition, used as starting point in many studies, is more explicit...
and includes the object of study and the procedure. It reads as follows: “a technique for identifying, classifying and systematically interpreting the unacceptable forms produced by someone learning a foreign language, using any of the principles and procedures provided by linguistics”.

Before proceeding, it must be said that the errors which are identified, classified, and interpreted in EA are not to be mixed up with mistakes. Researchers draw a clear line between the two notions. Both are unacceptable forms, but while mistakes can be attributed to the learners’ lack of attention, carelessness, or the failure to apply a rule he/she knows, an error is more systematic and may be due to a gap in the learner’s competence. The innovation brought in by EA, with respect to previous methodologies, was that the mother tongue (MT) loses the prominence it had in the identification of errors. EA assumes that errors can be fully described and explained in terms of the target language, independently of the mother tongue of the learner.

The notion that the language learner develops his/her own linguistic system at different stages of the learning process is critical to EA. References [2,5–7] speak about a “transitional competence” developed both in the process of acquiring the first language and in second-language acquisition. In the case of the mother tongue, this system is not the adult language; in the case of learners of a foreign language, it is not the target language. Selinker [8] referred to this mental grammar developed by learners at intermediate stages of the learning process as the Interlanguage. EA compares this interlanguage with the target language in order to find mismatches, i.e., errors.

Corder [7] claims that this interlanguage can be described linguistically and errors will play a crucial role in that enterprise. They are no longer regarded as something negative in the process of learning a language, they are rather regarded as evidence that the learners are testing hypotheses about the linguistic system of their target language. Studies on SLA have tended to focus on learners’ errors as they help to predict the difficulties involved in acquiring a second language. In this sense, errors become part of the learning process, as they provide valuable information (i) to design teaching methodologies, programs, and materials and (ii) to inform researchers about the process of acquisition [9].

Upon detection of the error [10], the methodological procedure which is activated is as follows [5]: (i) identification of errors in their context, (ii) classification and description, (iii) explanation and identification of the source of each error, (iv) assessment of the severity of the error, and (v) search for a possible solution. One of the most relevant achievements of EA is, therefore, to address the mistakes made by students in a systematic way. In this way, the areas of greatest difficulty and the different causes of errors are identified during the process of learning a foreign language, thus avoiding the fossilization of these errors [11,12].

However, this methodology also has some limitations and has been harshly criticized. By the mid-1970s, it showed already some signs of stress, as more and more voices were raised against EA [13]. One of the most important pieces of criticism was the fact that some mistakes may go unnoticed due to the so-called inhibition phenomenon, i.e., learners may avoid the use of certain language structures that they do not master in order not to make mistakes. As a result, these potential errors are dodged [13]. It may likewise be counterproductive to focus exclusively on the evaluation of mistakes, since by paying attention only to what students fail to achieve, what they do well may be underestimated or even overlooked [14], which may eventually lead to the frustration and demotivation of students. Finally, identifying a single source for an error is often difficult, if not impossible. Despite all the criticism, EA has not been abandoned. A quick search in any database will show a profusion of publications.

Although most widely used in the field of learning a second language, the evolution of EA is marked by its expansion to other areas of language education: mother-tongue literacy, assessment of oral and written production; language disorders and therapy work; and forensic linguistics [1].
Outside the realm of linguistics, EA has been extensively applied to the diagnosis and remediation of mathematical errors. Radatz [15] offers a systematic categorization of errors built around the difficulties arising from information processing. His influential classification has been adopted and modified in subsequent years by authors like References [16–20], among many others. Three major categories of errors are usually identified: (i) those referred to as the calculations themselves, (ii) those having to do with the technical description (and the use of the language specific to mathematics), and (iii) those involving decision-making strategies. The latter two were of crucial importance for the students participating in our study (more details in Section 2.3.1, below). In this area of study, like in linguistics, the importance of exploiting the educational potential of errors is often stressed [21–23]. It is also worth mentioning that researchers have very often focused on the difficulties involved in the resolution of word problems, which are caused to a great extent by the linguistic formulation of the exercise. The interested reader is referred to the complete literature review offered in Reference [24]. A very recent paper deals with the connection between mathematical errors and reading comprehension [25]. The authors find evidence that the poor understanding of the exercises is sometimes the source of mathematical errors. EA has likewise been applied in learning situations in which the language of instruction is the students’ mother tongue and at different educational stages, including university education [26]. In the case of Spanish, interesting studies are presented in References [27–30], just to mention some recent ones. In all of these studies, the authors analyze mainly linguistic aspects of the written production of university and pre-university students in subjects of the humanities.

In the case of STEM (Science, Technology, Engineering, and Mathematics) students, EA has greatly contributed to the identification and classification of the errors made by students fundamentally in writing and in the use of argumentative and rhetorical strategies [31,32]. In this sense, methodological approaches such as WID (Writing in Disciplines) [33], or WAC (Writing Across the Curriculum) [34], which consist of applying an instruction, developing the practice and proceeding with feedback, have played a fundamental role. All the studies in the field emphasize the importance of the optimal rhetorical organization of the texts written by students of engineering [35]. EA has also served as a basis for supervised feedback applied to thesis drafts and their presentations [36], mainly in English. The importance of developing abilities that go beyond the purely technical and theoretical contents of the different STEM disciplines is underlined in many papers (see, for instance, Reference [37] and the very interesting literature review of articles published between 2000 and 2018 in Reference [38]).

1.2. Our Study

In this study, we analyzed the written production of Spanish-speaking STEM students. The corpus to which EA was applied is quite extensive, as it consists of 79 written exercises and seven essays with an average 8020 words per participant. Our main aim was to identify the errors made by students in order to shape the feedback given in the classroom, to modify the contents of the course (if necessary) and to adapt our teaching strategies. All the efforts should lead to improve the students’ writing skills, but also to raise their awareness that being able to communicate the results of their work is crucial. The final objective will also be to prepare a rubric for future courses.

In order to achieve these ends, we applied the methodology of EA, in that we revised all the exercises in the corpus manually and identified and described the errors (see the categorization of errors in Section 2.3, below). They were systematically classified in the three main categories mathematical, linguistic, and rhetorical–organizational, which were in turn divided into subcategories, as explained below.

In the mathematical area, students were given precise instructions pertaining to the importance of developing a coherent explanation of all the exercises they had to solve (also of the practical cases in the essays). Deviances from these instructions were considered
errors. Of course, any mathematical errors having to do with calculations and formulas were also computed.

In the linguistic area, all the deviances from standard Spanish in terms of grammar and vocabulary were computed as errors. We also considered repetition of structures or words, and informal language a problem, as students had been instructed on the style of academic writing. Failure to comply with the standard rules of spelling and punctuation were also considered errors.

Students had been provided with a detailed description of the tasks (exercises and essays) concerning both the structure of the contents and techniques to guarantee the quality of the graphs and images that might be included in the different exercises. They had likewise been given a style sheet including the constituent parts of an essay (table of contents, etc.). Deviances from those rules and descriptions were considered errors in the rhetorical–organizational area.

In order to facilitate both the students’ and the professors’ work, a rubric was designed that compiled all the potential errors divided in the areas just mentioned [39]. Using this rubric appropriately the students could avoid, and the professors detect, the errors.

We are aware of a study by Conrad [40], in which she undertakes the analysis of a corpus of texts written by graduates in engineering and practitioners in the field. She includes in her study linguistic features (sentence structure, word choice, grammar, and punctuation) and genre organization (rhetorical moves in technical memos). Conrad’s analysis covers logical organization, grammar and communicative features such us conciseness, directness, standard spelling, grammar, and accurate word choices. Her conclusions are intended to be useful to develop materials that will be used in courses to help bridge the gap between the writing skills of engineering graduates and the skills required in the workplace. In our study, which shares the aim of improving materials and teaching strategies with Conrad’s analysis, we, however, focus on undergraduate students and the potential gaps in their training. Our study is likewise more extensive in that it covers more linguistic and rhetorical aspects than Conrad’s and, it incorporates the study of mathematical errors. Our corpus is also different, as it is made up of the written production of students in two different tasks (exercises and an essay).

To the best of our knowledge, never before has the EA methodology been applied in the field of engineering and STEM in a systematic way, encompassing rhetorical–organizational, linguistic, and mathematical aspects and Spanish as the language of expression. The present research looks at these three areas, using a typology of 15 errors in each. This theoretical analysis is implemented by systematically analyzing numerous texts written by several groups of third-year students of the Degree in Mechanical Engineering at the University of the Basque Country.

The aim of this study is twofold. First, we tried to find out if the data collected come to confirm that a complete evaluation of STEM students should involve linguistic, rhetorical, and mathematical analysis. Second, we tried to assess the suitability of EA for the elaboration of an accurate diagnosis of the writing needs of STEM students (in particular, of undergraduate engineering students), with the final objective of helping them to enhance their writing skills and for their professors to elaborate appropriate materials.

2. Experimental Details
2.1. Participants

The participants were 28 Spanish-speaking students enrolled in the course “Machine Design”, as just mentioned. All the students were in their third year of the Degree in Mechanical Engineering at the Faculty of Engineering in Vitoria (University of the Basque Country, UPV/EHU). In this course, students worked in groups of 4, which were formed by the professor at the beginning of the semester. Students were randomly allocated to groups by their instructor without following any specific criterion. The basic aim of this allocation procedure was to avoid that students of similar profiles (in terms of intellectual
and working capacity) would converge in the same group, thus minimizing the effects of this variable on the final results.

Prior to the data collection, students were informed about the application of the EA methodology to the submitted assignments. They were informed of the fact that the exercises were going to be marked twice: once by the professor of the subject in the usual manner and a second time by another professor of the degree who was not involved in their evaluation. The students’ personal information (in this case, just their names) was removed from all the exercises and only the code assigned to the group was kept for future reference. The students were also informed that the application of this EA methodology would not have any influence on their final grade, and that, by participating, they could help future students, as the results might be used for diagnostic purposes. All the students gave their consent to participate in the project. At the end of the course, still pending a detailed analysis of the data, general information on preliminary impressions was presented to the students. It was also emphasized that, by making this effort, students would understand the demands of workplace writing.

2.2. Data Collection

The data collected for the study consisted of a number of written assignments, produced by the student in the course of the second semester of the academic year 2019/20, that is, from January to June 2020. Due to the lockdown ordered by the authorities as a consequence of the pandemic, six weeks had a face-to-face format (40% of the total), and nine weeks were online (60%).

The first assignment consists of the exercises corresponding to each didactic unit covered in the course. Students were required to write an introduction of between 50 and 100 words and a detailed explanation of the problem or exercise of between 150 and 250 words. The groups had to hand in between 11 and 14 exercises. The variation in the number of tasks given to the different groups was determined by their complexity. Thus, on occasion, a group received two relatively simple exercises to compensate for the more complex task given to the other groups. The students had one week to complete each exercise and submit it via the virtual platform Moodle. This part of the collection of data was done during the first 10 weeks of the course.

The second assignment was an essay in which the students had to describe the design of a part of a machine. In this case, they were required to write a longer piece of writing of approximately 10,000 words. The machinery parts assigned to the different groups, by drawing lots, were (in parentheses, the letter-code assigned to each group): brakes (A), bearings (B), belts (C), clutches (D), bevel gears (E), springs and permanent unions (F), and screws and non-permanent unions (G). In the essays, an original and systematic development was required (including the revision of theoretical foundations), as well as practical applications of the different elements being analyzed, explaining various real cases of operation through proposed and solved exercises. The students had 5 effective weeks, from week 10 to the end of the semester.

The instructions given to the students were as follows:

1. Everything must be expressed in written form in a way that should be legible and understandable by a STEM student (in the first years).
2. The bibliography must include all the sources used.
3. The information to include is as follows:
   - Title of the essay and name of the students in the group;
   - Table of contents;
   - Introduction;
   - Description of the element, operation, classification, applications. If possible, videos of the item working can be attached;
   - Calculations: Propose 2 or 3 solved and explained examples of calculation and design so that they are understandable for self-study. In the case of the simplest
elements (such as bearings, screws, belts, and springs), a greater number of exercises should be presented;
- Conclusions;
- Bibliography.

Just as in the case of the exercises, the submission of the essays took place via Moodle. The groups were also asked to do a 15-minute oral presentation on the topic of their essay. It must be noted that, on receiving each assignment, students were reminded of the importance of their written expression, and of the fact that any technical student should be able to understand the explanations they provided simply by reading the text presented.

In this course, the medium of instruction is Spanish, so that all the interactions in the classroom are carried out in this language. It is therefore the language used in the lectures, the discussion of the tasks, and to provide feedback.

2.3. Categorization of Errors

The EA methodology requires a categorization of errors [41]. In the present study, 15 errors have been identified and categorized in each of the following areas: mathematical, linguistic, and rhetorical–organizational. The process was as follows. Careful reading of the exercises and essays submitted by the students allowed us to identify the errors, which were recurrent in the documents presented for assessment. This task was carried out by two of the authors of this article, the professors responsible for the course in which the data were collected. After that, the errors were classified according to their type. We agreed that the best course of action would be to continue with the analysis of the errors in separate groups, each concentrating on the category of errors in their area of expertise (mathematical, linguistic or rhetorical). The decision to include 15 errors in each category was taken in order to calculate the statistical value in a simple manner and to achieve a certain balance in the analysis of errors, thus guaranteeing that the three areas under study were given the same quantitative weight.

The tables displaying the typology of errors used in the present empirical study are presented in the remainder of this section.

2.3.1. Mathematical Errors

Before going into the classification of mathematical errors, a brief account of the specific character of the course “Machine Design” is in order.

This course deals with the practical application of basic mechanical concepts to specific cases. The technical–mathematical side of the subject is not so much about obtaining the right result for a specific calculation. The students are assumed to be able to calculate correctly, and they are closely monitored by their professors in tutorials. Although there is a certain amount of calculation involved, this is not the most important aspect of the course.

It is true, however, that in some of the exercises, the student has to dimension a machine, i.e., the student has to provide the minimal acceptable dimensions, which may involve the choice and purchase of different parts of the machine. The good designer must find and propose the optimal solution, which must comply with the technical requirements specified at a reasonable price, considering also aesthetic, environmental, and social factors. The key question, and the aspect which is evaluated, is the student’s capacity to explain the suitability of the solution proposed.

Students are also required to calculate the safety coefficient of the machine. In this case, it is vital that the student provides a comprehensible explanation of basic concepts of static calculation and fatigue-based component life estimation. There is likewise a business aspect to the subject, as the machine designers must be able to “sell” their product, i.e., they must be able to persuade potential purchasers that their solution is appropriate and optimal.

All of these requirements of the subject make the development of writing and oral skills essential and also justify the classification of errors made in this section, which mostly involve problems with notations, conversion of units, etc.
The mathematical errors, found in the sample, coded as MAT followed by a number, are described in Table 1. They can be subcategorized according to the following aspects:

a. Mathematical coherence in the management of formulas and units (from MAT-1 to MAT-9). They include errors in the system of units and their conversion, numerical calculations, errors in mathematical notation. As can be seen, this category is the one containing the highest number of errors, up to 9. After a first evaluation of the corpus, it was already observed that there were a large number of errors related to mathematical coherence. Given their number and variety, nine subcategories of these errors were proposed.

b. Technical description of operation and applications (from MAT-10 to MAT-12): These are errors in the interpretation of data and their incorporation in the formulas used, omissions in definitions of the parameters used in the design, use of tables, or experimental graphs, or design standards.

c. Decision-making (from MAT-13 to MAT-15): These errors are related to omissions in the explanation of the calculation method used, as well as omissions or mistakes in admissible convergence criteria in iterative mathematical processes.

| Mathematical Coherence: Managing Formulas and Units | Technical Description of Operation and Applications | Decision-Making |
|-----------------------------------------------------|--------------------------------------------------|-----------------|
| MAT-1 No International System of units              | MAT-10 Does not define uses, applications, restrictions | MAT-13 Type of design under static or dynamic loads |
| MAT-2 No English System of units                    | MAT-11 Does not define basic design criteria. Does not identify simplifying hypotheses | MAT-14 Justification of the applied design method (Goodman, Soderberg) |
| MAT-3 No engineering notation                       | MAT-12 Does not explain the coefficients of graphs or tables | MAT-15 Does not define convergence criteria in iteration problems |
| MAT-4 Decimals: non-significant digits               |                                                  |                 |
| MAT-5 Numerical calculation error                    |                                                  |                 |
| MAT-6 Unit conversion error                          |                                                  |                 |
| MAT-7 Does not enter / describe formula              |                                                  |                 |
| MAT-8 Dimensional analysis of wrong formulas         |                                                  |                 |
| MAT-9 Angle error (radians, degrees)                |                                                  |                 |

2.3.2. Linguistic Errors

Linguistic errors, coded as LING, are described in Table 2. They have been divided into three large categories:

a. The first one includes errors concerning the syntax of the sentence. Questions such as the complexity of the sentences, the use of connectors, and the coherence in the use of verb tenses are considered here (from LING-1 to LING-5).

b. The second group delves into the richness in the use of vocabulary (from LING-6 to LING-8).

c. In the third group, the grammatical and ortho-typographical errors are considered (from LING-9 to LING-15).

The linguistic aspect is quite complex, as it comprises numerous parameters. For this reason, we decided to make an exhaustive subclassification in this pilot study (15 types). This had the advantage of allowing us to gather very specific information and to shape our interventions in a manner that would guarantee improvements in all types of errors.
Table 2. Description of linguistic errors.

| Syntax | Lexicon | Grammar, Spelling, and Punctuation |
|--------|---------|-----------------------------------|
| LING-1 No connectors | LING-6 Poor lexicon, repetitions | LING-9 Spelling mistakes |
| LING-2 Person changes, incoherence (me/us) | LING-7 Colloquial and/or oral terms | LING-10 Accent errors |
| LING-3 Little variety of connectors | LING-8 Words empty of content | LING-11 Punctuation errors |
| LING-4 Verb tenses not coherent | | LING-12 Uppercase errors |
| LING-5 Syntactic errors that make understanding difficult | | LING-13 Subject-predicate concordance errors (anacolutha) |
| | | LING-14 Phrases without conjugated verb |
| | | LING-15 Lack of determiners, prepositions and/or conjunctions |

2.3.3. Rhetorical–Organizational Errors

The rhetorical–organizational errors are coded as ORG and are subdivided into two categories, as shown in Table 3:

a. A visual one, which refers to the clarity of reading and the organization of all graphic material in relation to the layout of the page, graphs, paragraphs, and graphic representation (from ORG-1 to ORG-8).

b. A second one, closely related to the rhetorical organization of the text, the coherence of the discourse and the interpretation of results, and its adaptation to the academic context (from ORG-9 to ORG-15).

Table 3. Description of rhetorical–organizational errors.

| Clarity of Reading-Layout of the Page | Rhetorical Organization of the Text (Logical Organization of Discourse). Interpretation of Results |
|--------------------------------------|-------------------------------------------------------------------------------------------------|
| ORG-1 In disorder | ORG-9 Does not answer the question |
| ORG-2 No paragraphs | ORG-10 Does not identify the type of analysis |
| ORG-3 No page numbers, no references to tables/figures | ORG-11 Does not identify simplifying hypotheses |
| ORG-4 Schema error | ORG-12 Does not interpret the results |
| ORG-5 Table error | ORG-13 Negative results without explaining incompatibilities |
| ORG-6 Aesthetic errors | ORG-14 Irrelevant conclusions |
| ORG-7 Visual errors and poor supporting material | ORG-15 Bibliography errors |
| ORG-8 No reference system, no references, no schemas, and no graphs | |

3. Results and Discussion

In this section, we present and discuss the results obtained, after applying the EA methodology to the tasks submitted by the seven groups of students participating in the study. The results about to be presented have been compiled and classified by type of error. The results obtained in the exercises and the essays are kept separate throughout.

3.1. Mathematical Errors

With the categorization of mathematical errors described in Table 1, above, the errors made by each of the seven groups in each of the exercises submitted in a period of 10 weeks were counted manually. Figure 1a, below, displays the total number of mathematical errors in the three subcategories described in Table 1 which were found in the exercises. This information is completed with the detail of the number of errors across groups in Figure 1b.
Mathematical errors, are associated with the technical description of operation and application. Certain mathematical aspects are taken for granted and explanation is omitted, even from MAT-13 to MAT-15, which represent 18% of all mathematical errors, the most common one is error no introduction or description of the formula. It quite often happens that the reader of their exercise needs explanations as to why the formula or the appropriate notation is introduced. It often happens that certain mathematical aspects are taken for granted and explanation is omitted, even though the need to include an explanation is indicated in the pre-submission instructions.

These are the most significant results. Errors associated with mathematical coherence and formula management (from MAT-1 to MAT-9) are clearly predominant in the sample. They represent 56% of all mathematical errors. Among them, the most common one is error no introduction or description of the formula, followed by MAT-3 no use of engineering notation. In the subject of “Machine Design”, the proper management and explanation of the formulas to be applied is a complex aspect, since it requires the integration of diverse technical knowledge. The use of various unit systems also adds considerable difficulty to the task. Nevertheless, students in the third year of the degree should be accustomed to explaining the formulas used, as well as the parameters involved. Formulas in mechanical engineering very commonly include sub-indexes, coefficients to the power of numbers, etc., and students should use these notations properly. The fact that they made errors involving these aspects may indicate that the students did not learn how to write formulas in previous years. These repeated errors may be due to the fact that students are not aware of the fact that the reader of their exercise needs explanations as to why the formula or the appropriate notation is introduced. It quite often happens that certain mathematical aspects are taken for granted and explanation is omitted, even though the need to include an explanation is indicated in the pre-submission instructions.

The second most frequently occurring errors in this category, representing 26% of all mathematical errors, are associated with the technical description of operation and applications (from MAT-10 to MAT-12). In particular, the error typified as MAT-10, which encompasses the non-explanation or definition of the uses, applications or restrictions, stands out from the rest. In this case, students show a total lack of habit when it comes to explaining or describing relations or applications. They develop the exercises without writing an introduction explaining the conditions under which the formulas applied may be used. This aspect is closely related to the previous one and reflects the lack of awareness on the part of students concerning the need to explain in detail the formulas being used.

As to the third category of mathematical errors, those related to decision-making (from MAT-13 to MAT-15), which represent 18% of all mathematical errors, the most common one is MAT-13, associated to the type of design under static or dynamic loads. Once again students seem to have difficulties to choose the correct expression to be used in a given exercise. They seem to reproduce what they have heard in the lectures without thinking about whether it is appropriate to the context in terms of both register and content.

Considering now the results across groups, in Figure 1b above, it can be seen that groups A and C account for 43% of the total mathematical errors, each making more than 20% of the mathematical errors, while groups F and G are the ones that made the fewest errors (10.6% of the total), making less than 4% and less than 7% each. Finally, the other groups (B, D, and E) contribute with circa 15% each.

Turning now to the essays, the number of mathematical errors made in this task is notably lower than in the exercises (see Figure 2, below). This may be due to the fact that the essays were written in weeks 10 to 15, after students had received feedback on the
mistakes made in the exercises. Although not so many as in the exercises, there are some mathematical errors in the essays. To them, we turn now.

**Figure 2.** Accumulated mathematical errors corresponding to the essays (a) by categories and (b) by groups.

In the category of mathematical coherence and formula management, MAT-1 *not using the International System of units* is the predominant error, with MAT-3 *not using engineering notation* in the second position. It is worth noting, in this respect, that the essays were written during the lockdown, so students did not have access to books in the university library. In order to compensate for this inconvenience and potential problem, the professor provided them with a list of references, some of which used the English System of units. The students did not make the effort of converting the units to the International System. In the subcategory associated with the technical description of operation and applications, the error that stands out is MAT-10 *does not describe uses, applications and restrictions*. Finally, in the last category, decision-making errors, there are not really noteworthy errors, but MAT-13 *related to the type of design under static or dynamic loads* can be mentioned as the predominant error in the category.

It is also worth noting that the distribution of errors is not homogeneous, as group A accumulated 43% of the total errors, whereas, group E made no mathematical errors.

All in all, as was the case in the exercises, the predominant error continues to be in the category of mathematical coherence and formula management, where MAT-1 *not using the International System of units*, and MAT-3 *not using engineering notation* stand out as the most frequently occurring errors, followed by MAT-10 *does not describe uses, applications and restrictions* of the technical description category.

The conclusion that can be drawn from these results is that the engineering students in our corpus seem to master the technical mathematical aspects, but not the coherence required to show and explain the technical results in an optimal way.

### 3.2. Linguistic Errors

The quantitative analysis of the linguistic errors in the exercises shows that their number is much higher (327) than the mathematical errors (188). The accumulated results are displayed in Figure 3. The essentially technical content of the course and the lack of habit of reinforcing the linguistic aspect of the tasks presented may be two of the factors behind the high figures.

Considering the categories of linguistic errors defined in Table 2, above, it can be seen that the repetition of errors is distributed almost homogeneously across the three categories analyzed: structure of the sentence, lexical selection, and grammar in general.
It is significant that both in the first category, syntax, and in the second, lexicon, the greatest number of errors reported are those related to LING-6 lexical poverty, accumulating (22.6%) of the total, and LING-3 little variety of connectors (14.1%). In other words, in most exercises, the vocabulary is poor and repetitive. Thus, for example, the term calculate “to calculate” is repeated quite often in some of the exercises. The same lack of richness is observed in the case of linking words. These results reveal what might be considered one of the main linguistic problems of STEM students: the limited use (and command) of vocabulary in their mother tongue. This may be aggravated by their reluctance to read books on the topic of design. Many students seem to be satisfied with the materials provided in the course and do not consider it necessary to read any additional materials. It has also been noticed that students have no previous experience in writing longer compositions. As they lack the skill, they make the effort to build a given structure once and, afterwards, they just repeat it, without trying to find synonyms or alternative phrasings. Besides, they show no intention of producing rich texts, but only present the results of the exercises without paying attention to the form.

In spite of the lack of command of vocabulary, the lexicon used is in general appropriate to the academic field. What is not so appropriate is the register used by the students in some texts, which is far too informal for the academic context. Thus, expressions like sacamos “we get” or cogemos “we take” appear frequently in the explanation of the problems.

Concerning the syntax, the exercises analyzed are usually quite concise and very often do not present complex sentences, which translates into LING-1 a lack of connectors or the already mentioned little variation in their use. By means of illustration, the only causal connector used in the sample was ya que “as, because”. This fact shows once more the lack of interest or habit in writing technical texts.

Turning finally to the third category of linguistic errors, it must be mentioned that, although spelling errors are not very common, which could be attributed to the spell-check utility of word processors, we found sentences whose ungrammaticality was related to the lack of an inflected verb, as well as serious deficiencies in punctuation, which is the most frequently repeated error (LING-11 punctuation errors). As is well-known, auto spell-check does not specifically correct punctuation errors.

In sum, the most recurrent errors in the exercises were lexical poverty and lack of punctuation, both an indication of poor previous training in writing and lack of command of the grammar and lexicon of their native language.

Turning briefly to the results displayed in Figure 3b, above, it can be seen that the groups that made the highest number of linguistic errors were B (20.8%), and C and F (17.7%, each). These three groups accumulate more than half of the total errors (56%). Group A and group E follow closely with 15.6% and 15.3% of the errors, respectively. Group G (0.9%) is the one with the lowest number of errors.

As in the case of mathematical errors, the number of linguistic errors decreases significantly in the essays (see Figure 4a, below). The number of errors found (36) is remarkably low given the amount of written production demanded.
We attribute the higher linguistic quality of the essays (when compared with the exercises) to three main reasons. First, the students now had the experience and training of the previous exercises. As mentioned before, the exercises had been discussed in class and the students had received feedback. Second, the students probably regarded the exercises as simple class tasks, but perceived the essay as a more significant piece of work, which made them be more careful in its elaboration. In this respect, it must also be mentioned that the relative weight of the two assignments in the final grade was different, 10% in the case of the exercises and 30% in the case of the essay. Finally, the use of the spell-checking tools of the word processors most probably played an important role in the correction of the task.

The most common error considering all the essays is the error coded as LING-6 related to the poor and repetitive lexicon (22.6%). Phrases like “there are two kinds of bearings” (group B), “this element” (group B) are often repeated and no attempt is made at using synonyms. The second error is LING-3, little variety of connectors (14%). The students used, as already mentioned in the discussion of the exercises above, the causal connector ya que “because”, without considering alternative linking devices. The third most frequently occurring error is LING-11 with 13% of the total, which is consistent with the results obtained in the exercises, where the most frequent errors were LING-6 lexical poverty (14%) and LING-11 punctuation error (14%). There, person shifts in pronouns (me, we) were also attested, 11%.

Concerning the group results, it is clearly seen in the graph in Figure 4b, above, that groups C and D accumulate almost 40% of the total errors, whereas groups B, E, and F contribute each with 16.7%. In this case, group G made no linguistic errors (0%).

All in all, the results show that the deficiencies and the proportion of errors are of the same characteristics in the exercises and in the essays.

3.3. Rhetorical–Organizational Errors

In the analysis of the results concerning organizational errors, it can be seen that some are far more common than others. Figure 5, below, shows the accumulated results in the exercises.

In the exercises, three were the most frequent errors, all of them in the category of rhetorical and logical organization. The first one is the failure to identify the type of analysis they have to carry out ORG-10, (19%). Students fail to do the data collection and the interpretation. They are used to receiving this part of the task from their lecturer in the classroom, and, in their autonomous work, they seem to consider it unnecessary to include this part. Generally speaking, the exercises are not badly planned but students do not explicitly identify details such as the type of resistant, static, or dynamic analysis, or which machine element is affected. The second most frequent error is the failure to identify the simplifying hypotheses or the definition of the parameters that appear in the formulas used in the exercise ORG-11 (21%). Students are unable to identify the simplifying hypotheses which accompany the formulas, which often leads them to make the wrong decisions. In general, values increased by certain safety coefficients are estimated, which leads to oversizing, as...
well as to higher-than-actual stress results. Thus, without providing the definitions of each parameter, it is not easy to know whether students understand the exercise or if, on the contrary, they are just following the pattern of the exercises done in class.

The third error in frequency of occurrence is the **non-interpretation of the results** ORG-12 (19%), that is, in the sample, students tend to provide the numerical result without interpreting its meaning or repercussion. For optimal learning, it would be beneficial for students to get used to discussing whether the resulting stresses are high or small, whether the mechanical element will withstand these stresses, whether any part of the element could be resized to adjust the stresses to the admissible limit value in order to increase the use of the material, etc.

Slightly less common are **support material** ORG-7 (15%) and **aesthetics** ORG-6 (11%) errors. These errors are related the fact that the arrangement of information on the page is chaotic, and that students often complete some tasks included in the essay in notebooks, instead of using the text editor, and take a picture of them. The photograph is then scanned and added to the document. From an aesthetic point of view, these images are often of poor quality (dark or blurred). Another error found in the sample is ORG-3 (7.7%), as students do not correctly reference scanned figures from works by other authors. Finally, a note on some errors which also appear in the exercises, although not so frequently. One is ORG-2 **lack of explanatory paragraphs**, and two further occurrences of different aspects of ORG-3 the **lack of pagination in the documents** and also the **lack of references to tables or figures**.

Errors found on very specific occasions in the sample include the following: first, ORG-1 **Lack of order**, the information provided was not properly organized in some of the exercises; second, ORG-8, related to the graphic representation of information **without scales or reference systems**. Without scales it is impossible to know the size of the object represented. Finally, the third type includes both ORG-4 **errors in schema** and ORG-5 **errors in tables**.

Focusing now on the distribution of errors across groups, Figure 5b, above, shows that group D made the highest contribution (19.1%) to the total number of organizational errors. At the other end are groups F and G, making the lowest number of errors (11.4% and 12.3%, respectively). In between, groups C and E, with 14.6% each, are followed by group A with 14.3%. As can be seen, the distribution of rhetorical–organizational errors in the exercises is quite homogeneous across groups, with an average of 14.3%.

The discussion turns now to the presentation of the results of the error count in the essays submitted by the students. In this case, two are the errors that stand out for their frequency: ORG-11 and ORG-12, both errors belonging to the category represented in Figure 6a. Rhetorical and logical organization errors represent 91% of the total. This can be due to the length of the essays.

---

**Figure 5.** Accumulated rhetorical–organizational errors corresponding to the exercises (a) by categories and (b) by groups.
In some cases, there were also inconsistencies in the citation style used. The students do not seem to be aware of the importance of references and of the fact that not providing them correctly constitutes plagiarism. In some cases, there were also inconsistencies in the citation style used.

A further problem often detected in the essays is that the students do not obtain or comment on relevant conclusions (ORG-14). As mentioned in Section 2.2, above, the essays consisted of the theoretical development of a type of machine element. Students received the instruction that, in a section of the paper, they had to show practical applications of that machine element. Although the exercises shown were well developed in most cases, the students copied information from the sources, without drawing any significant conclusions.

Another important error in the essays is ORG-15, that is, students do not define the bibliographic references correctly or sufficiently. In some essays, important details were missing in the bibliographic references. The students do not seem to be aware of the importance of references and of the fact that not providing them correctly constitutes plagiarism. In some cases, there were also inconsistencies in the citation style used.

Similarly, the results obtained after the development of the practical case were not explained (ORG-13). It is evident that students are not used to commenting on results and identifying inconsistencies in them, or drawing conclusions from the data.

As regards errors of type ORG-6, aesthetics, the same problems observed in the case of the exercises, and mentioned above, appear in the essays. Thus, some students included in their paper scanned images of such low quality that they were detrimental to the quality of the final paper.

Finally, error ORG-3, bad pagination or lack of references to tables or figures, is mentioned but this time because it occurs only once in the sample. It must also be noted that this error occurred frequently in the exercises (see the discussion above). It seems that students are more aware of the fact that they have to add page numbers to an essay than to the exercises.

Before moving on to the conclusions, the discussion turns briefly to Figure 6b, which displays the percentages of errors in the essays across groups. Groups A and B make each 17.1% of the total errors. Groups E, F, and G follow with 14.3% each, and finally, groups C and D contribute to the final count with 11.4% each. Notice also that four of the seven groups made no mistakes of the category Clear layout of the page. The use of computers in the essay may have contributed to this result. Notice that in the case of the exercises (discussed above), the errors in this category amount to 41%.

In a nutshell, the most common errors made by students were the lack of initial data collection, not identifying the type of analysis to be done, not answering the question,
and not detailing the assumptions and parameters used in the calculations. A potential explanation for all of these problems is that the students always do the exercises following certain steps (the order familiar from the lectures), and, for this reason, they tend to omit the identification of the type of analysis, as they take for granted that their professor, who will correct the exercises, shares this knowledge. Slightly less typical were aesthetic errors and sufficient data presentation in the tables. Finally, there were errors related to structure and the organization of ideas in paragraphs, as well as lack of paragraph division.

In addition, it must be noted that, in the essays, no exhaustive count of all the individual errors has been performed, as this task was very demanding and not necessary for statistical purposes. Therefore, no matter how many errors of a given type were made, only one was counted.

The lower proportion of errors in the essays is indicative of an improvement in the students’ skills and competence. It can also be interpreted as an indication that students have been successful in the learning process.

Considering all the results in Subsections 3.1–3.3, mathematical errors are not very abundant in the sample (22%), as opposed to linguistic (38%) and rhetorical–organizational errors (39%). This may be due to the fact that mathematics is more central to engineering studies whereas linguistic and rhetorical–organizational aspects, though necessary for their future professional development, are regarded as tangential by the students. It seems to be necessary to introduce in the syllabus teaching strategies and activities that may improve the students’ rhetorical–organizational and linguistic competences. What seems clear is that the different error types require different types of pedagogical intervention. Thus, for instance, error ORG-15 bibliography errors, which appears frequently in the sample, would probably require a brief intervention consisting of giving students some training in reference and citation styles. Other errors, as for instance LING-5 syntactic errors that make understanding difficult, which has to do with a deficient use of syntax, will require a longer-term pedagogical intervention before any real improvement is observed.

4. Conclusions

The main aims of this research were to find out whether the data would confirm the need to include rhetorical–organizational, linguistic, and mathematical analysis in the assessment or STEM students and to test the suitability of EA when it comes to identifying the specific needs of students regarding writing skills. In the light of the results reported on these pages, it can be concluded that both questions have been confirmed.

Application of the EA methodology allowed us to elaborate a complete typology and description of errors in the three areas under consideration. It has been shown that the list of all errors divided by areas and properly typologized makes it possible to achieve a precise and coherent diagnosis of the needs and strengths of the students. Furthermore, EA allows a clear identification of errors, since they are repeated in many of the texts analyzed in the corpus of this study.

The identification of the errors is the first step towards finding a solution. Thanks to the results obtained after the application of this methodology to the analysis of the corpus, rubrics can be created to facilitate the objective assessment of all parameters and pedagogical interventions can be designed to solve the problems identified.

In the present case, one of the conclusions that can be drawn is that STEM teaching, rather than just focusing on technical parameters, should also reinforce aspects pertaining to the linguistic and rhetorical competence of students, thus contributing to the development of their skills in written communication.

Coming down to the specific group studied, the results obtained here allow us to focus on these three points in subsequent intervention in the machine design classroom. As far as mathematics is concerned the objective would be to develop a correct explanation of formulas. The linguistics aspect would improve if students received training in different aspects of language use so that they build correct sentences and use a wider range of
Finally, we must insist on rhetorical organization with the interpretation of the results when an explanation is given. The findings also point at a partial correlation between the three error categories analyzed here. An in-depth analysis of this interesting question is left for future research. We are aware that this study (at this point just a pilot study) has limitations. Thus, for example, the corpus is rather limited, a question that we hope to address in the future, as we intend to gather more data in subsequent courses. The continuation of this study will likewise provide us with a stronger basis of comparison to see the results of the implementation of didactic strategies designed to address the problems detected in the course of this study. We are also open to reconsider the classification of some of the errors in the subcategories established for each area to meet the pedagogical needs of the different groups in which the taxonomy is used. This study has shown that it is possible to identify the strong and weak points in the writings of engineering undergraduates by using the EA methodology.

**Author Contributions:** Conceptualization, M.-J.A. and M.A.C.; methodology, M.-J.A., V.G.-M. and M.A.C.; software, V.G.-M.; validation, M.-J.A. and M.A.C. and M.A.-C.; formal analysis, M.A.-C.; investigation, M.A.C.; resources, M.A.C.; data curation, V.G.-M., M.-J.A. and M.A.C.; writing—original draft preparation, M.-J.A. and M.A.C.; writing—review and editing, M.A.-C.; visualization, V.G.-M.; supervision, M.A.-C.; project administration, M.A.C.; funding acquisition, V.G.-M. and M.A.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Written informed consent has been obtained from the students to publish this paper.

**Data Availability Statement:** The anonymized data used for the analysis of the present study are available from the corresponding author.

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**

1. James, C. *Errors in Language Learning and Use*; Routledge: London, UK, 1998; p. 320.
2. Corder, S.P. *Introducing Applied Linguistics*; Penguin Books: Harmondsworth, UK, 1975.
3. Cook, V. *Linguistics and Second Language Acquisition*; Modern Linguistics Series; Macmillan Education: London, UK, 1993; p. 333.
4. Crystal, D. *A Dictionary of Linguistics and Phonetics*; The Language Library: Wiley: Hoboken, NJ, USA, 2009; p. 529.
5. Corder, S.P. The significance of learners’ errors. *Int. Rev. Appl. Linguist. Lang. Teach.* 1967, 5, 161. [CrossRef]
6. Corder, S.P. Error analysis, interlanguage and second language acquisition. *Lang. Teach. Linguist. Abstr.* 1975, 8, 201–218. [CrossRef]
7. Corder, S.P. Error analysis and interlanguage. *Can. Mod. Lang. Rev.* 1981, 40, 649–650.
8. Selinker, L. Interlanguage. *Int. Rev. Appl. Linguist. Lang. Teach.* 1972, 10, 209–232. [CrossRef]
9. Ellis, R. *Second Language Acquisition*; Oxford Introduction to Language Study Series; Oxford University Press: Oxford, UK, 1997; p. 147.
10. Fernández-López, S. Errores e interlengua en el aprendizaje del español como lengua extranjera. *Didáctica Leng. Lit.* 1995, 7, 203–216.
11. Zhan, H. Frequent errors in Chinese EFL learners’ topic-based writings. *Engl. Lang. Teach.* 2015, 8, 72. [CrossRef]
12. Leonardo, C.L. Las preposiciones en el habla no nativa de nivel intermedio: Análisis de la interlengua basado en corpus. *RNLAEL* 2014, 8, 5–27. [CrossRef]
13. Schachter, J. An error in error analysis. *Lang. Learn.* 1974, 24, 205–214. [CrossRef]
14. Larsen-Freeman, D.; Long, M.H.; Molina-Martos, I. *Introducción al Estudio de la Adquisición de Segundas Lenguas*; Gredos: Madrid, Spain, 1994.
15. Radatz, H. Error analysis in mathematics education. *J. Res. Math. Educ.* 1979, 10, 163–172. [CrossRef]
16. Movshovitz-Hadar, N.; Zaslavsky, O.; Inbar, S. An empirical classification model for errors in high school mathematics. *J. Res. Math. Educ.* 1987, 18, 3–14. [CrossRef]
17. Socas, M. *Dificultades, Obstáculos y Errores en el Aprendizaje de las Matemáticas en la Educación Secundaria*; La Educación Matemática en la Enseñanza Secundaria; Rico, L., Ed.; Horsori: Barcelona, Spain, 1997; pp. 125–154.
18. Astolfi, J.P. El “error”, un medio para enseñar. In *Colectión: Investigación y Enseñanza*; Diada Editora: Sevilla, Spain, 2003.
19. Araya, G.R.; Sánchez, C.M.; Mora, H.R. Errores matemáticos de estudiantes que ingresan a la universidad. Actual. Investig. Educ. 2019, 19, 104–136. [CrossRef]
20. Abrate, R.; Pochulu, M.; Vargas, J. Errores y Dificultades en Matemática: Análisis de Causas y Sugerencias de Trabajo; Universidad Nacional de Villa María: Buenos Aires, Argentina, 2006; p. 198.
21. Borasi, R.; Siegel, M. Reading to learn mathematics: New connections, new questions, new challenges. Learn. Math. 1990, 10, 9–16.
22. Ashlock, R.B. Error Patterns in Computation; Using Error Patterns to Improve Instruction; Pearson: Columbus, NJ, USA, 2006; p. 98.
23. Legutko, M. An analysis of students’ mathematical errors in the teaching-research process. In Handbook for Mathematics Teaching: Teacher Experiment. A Tool for Research; Czarnocha, B., Ed.; University of Rzeszów: Rzeszów, Poland, 2008; pp. 141–152.
24. Daroczy, G.; Wolska, M.; Meurers, W.D.; Nuerk, H. Word problems: A review of linguistic and numerical factors contributing to their difficulty. Front. Psychol. 2015, 6, 348. [CrossRef] [PubMed]
25. Sanz, M.T.; López-Iñesta, E.; García-Costa, D.; Grimaldo, F. Measuring arithmetic word problem complexity through reading comprehension and learning analytics. Mathematics 2020, 8, 1556. [CrossRef]
26. Martín, R.A.M.; Nikleva, D. Análisis de la competencia lingüístico-discursiva escrita de los alumnos de nuevo ingreso del Grado de Maestro en Educación Primaria. Rev. Signos 2016, 49, 48–70. [CrossRef]
27. I Feixas, M.J.; Vinyoles, A.A.; Bargalló, J.I. El uso de evidencias en la escritura científica en los estudiantes de magisterio. Textos Didáctica Leng. Lit. 2013, 64, 40–46.
28. Rodríguez-Gómez, D.; Valdeoriola Roquet, J. Metodología de la Investigación; Universitat Oberta de Catalunya: Barcelona, Spain, 2009; p. 20.
29. Dolz, J.; Gagnon, R.; Mosquera, S.; Sánchez-Abdchi, V. Producción Escrita y Dificultades de Aprendizaje; Graó: Barcelona, Spain, 2013; p. 91.
30. Morales, A.O.; Espinoza, N. El desarrollo de la escritura de estudiantes universitarios. Lect. Vida 2005, 26, 26–37.
31. Berdanier, C.G.P.; Whitehair, C.; Kim, A.; Satterfield, D. Analysis of social media forums to elicit narratives of graduate engineering student attrition. J. Eng. Educ. 2020, 109, 125–147. [CrossRef]
32. Leydens, J.A. Novice and insider perspectives on academic and workplace writing: Toward a continuum of rhetorical awareness. IEEE Trans. Prof. Commun. 2008, 51, 242–263. [CrossRef]
33. Carter, M.; Fierzli, M.; Wiebe, E.N. Writing to learn by learning to write in the disciplines. J. Bus. Tech. Commun. 2007, 21, 278–302. [CrossRef]
34. Ramírez-Echeverry, J.J.; Dussan, O.F.A.; García-Carrillo, A. Effects of an educational intervention on the technical writing competence of engineering students. Ing. Investig. 2016, 36, 39–49. [CrossRef]
35. Fox, J.; Artemeva, N. From diagnosis toward academic support: Developing a disciplinary, ESP-based writing task and rubric to identify the needs of entering undergraduate engineering students. Esp Today 2017, 5, 148–171. [CrossRef]
36. Bastola, N.M.; Hu, G. Supervisory feedback across disciplines: Does it meet students’ expectations? Assess. Eval. High. Educ. 2020. [CrossRef]
37. Daly, S.R.; Adams, R.S.; Bodner, G.M. What does it mean to design? A qualitative investigation of design professionals’ experiences. J. Eng. Educ. 2012, 101, 187–219. [CrossRef]
38. Mazzurco, A.; Crossin, E.; Chandrasekaran, S.; Daniel, S.; Sadewo, G.R.P. Empirical research studies of practicing engineers: A mapping review of journal articles 2000–2018. Eur. J. Eng. Educ. 2020, 1–24. [CrossRef]
39. Canter, M.A.; Arevalo, M.; García-Marina, V.; Alves-Castro, M. A rubric to assess and improve technical writing in undergraduate engineering courses. Educ. Sci. 2021, 11, 146. [CrossRef]
40. Conrad, S. A comparison of practitioner and student writing in civil engineering. J. Eng. Educ. 2017, 106, 191–217. [CrossRef]
41. Schmitt, N.; Hemchua, S. An analysis of lexical errors in the English compositions of Thai learners. Prospect 2006, 21, 3–25.