Teaching the specialized language of Mathematics with a data-driven approach: what data do we use?

Cecilia Fissore[1[0000-0001-8398-265X]], Francesco Floris[1[0000-0003-0856-2422]],

Marina Marchisio Conte[1[0000-0003-1007-5404]], Matteo Sacchet[1[0000-0002-5630-0796]]

[1] University of Turin, Turin, Italy
{cecilia.fissore, francesco.floris, marina.marchisio, matteo.sacchet}@unito.it

Abstract. Numerous studies in Mathematics Education have shown that among the causes of disciplinary learning difficulties are the acquisition and understanding of its specialized language. Data-Driven Learning (DDL) is a didactic approach that treats language as data and sees students as researchers doing guided discovery activities. The exploration of corpora can effectively support reflection on the specialized languages of Mathematics. What data should be used? Students daily use the most recent Large Language Models (LLM) and Google, which can be used for linguistic investigations. However, it must be remembered that there is no total control over the data on which searches are carried out, the results, and the type of language they use. Control over the data is important, especially when teachers want to use these tools to design and deliver didactic activities. This paper presents a recent DDL research activity with 80 secondary school students on the specialized language of Mathematics. The students conducted linguistic investigations on a specially designed corpus and carried out corpus-based activities with automatic formative assessment within a Digital Learning Environment. The results show that the students appreciated the proposed activities. They develop linguistic and mathematical skills and become more aware of the importance of the language they use. Students developed digital skills in browsing, searching, and filtering data, as well as in evaluating data, information, and digital content. LLM, such as ChatGPT, could not be used for the same type of activity, but with appropriate design, they can be used as a starting point for investigation and linguistic reflection. In the future, given the notable diffusion of these AI tools, it is essential to train teachers and students on their strengths and weaknesses and how they influence teaching and learning.

Keywords: Data-Driven Learning, Large Language Model, Language for Specific Purposes, Mathematics Education.

1. Introduction

Every subject, from Geography to Mathematics, has a specialized language, which is the typical way of expressing itself in that area. Research in mathematics education has studied how learning difficulties in Mathematics depend on linguistic difficulties, that is, on the acquisition and understanding of its specialized language [1]. Understanding the language of Mathematics is considered by didactic research to be one of the major obstacles to learning the discipline at all school levels [2]. Furthermore, Mathematics has always been one of the school disciplines where students of all grades encounter the most significant difficulties. For example, the PISA 2022 international triennial survey (to be launched in 2020) shows critical issues in Mathematics for Italian students [3]. Mathematics scores were significantly lower than in 2018 and in previous cycles in almost all types of education. The growth trend seems to have been reversed compared to 20 years ago. For this reason, many researchers in mathematics education study the language that students use in class (with their peers or with the teacher) for questions or collaborative or problem-solving activities that teachers use during explanations, which is present in textbooks, and so on.

It is possible to use the Data-Driven Learning (DDL) methodology to help students develop language skills. DDL treats language as data and sees students as researchers doing guided discovery activities [4]. Students can discover grammatical patterns, word meanings, or other aspects of language by searching linguistic data and investigating large amounts of authentic language. This methodology was born for language teaching and learning but has recently been used for learning specialized languages [5]. DDL is closely linked to the concept of corpus, a collection of texts on which linguistic investigations are carried out.

In recent times, students' use of the Large Language Model (LLM) has spread considerably, also for educational purposes. LLMs are deep neural network models (deep learning) capable of acquiring vast knowledge of a language from enormous quantities of texts, mainly taken from the Web [6]. The impact of LLM on education can be enormous [7]. LLM could change the educational learning goals, activities, and assessment and evaluation practices. The advantages of LLM in education include personalized learning, timely support, continuous assessment, resource delivery, collaboration, and so on [8]. One of the most used LLM by students for learning is currently ChatGPT (https://chat.openai.com/). ChatGPT is just one of many ways to use LLM. Research is progressing rapidly in this area, and new models are constantly being developed [6]. The release of ChatGPT as a chatbot based on GPT models by OpenAI marks a pivotal milestone in the development of chatbots and conversational AI [8]. By conversing with ChatGPT, students also try to study and learn Mathematics. There is a lot of research studying the math skills of ChatGPT (see, for example, [9], [10]). Not as much research studies ChatGPT's linguistic skills in Mathematics. LLM represent a significant step towards integrating natural language understanding and generation in information technology and artificial intelligence, with applications that extend to multiple sectors and scenarios (mathematical, chemical, legal, etc.). However, ChatGPT does not always manipulate the specialized language of these disciplines. LLM, as well as Google, can be used for linguistic investigations [11]. However, it must be remembered that there is no control over the corpus data on which they carry out searches, the results obtained from the searches, and the language they use. Control over the data is important, especially when teachers want to use these tools to design and deliver didactic activities.

This paper illustrates a project connecting DDL with a Digital Learning Environment (DLE) integrated with an Automatic Assessment System (AAS) to develop competencies in the language of Mathematics for Italian secondary school students [12]. The DLE is a stimulating environment that boosts motivation and fosters constructivist acquisition while respecting students' learning pace [13] [14]. In this research, The students conducted linguistic investigations on a specially designed corpus and carried out corpus-based activities with automatic formative assessment within a DLE. We define automatic formative assessment as using formative assessment in a DLE by automatically processing student responses and providing feedback [14] [15]. With the AAS, it is possible to create adaptive questions with personalized, immediate, and interactive feedback giving information about how the DDL task was performed and the process to be mastered, thus enabling self-regulation and self-monitoring of actions [16] [17] [18]. The DDL approach gives students the language access keys to the content and, in the case of Mathematics, proves its effectiveness in helping them to understand and manage a language that can pose actual obstacles to problem solving activities and exercises.

Section 2 presents the state of the art; section 3 illustrates the research activity and methodology; and section 4 shows the main results emerging from the research activity and discusses the implications and conclusions.

2. State of the art

2.1 The specialized language of Mathematics and ChatGPT

Numerous studies in mathematics education have shown that the causes of difficulties are the acquisition, understanding, and management of its language [1]. Over time, Mathematics has developed a specialized language that has become increasingly universal, precise, concise, and effective. This language has its semiological code, which is realized in texts in which technical terms, figures and graphs, and symbolic expressions (equations, formulas, algebraic expressions, etc.) coexist; the latter are sometimes inserted in sentences which, for the rest, use the everyday language [19]. Mathematicians and researchers in mathematics education have different attitudes and points of view towards language, which derive from two apparently divergent aspects: the specificity of Mathematics and its language and the role of the context in mathematical communication [20]. Teaching is communication, and one of its purposes is to encourage student learning, using a language that does not hinder understanding and using metaphors to make concepts understood. At the same time, one of the main objectives is to acquire that specialized language.

Often, the linguistic aspects of teaching a language are considered only in a CLIL (Content and Language Integrated Learning) context, but teachers tend to forget that language education is a task that belongs to the training curriculum as a whole [5]. The characteristics of mathematical language often contrast with the students' linguistic habits, which are difficult to understand [2].

Secondary school students are often forced to deal with the languages of the disciplines without adequate linguistic scaffolding that allows them to make the necessary differentiations and categorizations of the language for specific purposes concerning more general language varieties. Students' linguistic difficulties may concern comprehending or producing verbal texts, symbolic expressions, and figures. In all cases, it is not easy to recognize students' linguistic difficulties because they often compete with other difficulties related to gaps in content, careless mistakes, etc. Language difficulties can cause behaviours that can be identified, such as misinterpretations or the production of incoherent texts, but they can also cause no behaviours, such as not responding, or behaviours that are difficult to interpret, such as random answers [20].

In recent times, the use of LLM by students to learn Mathematics has spread considerably [8]. Consciously or not, this also affects students' learning of its specialized language. However, LLM, such as ChatGPT, do not always use specialized language. If we ask ChatGPT, "Do you know the specialized language of Mathematics?" a response is ", Yes, I'm familiar with the specialized language of Mathematics. Mathematics has its technical vocabulary that includes specialized terms, symbols, and notations" [21]. For example, if we ask ChatGPT, "Can you define in 100 characters what "intersection points" are in Mathematics?" the answer is "points where lines or objects meet, sharing coordinates" [21]. The use of terms taken from a common language, such as "object" or "meet", can confuse students because, in a common language, these terms have different meanings. Furthermore, using the specialist verb "intersect" is missing, which students must acquire. Otherwise, students will have difficulty interpreting the meaning of the concept of intersection and tasks such as "at what point does the function intersect the ordinate axis?".

2.2 Data-driven learning

It is possible to use the DDL methodology to learn specialized languages [12], [22]. The main idea behind DDL is that learners can discover grammatical patterns, word meanings, or other aspects of language by searching for linguistic data. Corpora are proving increasingly influential in language teaching as sources of language descriptions [5]. A corpus is a collection of texts or parts of them in a finite number in an electronic format processed in a uniform way that makes it manageable and searchable using a computer [23]. Students have to deal with a "massive but controlled exposure to authentic input," which is fundamental for language learning [24], and such controlled and contextualized contact fosters more language awareness, noticing, and autonomy.

Tim Johns [4] argues that at the heart of the approach is the use of the machine as a rather special type of informant. Once the informant answers the question, students must make an effort to "make sense of that response and to integrate it with what is already known" [4]. Corpora provide data but do not interpret them: it is up to learners' work and responsibility to evaluate the information found. The easiest way to explore corpus data is through concordancers: end-users may display a list of words with their immediate context. A concordance based on KWIC (keyword in context) can reveal massive information about the language: idioms, collocations, fixed phrases, and frequency data. Since then, literature on the uses and benefits of corpora for language learning has rapidly grown, although there is still little field practice in Italy at least [5].

Recently, the number of corpora and specialised tools available to practising teachers and learners has increased. However, a common criticism is that many of them still require considerable investment in training for learners (and teachers) to understand their rationale and how to use them effectively. Even accepting the potential benefits of a DDL approach, technology is perceived as a major barrier to the implementation of DDL in classrooms around the world [25].

Google itself can be used for linguistic investigations, but positions on the nature of the Web as a corpus are multiple and conflicting [5]. The exploration of web resources as corpus responds to the lack of corpora available to study increasingly complex specific linguistic problems. At the same time, materials age quickly compared to the continuous evolution of language, as well as new technologies and new means of online communication. Problems often cited against treating the Web as a corpus are its unknown size (no finite dimension), ever-changing composition, hidden pages, etc. Its advantages for language teaching include its size (lots of data), timeliness, variety (whatever you want is probably there somewhere), availability (free), reliability (the Web itself doesn't crash or impose limits on the number of simultaneous users), speed, flexibility, and so on. Importantly, it is already familiar to learners, especially via internet search engines such as Google [26], [23]. Another tool that has recently become very popular with students are LLMs, such as ChatGPT. Although ChatGPT is not designed for linguistic research, it does allow users to do so. Much recent research has explored the potential of ChatGPT for language teaching and learning (see, for example, [27], [28]), analysing the digital skills needed by teachers and students to use this chatbot ethically and effectively. Some research explores the role of ChatGPT in DDL, as the world's largest corpus (in a sense) is publicly available for free querying, with an interactive chatbot available to query this corpus quickly and recursively [29]. LLM can also be used for DDL to generate texts at the appropriate language level and to observe repetitions and variations manually or with corpus analysis tools. It is important to note that when conducting linguistic research in a specialised field, designing prompts to produce the desired results is particularly important. Students, for example, might assume that if they talk about “functions”, ChatGPT understands that they are talking about math. However, if we ask ChatGPT, “What are the verbs that have “function” as the subject?”, we get an ambiguous answer like: “Verbs that have “function” as the subject are typically used when describing how something operates or works; like functions, operates, works, performs, runs, executes, etc.” [21]. If we ask, “What are the verbs that have "function" as the subject in Mathematics?" the answer is radically different (see Fig. 1). In this case, ChatGPT interpreted the task correctly, and the results refer exclusively to the mathematics field. As shown in Fig. 1, ChatGPT reports some verbs as a list, followed by an example sentence in which the verb is present. The presence of the example can help students understand in which context it is appropriate to use the verb, and they can reason about its meaning. However, as we have no control over the data, we do not know what determines the choice of verbs and the order in which they are presented. Furthermore, the concept of function is very transversal in Mathematics and is studied from lower secondary school to university. The results become more accurate if we add more details to the prompt.

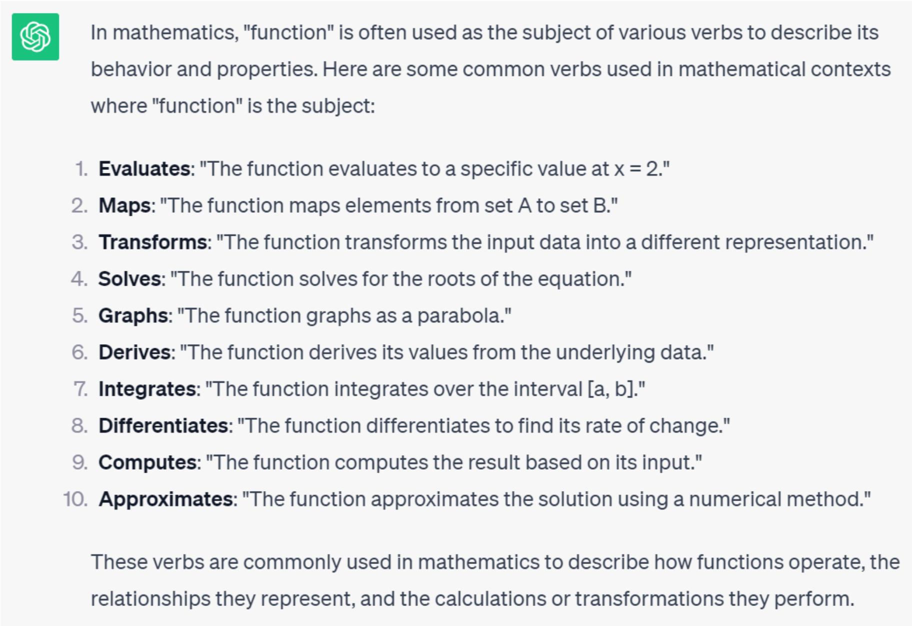

Fig. 1. Example of mathematical linguistic investigation with ChatGPT.

3. Methodology

3.1 Research context

The research activity involved four classes of two Italian secondary schools for 80 students in grade 11 and their teachers. Both schools are science lyceums, secondary schools with a curriculum focused on STEM subjects. It is also not the first experience of the research group in merging STEM assessment and languages [30]. The activities with the students involved two classes as experimental groups, while the other two classes comprised the control group. The students in the control group took a test and a questionnaire before the research activity started and a test and a questionnaire at the end of the research activity. These were the only activities that the control group performed. They did not do any classroom activities with the researchers; they only did normal lessons with their teachers. The experimental group participated also in the research activities compared to the control group.

The didactic experimentation, from November to December 2021, consisted of four two-hour meetings in the classroom. Teachers were involved in selecting topics to be covered and designing the teaching materials to ensure that they were in line with the programme. Teachers were not trained before the start of the experiment, and trainers from the University of Turin conducted the classroom activities. The teachers participated in all the meetings with the students and carried out the activities to study the proposed teaching methodologies. Two specific corpora were created for the activities, selecting texts relating to the concept of function from secondary school textbooks. The first corpus was in Italian, and the second was in English. Corpora were created and validated by the researchers, who verified that the results of the linguistic research were relevant to the planned research activity.

Corpus-based activities with automatic formative assessment are questions implemented within the DLE using the AAS [12]. From a formative point of view, the students have several attempts to answer each question. The questions are designed to guide students in consulting the corpus and to make them reflect on the mathematical concepts. To carry out the activities, students use the AntConc concordance tool (https://www.laurenceanthony.net/software/antconc/) and the DLE (available at the link https://linguaggispecialistici.i-learn.unito. it) side by side. Fig. 2 shows an example of a corpus-based activity with automatic formative assessment. The question is divided into two parts: in the first part, students are asked to search for the word "funzion*" (which means function) in the corpus by inserting 1R, 2R, and 3R as concordance levels and to select the verbs that have "function" as subject; in the second part of the question the students had to complete the proposed statements by inserting the verbs identified in the previous section.

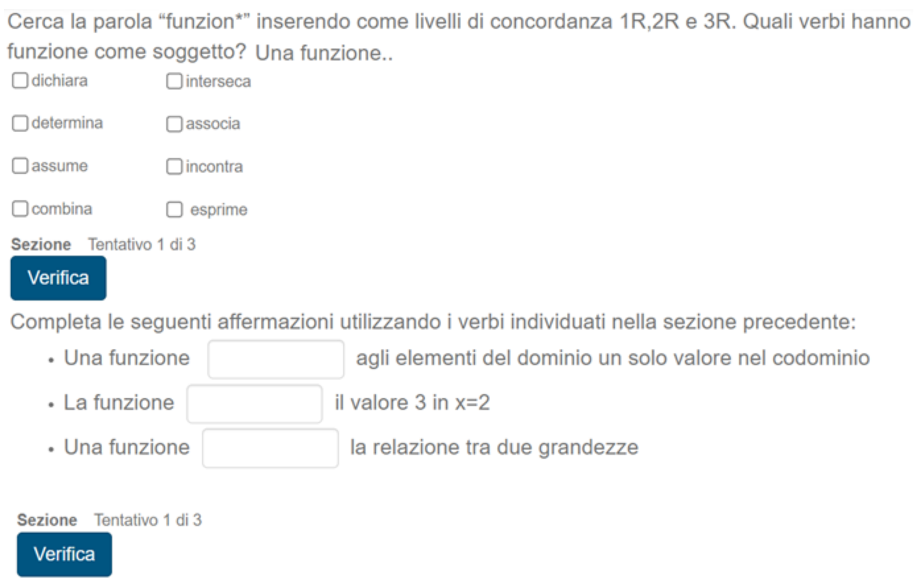

Fig. 2. Example of corpus-based activity.

Students see the second part after answering the first. At the end of each part of the question, there is a "verifica" button (which means "verify"). Students could click this button after answering the question to obtain immediate feedback on the correctness of the answer. If they made a mistake, they could try to answer the question again. In case of an incorrect answer after finishing the attempts, the students see the correct answer that is needed to carry out the next part. To answer the first part of the question, students had to use the AntConc software and enter the word "funzion*" in the search bar. Inserting the asterisk symbol is used to search for both the word function in the singular (“funzione”) and the plural (“funzioni”). Fig. 3 shows an example of what appears on the screen to students after the search. The levels of concordance specified in the text of the question are placed under the search bar and must be set by the students. Through the levels of concordance, one chooses which words to highlight in the search. In this case, by selecting the levels “1R”, “2R”, and “3R”, the search shows all the phrases of the corpus in which the word function appears, and the following words are highlighted: the words one position to the right concerning the searched word (in red), the words two positions to the right concerning the searched word (in green), and the words three positions to the right concerning the searched word (in purple). In step-by-step processes with multiple response attempts, students earn partial credits for the correctness of their answers. These points act as a motivational lever and, by expressing intermediate levels between “incorrect” and “correct”, also offer teachers and students more precise information about the student's competence in a particular domain.

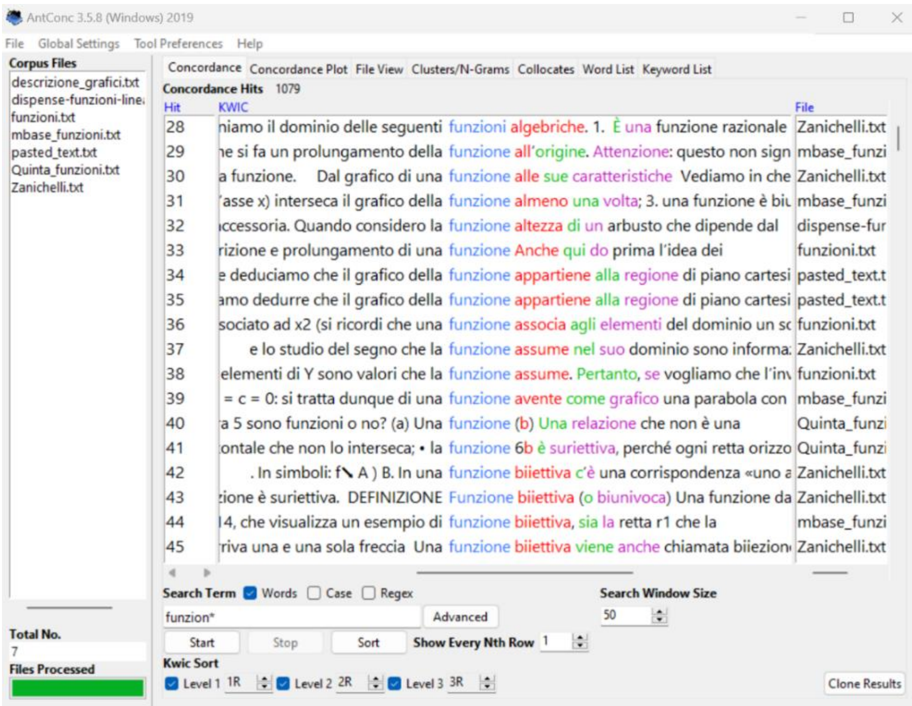

Fig. 3. Example of linguistic investigation with AntConc.

3.2 Data collection and analysis

The research methodology of the didactic experimentation included:

- Initial test (before starting the research activity) and final test (at the end of the research activity) to evaluate students' mathematical and linguistic skills;

- Initial questionnaire (before starting the research activity) and final questionnaire (at the end of the research activity) to evaluate students' opinions about mathematics and mathematics lessons.

All 80 students took an initial test and questionnaire before starting the research activity. The test consisted of ten questions and included linguistic questions and mathematical exercises. The questionnaire was composed of 40 Likert-scale questions (on a scale from "1 = very disagree" to "4 = very agree") covering opinions on Mathematics, mathematics lessons, the study of Mathematics for school, school and extracurricular activities in Mathematics, and students' attitudes towards Mathematics. At the end of the experimentation, all students took a final test structured like the initial test and a final questionnaire. The questions in the final questionnaire were the same as in the initial questionnaire. Questions were added to the questions of the students who carried out the classroom activities to analyze the satisfaction of the proposed activities and methodologies.

This activity was carried out as part of a curricular activity, under the educational responsibility of the schools that provide the instruction in the subject. Students in both groups could choose whether or not to participate in the educational experiment and change their minds at any time. In any case, their decision did not impact their school evaluations and the educational activities proposed to them. The data were collected through learning tests and paper questionnaires, tests with automatic assessment, and digital questionnaires. Once collected, the data were irreversibly anonymised and processed anonymously. The data were digitised, and the paper copies were destroyed. Qualitative and quantitative analyses can be carried out on the anonymised data in order to pursue the research objectives. Only aggregate statistics will be presented.

The research questions are:

- (RQ1) Did students value the methods proposed?
- (RQ2) Did the research activities increase awareness of mathematical language and mathematics?
- (RQ3) Was the impact of the research different for males and females?

Only the responses to the final survey of 35 students who participated in the experimental activity were considered to answer the research questions.

Questions of the final questionnaire concerning the satisfaction with the proposed activities and methods were analysed. In this question, the students had to indicate how much they agreed with various statements on a Likert scale from "1=strongly disagree" to "4=strongly agree". The items concerned:

- The satisfaction with the proposed activities (attention in class, satisfaction, interest in the activities and in Mathematics, appreciation of group activities);
- The usefulness of the activities for the acquisition of awareness of mathematical language and reflection on mathematical content;
- Appreciation of DDL (use of AntConc, interest in text analysis and usefulness for the future);
- Appreciation of automatic formative assessment (immediate and interactive feedback, multiple attempts possible);
- Understanding and appreciation of mathematical topics.

These questions were analysed by reporting the mean and standard deviation. To analyse the correlations between the items, we calculated the Pearson correlation coefficient between all possible pairs of items. In the results, “r(df)” indicates Pearson coefficient with “df” degrees of freedom. To search for useful insights on gender studies, we also investigated the presence of differences in the responses of females and males to the items and the statistical significance of those differences through a two-tailed t-test.

A comparison between the results of the initial and final tests is planned to quantitatively assess the methods' impact, but the analyses are still ongoing.

4. Results

The data that we will analyze consists of the answers to the final survey of 35 students who participated in the experimental activity. The remaining students among the 80 students we considered were part of the control group. Among the 35 students, there are 20 females and 15 males. Students appreciated the proposed methodologies and the mathematical activities, which were very different from the traditional ones. The use of the DDL approach had a very positive and motivating impact on the students: according to them, the activities were practical and not theoretical, interactive and non-transmissive, and based on real data.

Table 1. Mean and standard deviations of the 35 respondents to the final questionnaire

Item	Mean	Std. dev.
I paid attention during the activities	3.51	0.56
The proposed activities were interesting	3.23	0.60
The proposed activities were understandable	3.43	0.61
The proposed activities stimulated my interest in mathematics	2.60	0.60
The proposed activities were useful for gaining awareness of the language	3.14	0.65
I enjoyed the group activities	3.34	0.84
It was easy to use AntConc.	3.09	0.85
It was easy to understand how to use AntConc to answer questions.	3.14	0.85
It was interesting to learn how to use AntConc.	3.40	0.74
Learning to use text analysis software is useful for the future.	3.14	0.91
The exercises helped me to better understand the topics covered.	3.06	0.80
The exercises made me appreciate the topics covered more.	2.63	1.00
The exercises helped me to realize my preparation.	3.03	0.71
It is helpful to view the correct answer immediately after answering a question.	3.63	0.55
When I gave a wrong answer, I would try the exercise again.	3.40	0.69
Having the assessment immediately after solving an exercise helped me understand how to answer the question.	3.51	0.66
The exercises helped me to be more aware of the language to use.	3.14	0.77
The exercises on language also made me reflect on meanings and contents.	2.77	0.77

Table 1 presents an analysis of the student's responses to the question: "Thinking about the experimental sessions, on a scale from '1= strongly disagree' to '4= strongly agree', to what extent do you agree or disagree with the following statements?". The students found the activities interesting (3.23) and understandable (3.43). The activities stimulated their interest in Mathematics (2.60) and raised students' awareness of language (3.14). They appreciated the group discussions and the group activity (3.34). According to students, it was easy to use AntConc (3.09) and to understand how to use AntConc to answer questions (3.14). They also believe that learning to use text analysis software is useful for the future (3.14). According to students, the exercises helped them to become more aware of the language to use (3.14), and the exercises on language also made them reflect on meanings and contents (2.77).

Students appreciated the formative aspect of the activities: the possibility of viewing the correct answer immediately after answering a question (3.63) and retrying the exercise in case of a wrong answer (3.40). They agreed that having the assessment immediately after solving an exercise helped them understand how the question should be answered (3.51). The DDL approach gives students the language access keys to the content and, in the case of Mathematics, proves its effectiveness in helping them understand and manage its language that can pose obstacles to problem solving activities and exercises. Students exposed to DDL engaged in an activity that improved their language skills in their actual working practice. The study of language is concerned not only with the choice of words to use in a more specialized language but also with the terms conveying meanings, and by exploring the language, one explores the meanings. The new and interdisciplinary nature of the activity greatly impressed the students and their teachers, who, for the first time, discovered these methodologies and carried out linguistics activities. These activities allow students to better understand mathematical topics and to be more aware of the importance of using correct language to understand and be understood correctly.

We also looked for correlations among items and calculated the Pearson Correlation Coefficient among all the possible pairs of items. We did not detect any strong correlation among items, and the correlation coefficient was not higher than 0.75. However, we found some interesting weak correlations among the following pairs of items.

Couple #1 is “The proposed activities stimulated my interest in mathematics” and “The exercises helped me to better understand the topics covered” [$r(33)=0.65$, $p<0.001$]. This correlation shows that interest and understanding grow together; engaging and motivating students can lead to a better understanding of the topics.

Couple #2 is "The exercises helped me to better understand the topics covered." and "The exercises made me appreciate the topics covered more." [$r(33)=0.65$, $p<0.001$]. This couple shows that understanding and appreciation are related. However, we can think of this relation mainly in one direction: the more students understand, the more they appreciate what they studied. This is of great importance for teachers who design activities for their students, as there is a need to focus on the understandability of the activity. Moreover, the immediate and interactive feedback that helped students while performing the activity played an important role in this correlation.

Two couples, couple #3, "It is helpful to view the correct answer immediately after answering a question." and "When I gave a wrong answer, I would try the exercise again." [r(33)=0.63, p<0.001] and couple #4 "When I gave a wrong answer, I would try the exercise again." and "Having the assessment immediately after solving an exercise helped me understand how to answer the question." [r(33)=0.63, p<0.001], concern the formative aspect of the activity: on one side if the student have the chance to retry, they are willing to do it, and this is how a learning environment should be, a place where students can try without the fear of mistakes, on the other side, when having multiple attempts, students understand better how to deal with the activity, viewing the whole process and then going back from the beginning to start a good performance.

As a last analysis, to search for useful insights on gender studies, we also investigated the presence of differences in the responses of females and males to the items of the final questionnaire, together with the statistical significance of those differences through a two-tailed t-test. The results are presented in Table 2.

Table 2. Mean values of females' and males' responses with a t-test significance

Item	**Mean (females)**	**Mean (males)**	**Difference (females-males)**	**p-value**
I paid attention during the activities	3.65	3.33	0.32	0.113
The proposed activities were interesting	3.35	3.07	0.28	0.147
The proposed activities were understandable	3.45	3.40	0.05	0.815
The proposed activities stimulated my interest in mathematics	2.80	2.33	0.47	0.026
The proposed activities were useful for gaining awareness of the language	3.35	2.87	0.48	0.039
I enjoyed the group activities	3.50	3.13	0.37	0.206
It was easy to use AntConc.	2.80	3.47	-0.67	0.022
It was easy to understand how to use AntConc to answer questions.	2.95	3.40	-0.45	0.111
It was interesting to learn how to use AntConc.	3.45	3.33	0.12	0.658
Learning to use text analysis software is useful for the future.	3.05	3.27	-0.22	0.491
The exercises helped me to better understand the topics covered.	3.30	2.73	0.57	0.047
The exercises made me appreciate the topics covered more.	2.85	2.33	0.52	0.157
The exercises helped me to realize my preparation.	3.20	2.80	0.40	0.111
It is helpful to view the correct answer immediately after answering a question.	3.70	3.53	0.17	0.403
When I gave a wrong answer, I would try the exercise again.	3.55	3.20	0.35	0.159
Having the assessment immediately after solving an exercise helped me understand how to answer the question.	3.60	3.40	0.20	0.398
The exercises helped me to be more aware of the language to use.	3.25	3.00	0.25	0.364

The exercises on language also made me reflect on meanings and contents.	3.00	2.47	0.53	0.050

In general, there are no peculiar differences among the two components of the sample. The highest difference (-0.67) with the highest significance (0.022) appears in the item “It was easy to use AntConc”. In this specific case, the males signalled a larger agreement on this item. Conversely, the females signalled greater agreement on "The proposed activities stimulated my interest in mathematics” with difference 0.47 and significance 0.026. The other two elements with a p-value lower than 0.05 are represented by "The proposed activities were useful for gaining awareness of the language” (0.039) and “The exercises helped me to better understand the topics covered.” (0.047).

5. Conclusions

In this paper, we present a didactic experiment characterised by the DDL and automatic formative assessment methodologies to study the language of mathematics. The results of the final questionnaire show that the students appreciated the mathematical activities, which were very different from the traditional ones. The use of technology had a very positive and motivating impact on the students.

The students highly appreciated the proposed methods to answer the first research question (RQ1). The DDL approach had a very positive and motivating effect on the students. According to the students, learning how to do linguistic analysis and useful for the future was interesting. They had no difficulty using AntConc and understood the linguistic investigations needed for corpus-based activities with automatic formative assessment. The students also appreciated the automatic formative assessment methodology. They appreciated the possibility of getting immediate feedback on the correctness of each answer, correcting themselves in case of a wrong answer by having multiple attempts and knowing the correct answer at the end of each activity. All these aspects are closely related.

About the second research question (RQ2), the activities helped students become more aware of the language used. Not all students agree that the proposed activities can increase their interest in mathematics, but they agree that it is useful to do this kind of activity to reflect on the language used. The results show a correlation between students' understanding of mathematical topics and their appreciation of the topics themselves. These couple of items show that understanding and appreciation are related, and this relationship is very important when planning educational activities. The correlation between understanding the proposed activities and understanding the mathematical topics was another interesting result in this respect.

There are no particular differences between males and females (RQ3). The most significant differences concerned the use of AntConc and the impact of the proposed activities on the students' understanding of mathematics and its language. According to the results, males found it easier to learn to use AntConc. The activities further stimulated females' interest in mathematics and their awareness of mathematical language. Furthermore, the activities helped females more to increase their awareness of the proposed topics.

Students acquire language and digital skills using the tools and techniques of corpus linguistics for pedagogical purposes. The use of the digital learning environment and an automatic assessment system enables the generation and collection of data on student learning: about learning processes and not just about results. These data are useful for driving and adjusting the learning path, making choices and decisions, and supporting learning in several ways. It would be important to train teachers on using these methodologies and technologies in the future. In particular, for Mathematics teachers, it could be very interesting and formative to learn the analysis of specialized texts through the consultation of a corpus and the design of DDL activities with formative assessment to make students reflect on the specialized language of Mathematics. It is important to promote reflection among all teachers of all levels on the importance of acquiring specialized languages and how linguistic difficulties can influence disciplinary difficulties.

This research activity, characterised by DDL methods integrated with automatic formative assessment systems, could also be carried out on a larger scale, involving many teachers and students. The main challenge for mathematics teachers might be to reflect on the linguistic aspects. For this reason, training teachers in the proposed methodologies and to help them acquire linguistic skills would be essential. A second important aspect is to provide teachers with ready-made materials (corpus and corpus-based teaching activities with automatic formative evaluation). Indeed, one of the biggest obstacles could be the corpus creation and the consequent selection of representative texts. If teachers wanted to cover a topic other than the concept of function, they would have to create a new corpus and new activities based on this corpus. The situation is similar if one wants to apply this approach to other subjects, such as chemistry or biology, which have specialised language. For the future development of this research activity, the training of mathematics teachers is fundamental, possibly integrated with the training of language teachers and creating databases of ready-made activities.

Using technologies and a student-centred DDL approach allows students to be involved in linguistic tasks, especially in a STEM subject to which they are not accustomed. The exploration of corpora can effectively support reflection on the specialized languages of Mathematics. Many technologies used daily by students support linguistic investigations, but it is important to question the types of data on which they are carried out. Students use the most recent LLM and Google daily, and these tools can be used for linguistic investigations. However, there is no control over the data on which searches are carried out, the results of the searches, and the type of language they use.

Given the considerable diffusion of AI tools, training teachers and students on their strengths and weaknesses and how they influence teaching and learning is important. Students can use these tools critically and profitably even for less standard tasks such as linguistic research in Mathematics if appropriately trained. LLM, such as ChatGPT, could not be used for the same type of tasks since designing tasks with automatic formative assessment to guide students in linguistic investigations requires knowing the correct answer. ChatGPT answers may differ from student to student and may change over time. However, it would be interesting to use ChatGPT, through appropriately designed activities, to make students reflect on the language of Mathematics as a starting point for linguistic and disciplinary reflections. It is important to note that

Generative AI has strengths in advancing our understanding of language in se that corpora have struggled to address. By combining corpus and Generative AI approaches, language learners can better understand how language works in different contexts than is currently possible with either approach alone. DDL researchers are well placed to take advantage of this renewed mainstream interest in language data, as we understand both the power of such data for language teaching and the conditions necessary for meaningful learning to take place with such data.

AI tools are changing mathematics education just as they will change the way of researching mathematics education. It is important to reflect on how these tools can influence students' language, both specialist and non-specialist, and their linguistic and not just disciplinary skills.

Acknowledgments. The research was carried out within Indam - Istituto Nazionale di Alta Matematica “Francesco Severi”. The authors thank all the students who contributed to the research by responding to the questionnaire, and to all the teachers who showed passion for their work and expressed interest in improving their teaching.

References

1. Ferrari, P. L.: Costruzione di competenze linguistiche appropriate per la matematica a partire dalla media inferiore. L’insegnamento della matematica e delle scienze integrate, 26A, 4, 469-496 (2003).
2. Sbaragli, S., Demartini, S., Franchini, E.: Le difficoltà di comprensione e di gestione dei termini specialistici della geometria all’ingresso della scuola secondaria di primo grado. La matematica e la sua didattica, 29(1), 7-37 (2021).
3. OECD: How did countries perform in PISA? In PISA 2022 Results (Volume I): The State of Learning and Equity in Education, OECD Publishing, Paris, https://doi.org/10.1787/9149c2f5-en, (2023).
4. Johns, T.: Should you be persuaded: two examples of data-driven learnin. Classroom Concordancing. ELR Journal, 1-12 (1991).
5. Corino, E., Onesti, C.: Data-driven learning: a scaffolding methodology for CLIL and LSP teaching and learning. In Frontiers in education, 4, article 7 (2019).
6. Attardi, G.: Il Bello, il Brutto e il Cattivo dei LLM. Mondo Digitale, 1-16 (2023).
7. Zhai, X.: ChatGPT user experience: Implications for education. SSRN 4312418, 1-10 (2022).
8. Abedi, M., Alshybani, I., Shahadat, M. R. B., Murillo, M.: Beyond Traditional Teaching: The Potential of Large Language Models and Chatbots in Graduate Engineering Education. Qeios (2023).
9. Frieder, S., Pinchetti, L., Griffiths, R. R., Salvatori, T., Lukasiewicz, T., Petersen, P. C., Berner, J.: Mathematical capabilities of chatgpt. arXiv:2301.13867 (2023).
10. Shakarian, P., Koyyalamudi, A., Ngu, N., Mareedu, L.: An Independent Evaluation of ChatGPT on Mathematical Word Problems (MWP). arXiv:2302.13814 (2023).
11. Corino, E.: Didattica delle lingue corpus-based. Educazione Linguistica - Language Education, 8, 231-248, (2014).
12. Corino, E., Fissore, C., Marchisio, M.: Data Driven Learning activities within a Digital Learning Environment to study the specialized language of Mathematics. In: O’Conner, L. IEEE 46th Annual Computers, Software, and Applications Conference (COMPSAC), 167-176 (2022).
13. Barana A., Marchisio, M.: Analyzing Interactions in Automatic Formative Assessment Activities for Mathematics in Digital Learning Environments. In Csapó, B., Uhomoibhi, J.

13th International Conference on Computer Supported Education (CSEDU 2021), 1, 497-504 (2021).
14. Barana, A., Marchisio, M., Sacchet, M.: Interactive feedback for learning Mathematics in a digital learning environment. In Education Sciences, vol. 11(6), 279 (2021).
15. Barana, A., Marchisio, M., Sacchet, M.: Effectiveness of automatic formative assessment for learning mathematics in higher education. In Domenech J., Merello, P., De la Poza, E. 7th International Conference on Higher Education Advances (HEAd'21), 1-8 (2021).
16. Barana, A., Fissore, C., Marchisio, M.: Automatic Formative Assessment Strategies for the Adaptive Teaching of Mathematics. In Communications in Computer and Information Science, 1473 CCIS, 341–365 (2021).
17. Marello, C., Marchisio, M., Pulvirenti, M., Fissore, C.: Automatic assessment to enhance online dictionaries consultation skills. In 16th International Conference on Cognition and Exploratory Learning in Digital Age, CELDA 2019, 331–338, (2019).
18. Corino, E., Fissore, C., Marchisio, M.: Adaptive Exercises and Formative Assessment for English Remedial Action. Orchestration of Learning Environments in the Digital World, 3-19 (2022).
19. Sbaragli, S., Demartini, S.: Lingua e strutture dei testi scolastici di matematica. Italmatica. Dedalo Editor (2021).
20. Ferrari, P. L.: Educazione matematica, lingua, linguaggi. Costruire, condividere e comunicare matematica in classe (2021).
21. OpenAI, ChatGPT, https://chat.openai.com, last accessed 2023/12/10.
22. Fissore, C.: Data driven learning activities in English on the specialized language of Mathematics in secondary school. In: Chova L. G., López Martínez A., Lees J. 14th International Conference on Education and new Learning Technologies (EDULEARN), 9326-9333 (2022).
23. Barbera, E. F.: Linguistica dei corpora e linguistica dei corpora italiana. Un'introduzione, pp. 1-113 (2013).
24. Cobb, T., Boulton, A.: Classroom applications of corpus analysis. Cambridge Handbook of English Corpus Linguistics, 478–497 (2015).
25. Boulton A.: Testing the limits of data-driven learning: Language proficiency and training, ReCALL, 21(1), 37-51 (2009).
26. Boulton, A.: What data for data-driven learning?. EuroCALL Review, 20(1), 23-27 (2012).
27. Kohnke, L., Moorhouse, B. L., Zou, D.: ChatGPT for Language Teaching and Learning, RELC Journal, 54(2), 537-550 (2023).
28. Baskara, R.: Exploring the implications of ChatGPT for language learning in higher education. Indonesian Journal of English Language Teaching and Applied Linguistics, 7(2), 343-358 (2023).
29. Crosthwaite, P., Baisa, V.: Generative AI and the end of corpus-assisted data-driven learning? Not so fast!. Applied Corpus Linguistics, 3(3), 100066 (2023).
30. Marchisio M., Barana A., Floris F., Pulvirenti M., Sacchet M., Rabellino S., Marello C.: Adapting STEM automated assessment system to enhance language skills. Proceedings of the eLearning and Software for Education Conference (eLSE 2019), pp. 403 – 410 (2019).