

Mapping Elements of the Field and Language in Texts of Practical Class Reports in Organic Chemistry Produced by Students

Mapeando Elementos do Campo e da Linguagem em Textos de Relatórios de Aulas Práticas em Química Orgânica Produzidos por Estudantes
Mapeo de Elementos del Campo y del Lenguaje en Textos de Informes de Clases Prácticas de Química Orgánica Elaborados por Estudiantes

Edinilza Maria Anastácio Feitosa,^{id} e Eduardo Fleury Mortimer^{id}

Abstract

In the field of chemistry, writing must move between macroscopic, submicroscopic and representational levels, using specific elements of language and the context in which language is produced. Analysis models based on Systemic Functional Linguistics, one that analyzes the construction of the field, and another that analyzes the construction of entities within discursive semantics, were used to analyze the writing of undergraduate chemistry students. Our objective was to evaluate whether these models could be applied to chemistry texts produced by students, identifying the linguistic choices used in the textual production, through the mapping of specific entities of the field of chemistry and to make adaptations to the evaluated models so that they are suitable for the context of Chemistry. The corpus of analysis consisted of 31 texts divided into three groups of experiments in organic chemistry. For each group, elements related to the field (taxonomic relations, types of activities and properties) and elements of language (entity, dimensionality and live grammatical metaphors) were mapped. The analysis of the texts showed that students have difficulty describing the field of chemistry in the submicroscopic world and this impacts the construction of taxonomies and the type of activities described. The mapping of entities indicated that the diversity of entities depends directly on the type of experiment and that students do not construct living grammatical metaphors. Therefore, the analysis models adopted in this article can be used to analyze texts produced by chemistry students. However, the texts analyzed here refer to a specific area of chemistry, organic chemistry. One possible unfolding of this research would be to extend this study to other areas of chemistry.

Keywords: language, chemistry, entity, systemic functional linguistics, higher education

Resumo

No campo da química, a escrita precisa transitar entre os níveis macroscópico, submicroscópico e o representacional, utilizando-se de elementos específicos da linguagem e do contexto no qual a linguagem é produzida. Modelos de análise baseados na Linguística Sistemico Funcional, um que analisa a construção do campo, e outro que analisa a construção de entidades dentro da semântica discursiva foram utilizados para analisar a escrita de estudantes de graduação em química. Nosso objetivo foi avaliar se estes modelos podiam ser aplicados a textos de química produzidos pelos estudantes, identificando escolhas linguísticas utilizadas na produção textual, mapeando entidades específicas do campo da química e realizar adaptações nos modelos avaliados de forma que se adequem ao contexto da Química. O corpus de análise foi constituído de 31 textos divididos em três grupos de experimentos em química orgânica. Para cada grupo, foram mapeados elementos relacionados ao campo (relações taxonômicas,

tipos de atividades e propiedades) e elementos da linguagem (entidade, dimensionalidade e metáforas gramaticais vivas). A análise dos textos mostrou que os estudantes têm dificuldades em descrever o campo da química no mundo submicroscópico e isso impacta na construção de taxonomias e nos tipos de atividades descritas. O mapeamento de entidades indicou que a diversidade das entidades depende diretamente do tipo de experimento e ainda que os estudantes não constroem metáforas gramaticais vivas. Portanto, os modelos de análise adotados neste artigo podem ser utilizados para analisar textos produzidos pelos estudantes de química. No entanto, os textos aqui analisados se referem a uma área específica da química, a química orgânica. Um desdobramento desta pesquisa seria estender este estudo à outras áreas da química.

Palavras-chave: linguagem, química, entidade, linguística sistêmico funcional, ensino superior

Resumen

En el campo de la química, la escritura debe moverse entre niveles macroscópicos, submicroscópicos y representacionales, utilizando elementos específicos del lenguaje y del contexto en el que se produce el lenguaje. Para analizar la escritura de estudiantes de química de pregrado, se utilizaron modelos de análisis basados en la Lingüística Funcional Sistémica, uno que analiza la construcción del campo, y otro que analiza la construcción de entidades dentro de la semántica discursiva. Nuestro objetivo fue evaluar si estos modelos podrían aplicarse a textos de química producidos por estudiantes, identificando las elecciones lingüísticas utilizadas en la producción textual, a través del mapeo de entidades específicas del campo de la química y realizar adaptaciones a los modelos evaluados para que sean adecuados al contexto de la Química. El corpus de análisis consistió en 31 textos divididos en tres grupos de experimentos en química orgánica. Para cada grupo, se mapearon elementos relacionados con el campo (relaciones taxonómicas, tipos de actividades y propiedades) y elementos del lenguaje (entidad, dimensionalidad y metáforas gramaticales vivas). El análisis de los textos mostró que los estudiantes tienen dificultades para describir el campo de la química en el mundo submicroscópico y esto impacta en la construcción de taxonomías y el tipo de actividades descritas. El mapeo de entidades indicó que la diversidad de entidades depende directamente del tipo de experimento y que los estudiantes no construyen metáforas gramaticales vivas. Por lo tanto, los modelos de análisis adoptados en este artículo pueden usarse para analizar textos producidos por estudiantes de química. Sin embargo, los textos analizados aquí se refieren a un área específica de la química, la química orgánica. Un posible desarrollo de esta investigación sería extender este estudio a otras áreas de la química.

Palabras clave: lenguaje, química, entidad, lingüística sistémica funcional, educación superior

Introduction

In addition to learning the phenomena, concepts, theories, and representations of a discipline of the natural sciences, students must also know how to speak, argue, read, and write about the content (Carvalho, 2018). In this study, we focus on the dimension of writing as a way to describe and explain phenomena, concepts, theories, and representations of the discipline of Chemistry, particularly organic chemistry. However, writing requires the development of a specific linguistic competence of the discipline, and in the case of chemistry, in order to participate in the scientific discourse of this discipline, students must also be able to move across three levels of knowledge: the macro and tangible level of what can be seen, touched, smelled; the submicroscopic level of atoms, molecules, ions, and structures; and the representational level, which includes symbols, formulas, and equations (Johnstone, 2000). In addition, students need to master a specific vocabulary of the discipline, whose terms may have different meanings, depending on the context in which they are used.

The exercise of writing occurs at various moments throughout school life. In chemistry courses, for example, it is common, after a practical activity, for students to be asked to produce a written text in the form of an activity report. It is in the “results and discussion” section of this report that the teacher expects the student to describe the results obtained in the activity and, in addition, to give explanations for the phenomena or transformations involved in the activity experienced. Another expectation of the teacher is that students use, in this writing, a language very close to scientific language, which has specific characteristics in the way meanings are organized and written (Halliday & Martin, 1993).

There is a concern regarding students' writing in the reports of practical laboratory classes (Bortnik et al., 2017; Clary-Lemon et al., 2019; Crawford & Klooper, 2019; Hoffa & Freeman, 2008; Paton-Walsh, 2015; Sampson & Walker, 2012). In all these studies, the authors observe that students have difficulties in producing written texts about laboratory experiments, especially regarding the explanations and discussion of data obtained in the experiment. Writing requires students to engage and reflect on what was done in the laboratory (Crawford & Klooper, 2019), and works as a form of learning (Clary-Lemon et al., 2019). To improve writing, the studies describe several strategies, such as writing built gradually, going through the writing process (learning to write) and then producing a written product (writing to learn) (Clary-Lemon et al., 2019); student engagement in peer review (Clary-Lemon et al., 2019; Sampson & Walker, 2012); participation in virtual lab classes prior to face-to-face lab class (Bortnik et al., 2017); or the use of a report template to guide students in improving their writing skills (Paton-Walsh, 2015). However, none of these studies considered the linguistic elements specific to the field of chemistry in the students' writing.

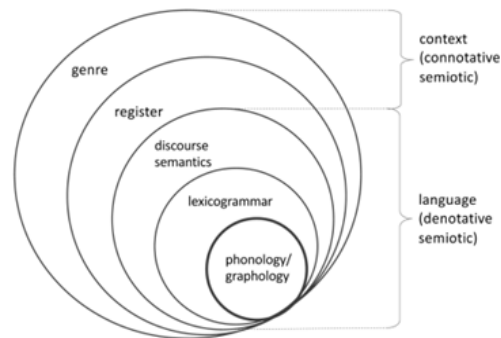
Unlike these studies, we are concerned with analyzing how students construct their writing of a practical class report from a linguistic point of view, or, more precisely, which linguistic choices are made by these students to describe and explain a chemical phenomenon. In the Results and Discussion section, we analyze texts from organic chemistry reports involving the study of chemical reactions (esterification), the study of a technique (extraction), and the study of a chemical property of organic compounds, solubility. For this, we adopt the framework of Systemic Functional Linguistics (SFL).

In Systemic Functional Linguistics, language is a stratified system dependent on context (Figure 1). Halliday and Hasan (1985) define context in their book **Language, Context, and Text: Aspects of Language in a Social-Semiotic Perspective**. They developed their ideas from anthropologist Bronislaw Malinowski and his theory of the context of situation and the context of culture. Until 1923, when Malinowski (1923) published his article, the word 'context' in English meant **co-text**, that is, the words and sentences before and after the specific sentence being analyzed. Malinowski introduced the term **context of situation**, referring to the environment of the text, and also introduced the term **context of culture**, referring to the total cultural context involved in interpreting the meaning of a text. While the context of situation requires attention to the immediate images and sounds surrounding the interactions in the production of, for example, an oral text, the context of culture requires that the entire cultural history behind the participants of an interaction be taken into account.

Halliday (1993) establishes the relationship between social context and language through a model in which he uses concentric circles, as shown in Figure 1. The model is used to establish a semiotic system (language) as the realization of another, more abstract semiotic system (context). Halliday (1978) proposes that what we should know about the context in order to predict the linguistic features that are likely to be associated with it are what is actually happening (the field), who is participating (the relations), and what part of the language is being used (the mode). These three variables are called the **Register** of the context of situation. "Register is, therefore, the configuration of meanings that happen because of the situation" (Fuzer & Cabral, 2014, p. 25) and are related to the metafunctions of language: ideational, interpersonal, and textual. These, in turn, construct our experience of the external and internal world (ideational meanings); the social relationships we maintain with the people around us (interpersonal meanings); and how ideational and interpersonal metafunctions are composed in a text (textual meanings) (Figure 2).

Figure 1

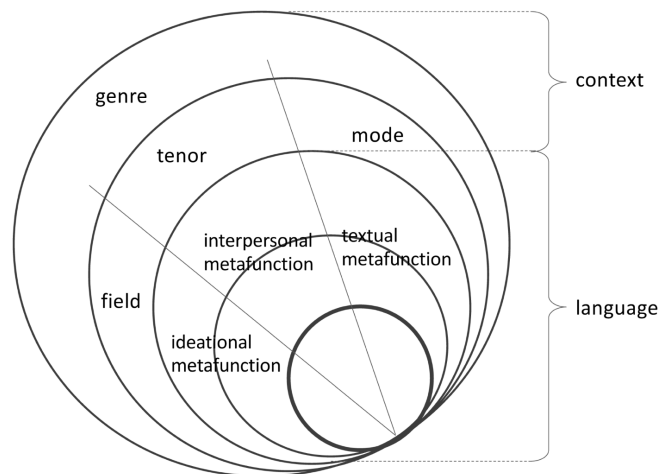
Stratified system according to SFL



Source: Hao (2020, p. 9).

Figure 2

Metafunctional relations between language and context



Source: Hao (2020, p. 11)

The understanding of language and context stratified from the perspective of SFL rearticulated how meanings are organized both in the stratum of the field, which is part of the context, and in the stratum of discourse semantics, which is part of language itself (Dreyfus & Hao, 2020). Based on this understanding, some models have been developed to analyze science texts, for example, Doran and Martin's (2021) field model that analyzes science textbook texts in relation to the context field variable, and Hao's (2020) textual choices model, which analyzes the construction of meanings within the stratum of discourse semantics that carries out the field, and was prepared from texts produced by biology students.

Our objective here is to evaluate whether these models can also be applied to chemistry texts produced by students in the Results and Discussion session of experimental activity reports, identifying the linguistic choices used in the construction of these texts, through the mapping of specific entities used in the field of chemistry. Understanding how students describe, explain, and argue about experienced chemical phenomena helps us to understand the language they use to move between the macroscopic and submicroscopic worlds, and to identify learning gaps, which in the future may contribute to the proposal of appropriate teaching strategies to overcome these failures. In addition, adapting models of language assessment, which are based on Systemic Functional Linguistics and, at the same time, are suitable for the field of chemistry, constitutes an important contribution to the field of science education.

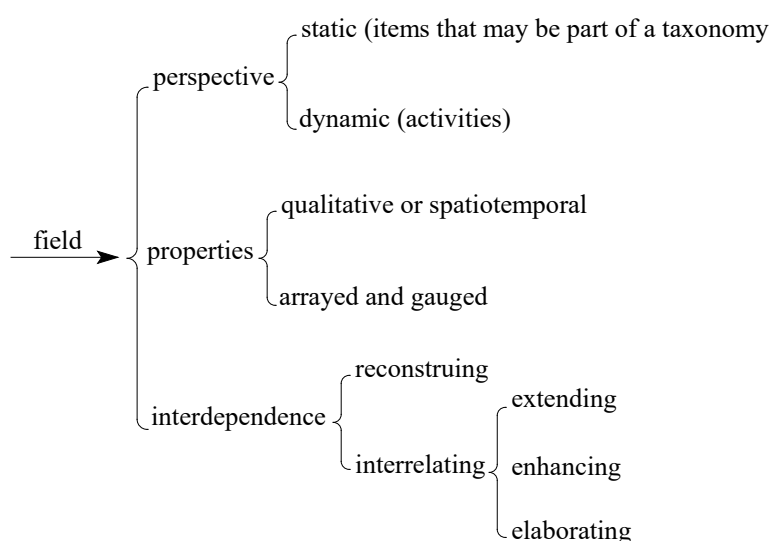
Theoretical Frameworks Supporting the Research

Doran and Martin's Field Model

Doran and Martin's (2021) field model, which applies to context, consists of three systems (Figure 3). In the first system, field can be interpreted from two different perspectives: a static one, which sees knowledge as items; and a dynamic one, which sees knowledge as an activity. Taxonomy is modeled as an aspect of field organization, along with the sequences of activities in which taxonomies can be included (Hao, 2020). As an example, knowledge as items, from a static perspective, includes the various items that populate an organic chemistry laboratory: salicylic acid, test tube, base, etc. Knowledge as an activity includes, for example, reactions of addition, esterification, electron migration, etc.

Figure 3

Doran and Martin's Field Model (2021)



Source: Prepared by the authors from Doran and Martin (2021).

Taxonomies involving items are of two types: classification taxonomy, in which types and subtypes of items are identified, for example, the item **Base** can be classified as a strong base or a weak base; and compositional taxonomy, with multiple levels of part-whole relationship between items, for example, when naming parts of a molecule. For instance, the active site and the output group are parts of a whole, the substrate. In the static perspective of the field, taxonomies can be indefinitely expanded or deepened, and the expansion of taxonomies is one of the main characteristics of scientific knowledge.

In the dynamic perspective, phenomena are constructed in terms of activities. The activity involves some kind of change that is oriented towards some global purpose. Activities can be classified as unmomented and momented. Unmomented activities are those that are not part of a sequence of activities; they are single activities, built through a figure (Hao, 2020), for example, sulfuric acid acts as a catalyst. According to Doran and Martin (2021), activities can also encapsulate other smaller activities; in an example given by the authors, “the unmomented activity inflammation is momented through various subheadings into four stages: increased diameter and permeability of blood vessels; phagocyte migration; phagocytosis; and tissue repair.” (Doran & Martin, 2021, p. 110). Momented activities, on the other hand, constitute a sequence of activities and may be of implication, when one activity implies another activity. These are more common in scientific explanations, in which events tend to be described in terms of causal or conditional relations, for example, the heating of the reaction mixture (activity 1) leads to the breaking of the bonds in the reactants (activity 2). Momented activities may also be of expectancy, which are activities that do not have a cause-and-effect relationship between them and are more common in the description of experimental procedures.

Associated with this system, from the perspective of field, is the system of properties. An item or activity may or may not have property, which may be qualitative or spatio-temporal. Qualitative properties are those that attribute some type of qualitative attribute or identifier to an item or activity, for example, metals that have a silver color. A spatio-temporal property positions the item or activity in a specific time and space; for example, in one of the reports analyzed, the student uses the term **adjacent** to indicate the position of an atom: “In the third step, we have the migration of hydrogen from the oxygen that has bound to the substrate to an adjacent oxygen” (R3). These qualitative and spatio-temporal properties can also be arrayed and gauged; for example, sodium chloride is a white, hygroscopic solid with a density of 2.16 g/mL. In the field of chemistry, numerous properties are directly related to a quantity and are called quantitative properties, for example, the boiling temperature of water at sea level, which is 100°C. We understand that these properties are included in the model of Doran and Martin (2021) as properties that can be arrayed, for example, hydrochloric acid is a stronger acid than acetic acid, and gauged, for example, the boiling temperature of ethoxyethane (34.6°C) is lower than that of ethanol (78.4°C).

Items, activities, and properties can participate in a third system, that of interdependency, which can be construed in two ways: reconstruing and interrelating. In reconstruing, one resource may be reconstrued as another. An activity, for example, to react, can be reconstrued as the item, reaction (itemized activity). Thus, the reaction activity, as an item, can enter a taxonomy and be classified as a saponification reaction, a hydrolysis reaction, and an addition reaction. Interrelating refers to the way in which the different elements of the field are associated with one another. The elements can interrelate, positioning themselves as relatively independent of one another, as in the case of the different types of interactions between molecules. On the other hand, some elements are dependent on others. The physical property of boiling temperature, for example, is relatively dependent on the type of intermolecular interaction that the molecule presents. A summary of the three systems of field analysis, together with examples from the reports, is presented in Figure 4.

The interrelating among the field elements may occur by extending, in which the meaning of an isolated element can be complemented by another. This interrelation of meanings by extending may or may not lead to an enhancing, which in turn may undergo a process of elaborating or be rewritten in another way, giving rise to another field element. In the example cited by Doran and Martin (2021), the itemized property, Earth's tilt, has its meaning complemented by extending, by the elements hemispheres (item), Earth's orbit (activity), and reception of sunlight by the Earth (activity). The combination of these elements leads to a variation in solar energy (itemized property) around the Earth (enhancing). In turn, the variation in solar energy can be rewritten as the seasons of the year (elaborating), another itemized property, which, as an item, can be taxonomized by classification into summer, winter, autumn, and spring. We did not find many examples of interrelating for the content of the reports analyzed.

Figure 4*Doran and Martin's model for the field of chemistry*

System			Example
Perspective	Static (items)	Classification	Weak base Strong base
		Part/whole composition	Output group and active site are parts of a molecule.
	Dynamics (activities)	Momented implication	Heating the reaction mixture leads to the breaking of the bonds in the reactants.
		Momented expectancy	Add water to the reaction mixture. Filter.
		Unmomented	Sulfuric acid acts as a catalyst. Solubilization
	Property	Qualitative	
Spatio-temporal		In the third step, we have the migration of hydrogen from the oxygen that has bound to the substrate to an adjacent oxygen.	
Arrayed and gauged		Sodium chloride is a white, hygroscopic solid with a density of 2.16 g/mL. The boiling temperature of ethoxyethane (34.6°C) is lower than that of ethanol (78.4°C).	
Interdependency	Reconstruing	Activity as an item	Saponification reaction Hydrolysis reaction
	Interrelating	They do not relate	Hydrogen bonding Ion-ion interaction
		Relating	The physical property of boiling temperature depends on the type of intermolecular interaction that the molecule presents.

Hao's Discourse Semantic Model of Entities and Dimensions

Items and itemized activities, from the perspective of field, are carried out at the stratum level of discourse semantics as entities. Hao (2018) has constituted several types of entities that depend on the field. In Hao's (2020) studies, six types of entities were listed: source entity, thing entity, activity entity, semiotic entity, place entity, and time entity. In this work, we limit the study of entities to two types: the thing entity and the activity entity. The thing entity is associated with Doran and Martin's (2021) static field view, while the activity entity is related to activities in the dynamic perspective of field. Doran and Martin (2021) suggest that activity entities allow field activity to be reconstrued from a static perspective, becoming discriminated or itemized activities which, as items, can enter classification and composition taxonomies.

According to Hao's (2020) categorization, the thing entity can be classified into two distinct groups: instrumental thing entity and observational thing entity. Those that are typically related to the instrumental part of an experiment are the instrumental thing entities and include the names of substances and equipment. Observational thing entities involve items that need to be defined linguistically. In the field of biology, these items include what can be seen with the naked eye, such as an insect, what can be seen through a microscope, for example, a cell, and what is only inferred, such as an enzymatic site. In the field of chemistry, the thing entities of the observational category can be seen from two perspectives: those that can be observed, for example, a portion of liquid linguistically defined in the macroscopic world as sulfuric acid, and the sulfuric acid molecule that is located in the submicroscopic world and, therefore, is inferred.

Activity entities are related to processes, that is, activities that are carried out both in the macroscopic world and in the submicroscopic world. Activity entities can be of two types: enacted activity entity and observational activity entity. The enacted activity entity is related to **doing science**, for example, an **experiment**, study, or how to do science, that is, the **method**. An enacted activity entity is an activity performed by an **agent** that is a people entity. Observational activity entities are activities that, in the field of chemistry, usually occur at the submicroscopic level and represent usually nominalized processes, for example, a reaction. The agent that performs the observed activity entity is a thing entity.

Thing entities and activity entities can have their meanings enriched by what Hao (2020) refers to as dimensions of an entity. The system of dimensionality proposed by the author comprises four types of dimension: categorised, structured, measure, and perceived. The categorised dimension names the relationship between entities through classification taxonomy, while the structured dimension relates entities through part-whole composition taxonomy. These dimensions, categorised and structured, provide important resources for explicitly naming taxonomic relationships among field items, including activities reconstrued as items, in semantic discourse.

The measured and perceived dimensions name the properties based on which items are distinguished from one another. The measured dimension enriches an entity by showing how it can be quantified. This measurement can be specific, made numerically, for example, 0.5 mL of sulfuric acid, or generic through terms that indicate a non-specific amount, such as, for example, **various** or **some**. The measured dimension can be associated with a quantitative property, since this property can be compared, arrayed, and gauged, for example, “caffeine is 9 times more soluble in dichloromethane than in water” (S8). The perceived dimension provides a resource to name qualitative properties that can be perceived and observed through the human senses, such as sight, smell, hearing, and touch, and do not need to be quantified, for example, “the formation of a white precipitate was observed” (R1). Thus, from the perspective of the field, **color** is a property, whereas from the perspective of discourse semantics, **color** is a dimension that enhances the meaning of the entity precipitate.

The dimensions **amplify** the entities (thing and activity) and enter the discourse because they are dependent on an entity, and this entity-dimension relationship, as a whole, interacts with the other interpersonal meanings in the text and is thus treated as a semantic unit of discourse (Hao, 2020).

Adaptations to Hao’s Model from the Perspective of Chemistry

In the stratum of discourse semantics, Hao (2020) proposes the construction of meanings through entities and dimensions. Initially, we divided the entities into two groups: The thing entities (instrumental and observational) that perform items; and the activity entities (reported and observed) that perform activities. In our analysis, in the field of chemistry, we realized that this classification needed to be reformulated. In chemistry, we work with the concepts of the macroscopic world related to the submicroscopic world. In the submicroscopic world, many items, activities, and properties are inferred. Another aspect considered is that some occurrences can be classified as an observed thing entity, such as precipitate, which is the product of a chemical reaction and consists of a chemical substance. Thus, we classify the thing entities as instrumental, of occurrence, and inferred (Figure 5).

Figure 5

Reclassification of thing entities

Thing entity	Definition	Example
Instrumental	Ostensibly defined in the macroscopic world	Test tube, analytical scale
Instrumental	Linguistically defined in the Macroscopic world	Sulfuric acid, sodium hydroxide
Occurrence	Linguistically defined in the Macroscopic world	Precipitate
Inferable	Linguistically defined in the submicroscopic world	Atom, ion

When we move on to the second group proposed by Hao (2020), of activity entities, it is also subdivided into enacted activity entities and observed activity entities. For us, this classification likewise does not contemplate the specificity of the field of chemistry. Some activities, such as filtration and heating, are very common in experimental procedures and were allocated to the category of reported activity entity. Activities that occur in the submicroscopic world were classified as inferred activity entities (Figure 6).

Figure 6

Reclassification of activity entities

Activity entity	Definition	Example
Reported	Activity entity, whose agent performing the activity is a person entity	Method, experimental procedure, filtration
Inferred	Linguistically defined, having a thing entity as the agent performing the activity	Acetylation reaction, hydrolysis, solubility

The Grammatical Metaphor

An important element of language that we also considered mapping was the construction of grammatical metaphors by the students. Scientific language profoundly alters the way of talking about the world, while operating through a new grammar. One of the main characteristics of this grammar is the use of grammatical metaphor (Halliday & Martin, 1993). Unlike the ordinary metaphor, in which a name is used metaphorically, without any change in the grammatical structure, such as in the use of the word **deluge** in the sentence **He presented a deluge of ideas**, the grammatical metaphor involves the change of the grammatical function. Nouns or nominal groups, which normally designate things, animals, and people, come to designate processes, which is usually done through verbs.

Grammatical metaphor is a linguistic resource that allows semantic meanings (of discourse) to be performed through multiple grammatical choices. Halliday (1998) recognized two functions of grammatical metaphor: the instantial (known as **live** metaphor), which contributes to the organization of the text, and the systemic (known as **dead** metaphor), which contributes to the construction of technicism.

The discursive flow composed through the grammatical metaphor is textually oriented, packaging a clause in the text into a piece of information that will be given in the next clause involving a process of **nominalization**, which transforms, for example, a verb into a noun. In an example given by Halliday (1993, p. 81), “The atomic nucleus absorbs and emits energy only in quanta, or discrete units. Each absorption marks its transition to a state of higher energy, and each emission marks its transition to a state of lower energy.” The information that the atom emits and absorbs energy in the first clause

is packaged in two nominalized terms, **emission** and **absorption**, in the second clause. These nominal groups come to designate processes that are related by **relating verbs**. There is, therefore, a change in the grammatical function of the nominal groups that come to designate processes and of the verbal groups that come to designate relations between the two nominal groups. Since the clauses can be rewritten congruently, for example, **by emitting energy, the atom transit to a state of lower energy**, this grammatical metaphor can be grammatically unpacked, and so it is instantiated and, therefore, a live grammatical metaphor.

In scientific texts, there are a large number of nominalizations that are not easily grammatically unpacked and are used as technical terms to develop scientific knowledge. Halliday (1998) suggested that they initially emerged as grammatical metaphors, but were conventionalized or **distilled** in their use over time and, therefore, ceased to be metaphorical, that is, they became dead, for example, **hydrolysis**. Dead metaphors, as technical terms, are usually introduced by a definition (Martin, 1993, p. 183), for example, “diffusion is the process whereby a substance in *high concentration* moves to a place of *low concentration*” (italic added by the author).

Methodological Procedures

In this study, we delimit the analysis of the Results and Discussion section of experimental activity reports produced by students in the discipline of organic chemistry. The choice of this particular part is because it is in this section that the student describes the results obtained in the experiments and explains the observed phenomena.

Experiments in organic chemistry are grouped into three types: study of techniques, such as extraction and recrystallization; study of chemical reactions, such as substitution reactions; and the study of properties, such as acidity and basicity of organic compounds. In this study, the *corpus* of analysis consists of texts from each of these types of experiments (Figure 7), comparing these texts produced by the students with synthetic texts elaborated and validated by organic chemistry teachers. Whenever we quote excerpts from the students' texts or summary reports/synthetic texts, we will use quotation marks, followed by the letter **R** or **S** and a number, from 1 to 31, in parentheses, indicating the number of the report. Reports from R1 to R7 refer to the Chemical Reaction Study experiment and S8 to the corresponding summary report; reports R9 to R17 refer to the Technique Study experiment and S18 to the corresponding summary report; and those from R19 to R30 refer to the Property Study experiment and S31 to the summary report on this theme. The *corpus* of analysis was built from reports on the discipline of organic chemistry prepared in 2019 and 2023, provided by students of the Degree in Chemistry through an informed consent form and by professors who teach the discipline in the same course. The research project that gave rise to this article was approved by the Research Ethics Committee of the State University where it was carried out.

Figure 7*Composition of the corpus of analysis*

Experiment Group	Type of experiment	Subject of the experimental activity	Identification of student reports	Identifying Summary Reports
Group 1	Chemical reaction study	Esterification	R1 to R7	S8
Group 2	Technique study	Caffeine extraction	R9 to R17	S18
Group 3	Property Study	Solubility	R19 to R30	S31
Total Reports			28	03

Applying the Field Model of Doran and Martin (2021)

In terms of field, the texts were analyzed using the system of perspective, the system of property, and the system of interdependency. Regarding the first system, of perspective, we sought to identify the potentialities of constructing taxonomies in the text, whether of classification or composition, and what type of activities were chosen by the students to describe and explain phenomena. We adopted Doran and Martin's classification (2021) for unmomented activities and for the momented activities of expectancy and implication (Figure 8).

Figure 8*Momented activities and unmomented*

Type of activity	Description
Unmomented activity	A single activity that is not part of a sequence of activities and can encapsulate other activities
Momented expectancy activity	A sequence of activities in which one activity does not imply or depend on another
Momented implication activity	A sequence of activities in which one activity implies or depends on another

We mapped the qualitative and spatio-temporal properties in the analyzed texts. All terms that conferred an attribute to an item or an activity were considered as property. When this term indicated some degree of intensity of the property analyzed, or when the same property was compared across different elements of the field, we considered this property as arrayed, and when a numerical value was associated with the property, we considered it as arrayed and gauged.

We also applied the model presented in Figure 3 to map the field elements that underwent the reconstrual process in the form of itemized activities and itemized properties. The recognition of this reconstrual was due to the potentiality of these elements (items, activities, and properties) to construct taxonomies. For example, when a student describes that in the experiment an addition reaction occurred, they put the term **reaction** within a taxonomic classification, even if it does not indicate other types of reaction. It was also analyzed how students built relationships among the field elements.

Mapping Entities Using the Model Adapted from Hao (2020)

Regarding the performance of items and activities within the discourse semantics, we sought to identify which linguistic choices are made by students in terms of entities, following an adaptation of the classification proposed by Hao (2020). The characteristics of each entity are described in Figure 9.

Figure 9

Characteristics of entities and occurrences

Entity	Type	Feature	Example
Thing	Instrumental	Ostensively defined	Name of glassware, equipment, and materials needed for the experiment, e.g., test tube, scale.
		Linguistically defined	Names, molecular formulas, and class of substances involved in the experiments, e.g., sulfuric acid, base, and Na_2SO_4 .
	Occurrence	Linguistically defined	Terms that refer to substances that are products of chemical phenomena, e.g., precipitate, product, crystal.
	Inferable	Linguistically defined	Names of chemical species, for example, names of molecules, ions, and atoms.
Activities	Reported	Ostensively defined	Any activity that has a person as its performing agent, for example, an identification test, experiment, method, or filtration.
	Inferable	Linguistically defined	Any activity whose performing agent is a thing entity, e.g., reaction, bonding, intermolecular interaction.

Some activities performed by a noun are the result of the process of nominalization of a verb. There is a wide number of terms that are the result of this process of nominalization, for example, fusion (from fuse), boiling (from boil), and reaction (from react). Nominalization can be a resource for decontextualizing the text in order to bring writing closer to scientific language. Depending on how the nominalization was constructed by the student, the nominal group can indicate a grammatical metaphor, in which one grammatical class or structure is replaced by another (Halliday & Martin, 1993). The construction of grammatical metaphors was also mapped, and for this, we analyzed the nominalized terms used by the students in search of live grammatical metaphors.

In the analysis of the texts, we initially considered the word count available in the text and used the word count tool provided by a text editor. When the text presents figures, we take into account the words contained in the figure since some students, instead of

writing a text, prefer to explain a certain phenomenon through diagrams or schemes. We also considered the words contained in tables or charts produced by the students with relevant information to explain the phenomenon. Symbolic representations, such as chemical equations, were not considered in the word count.

After reading each text, we identified the words or groups of words that fit the characteristics defined for the entities and also identified the nominalizations. In many cases, when these elements were represented by two or more words, for example, hydrogen bonding, we counted this group of words as a single term and made deductions from the total number of words in the text. Thus, in the sentence **The molecules of water and ethanol form hydrogen bonds with each other**, we only counted ten words instead of the twelve counted in the word count tool. We used the same criterion when a property was associated with an entity, including value or rank, for example, **polar molecule** or **very strong acid**.

In some cases, we classified the entity according to the final word of the nominal group. For example, in **acetylsalicylic acid molecule**, we consider the word molecule as an entity, and the rest of the term was considered as an identifying attribute of this molecule. In this example and in similar cases, all the words were counted as a single term.

The number of entities identified in each text was converted into percentage values to enable the comparison of students' linguistic choices regardless of the subject of the practical activity. From these values, a profile of the language used by the students to describe and explain an experienced phenomenon was established. The data were organized into tables, following the coding described in Figure 10.

Figure 10

Entity coding

Entity	Code
Instrumental thing entity ostensibly defined	ECIod
Instrumental thing entity linguistically defined	ECIld
Thing entity Occurrence	ECO
Thing entity Inferable	ECInf
Activity entity Reported	EAR
Activity entity Inferable	EAIInf

Results and Discussion

Field Relations

For a better understanding of the discussion, in this section, we use a coding system for properties and terms that refer to dimensions when we present examples from students' texts and need to highlight these elements. Properties are underlined, and terms referring to dimensions are italicized.

From a Static Perspective

When analyzing the reports from group 1, which involve a chemical reaction, it was possible to identify some taxonomies of part-whole composition when the student explained, at a submicroscopic level, the mechanism of the reaction. For example, in R1, when the student identifies that hydroxyl is part of the salicylic acid molecule; in R3 when it is described that the “migration of hydrogen makes this part of the intermediate a good output group”; and in R6, when the student classifies the salicylic acid molecule as “bifunctional because it has two functional groups: phenol and carboxylic acid”. In each of these examples, there is a part, for example, an output group, which is part of a whole, the intermediary. Therefore, it is a taxonomy by part-whole composition. It was not possible to identify classification taxonomy in the students’ texts for this group of reports; however, we identified, in the summary report (S8), that the ordering of a property linked to an item allows this type of taxonomy to be inferred. The nominal group, **very strong acid**, suggests the existence of another acid with lower strength.

In the reports from group 2, which deal with the extraction technique, the classification taxonomy was evidenced when the student describes that it is “necessary to add dichloromethane to separate the organic phase from the inorganic phase” (R11). In the other texts, including the synthetic one, the classification taxonomy is also only suggested, for example, “the organic phase containing caffeine was treated with sodium sulfate” (S18). No other examples of classification taxonomy were found in the texts of this group.

The reports from group 3, which deal with a property, provide several examples where items and activities constitute classification and composition taxonomies. For example, in R19, the student explains that, in a molecule, negative poles and positive poles can be formed, showing that there are two types of charges associated with the poles, “in addition to forming negative and positive poles in the hexane”. They place these poles in a classification taxonomy. When the student explains the solubility of a chemical substance, describing that the structure of this substance has a polar part and a nonpolar part, for example, “acetanilide is soluble in methanol. This happens because acetanilide has a nonpolar and a polar part in its structure” (R19). In this case, the student is constructing a part-whole composition taxonomy, because the polar part and the nonpolar part belong to the same molecule.

From a Dynamic Perspective

Considering that, in the reports from group 1, some students use the description of the reaction mechanism through a picture with chemical structures, the activities described through verbs, for example, **they attack**, or nominalized verbs, such as **the attack**, were classified as unmomented activities. First, because they are part of the title of a stage of the reaction, for example, “protonation” (R1). Second, because these activities, for instance, protonation, encapsulate other activities, “an oxygen of the carbonyl attacks a proton of sulfuric acid (catalyst of the reaction), so that the oxygen retains the pair of electrons of the pi bond and leaves the carbon positive (active site)” (R3).

Most of the activities in these texts were identified as momented activities. These activities were analyzed as a sequence of two or more activities. We found that, in most of these sequences, one activity implies another. Students use this type of momented implication activity to describe what has been done, or to describe the results, at the macroscopic level, for example, “there is a need to cool the reaction mixture to crystallize the product formed” (R3), as well as to explain the mechanism of reaction at the submicroscopic level, for example, “ferric chloride reacts with the phenolic hydroxyl of salicylic acid forming an iron-phenol complex” (R2). Perhaps because it is a section in which the result of the experiment is described and explained, involving the obtaining of the product through a chemical reaction, the momented expectancy activities have been little evidenced in this group of reports. This type of activity is common in the description of experimental procedures (Doran & Martin, 2021), but this was not confirmed in the data from group 1 reports.

When we analyzed the reports from group 2 with regard to the type of activity, we noticed that the students inserted part of the experimental procedure in the Results and Discussion section. As a consequence, most of the activities were classified as momented expectancy activity, for example, “after stirring, two phases were formed, and we collected only the organic phase in the Erlenmeyer, then we added sodium sulfate” (R14). Momented implication activities were evidenced when the student explains the function of an item used in the experiment, whether this explanation is at the macroscopic level, for example, “with the evaporation of the solvent, the formation of crystals occurred, which is pure caffeine” (R12), or at the submicroscopic level, for example, “calcium carbonate is used because the basic medium promotes the hydrolysis of the caffeine-tannin salt, increasing the yield of caffeine extraction” (R17).

In the reports from group 3, students rarely describe the field using sequences of activities. In most of the text, students explain the solubility of one substance in another by referring to their characteristics, for example, “acetanilide is insoluble in water. This happens because acetanilide has eight carbons in its molecular structure” (R19). However, implication activities were identified, for example, “methanol has a hydroxyl in its structure that interacts with a hydroxyl of salicylic acid, thus forming a hydrogen bond” (R19) and “there will be the donation of a proton from hydrochloric acid to aniline, thus forming a salt” (R13). We identified, in this type of activity, a subtype, in which the intensity of one activity does not imply another, but prevents another activity from happening, for example, “the strong hydrogen bonds (molecular interaction) of (between) water molecules hinder solubilization” (R25).

Most of the time, the activities aim to verify whether the solute studied is soluble or not in a certain solvent, for example, “using hexane as a solvent for acetanilide, the solute is partially soluble” (R20). If we consider that the student, in a deviation from standard language, omits some words, we could rewrite the previous sentence as a sequence of expectancy activity by adding the activity to verify or observe, “using hexane as a solvent of acetanilide, (we verified that) the solute is partially soluble”. It was observed that in the writing of the reports from this group, students basically used five construction structures involving activities (Figure 11).

Figure 11*Construction of sentences in the text involving activities*

Type	Example
Activity leading to property	Both salicylic acid and aniline, used as solutes, are <u>insoluble</u> in all solvents (R22) As this interaction is very strong, these liquids are completely <u>soluble</u> (R24)
Property leading to activity	Considering that they are highly <u>nonpolar</u> , solubilization was total (R5)
Ownership not leading to an activity	Water and hexane are also immiscible because one compound is <u>polar</u> and the other compound is <u>nonpolar</u> , making it difficult for their respective molecules to interact (R22)
Chemical structure leading to an activity	The presence of many carbons in their structural form, and a hydrophilic group (NH), allows a small interaction with water (R21)
Chemical structure preventing an activity	It will not be possible to make a hydrogen bond, because hexane does not have hydrophilic groups (R21)

These isolated activities do not seem to fit the classification proposed by Doran and Martin (2021), for example, for unmomented activities, because, despite being isolated activities, they do not encapsulate other activities as proposed by the authors. Other activities were also not classified because the students used other verbs to indicate an identifying process, for example, in “acetanilide presented itself as partially soluble in water” (R21), the verb “presented” does not represent an activity, as it is only indicating a property, and can be replaced by the verb “to be”, so that the sentence would be rewritten as “acetanilide **is** partially soluble in water”.

Property System

If we consider the two types of properties: qualitative and spatio-temporal, we can say, based on the analysis of the texts, that students make little use of properties associated with items and activities. In reports involving a chemical reaction (group 1), most of the properties are those observed in the macroscopic level, such as the shape and color of the product of the reaction, for example, “after 30 minutes, we noticed the formation of a precipitate, white-colored crystals” (R4), and coloration of the system submitted to the product identification test of this reaction, for example, “the mixture turned violet” (R5).

Two examples were identified in which students used the spatio-temporal property. In both cases, the property refers to the position of an atom and is used when students are explaining the mechanism of the reaction. In R3, the word **adjacent** It is used to make a distinction between two oxygen atoms, “in the third stage, we have the migration of hydrogen, from the oxygen that has bonded to the substrate to an adjacent

oxygen”. The same occurs in R7, with the word **central** in “a pair of electrons from the central oxygen of the anhydride group attacks a hydrogen that is bonded to a protonated oxygen”. Some qualitative properties were arrayed using words that denote intensity, for example, “acetylsalicylic acid is **slightly soluble** in water” (R1) and “strong acids **accelerate** the speed of reactions” (R7).

In group 2, we also identified the limited use of property by the students. In addition to the qualitative properties of shape, solubility, and color, some students, when describing an item, attributed to it an arrayed and gauged property, for example, “The need to let the mixture reach room temperature is due to this solvent having a boiling point around 40°C” (R12) and “caffeine is about 9 times more soluble in dichloromethane than in water” (R15). In none of the students’ texts or in the synthetic text of this group were spatio-temporal properties evidenced. Some properties seem to depend on a circumstance, for example, “at elevated temperatures, caffeine solubility increases” (R10); however, the circumstance factor is not addressed in Doran and Martin’s (2021) model.

By analyzing the property system of group 3, it was observed that in these texts, perhaps because it is an experiment that analyzes a property, the number of properties mentioned was much higher than that found in the texts from groups 1 and 2. All the properties identified are of the qualitative type. Some of these properties are described as characteristic of the item, for example, “methanol is a polar molecule” (R19) and “it is a white or colorless ionic solid, which is the potassium salt of phthalic acid” (R20). Other properties are arrayed, for example, “*weak* dipole moment” (R20), and others are arrayed and gauged, for example, “hexane is a hydrocarbon with a dipole moment equal to zero” (R20) and “aniline has a lower density (1.02 g/m³) than NaOH, which has *higher density* (2.13 g/cm³)” (R25). Properties not arrayed but gauged were also identified, for example, “the pKa of water is about 16” (R27).

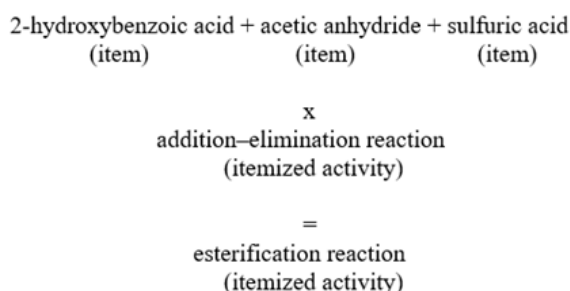
Interdependency System

The interdependency system for field analysis proposed by Doran and Martin (2021) is perhaps the most difficult to identify in the texts. In the reports from group 1, when we analyzed the elements of the field (items, activities, and properties) that went through the reconstrual process, we found the “reaction” activity. This activity has the characteristic of an item, that is, it constitutes a classification taxonomy, “esterification reaction” (R3), “acetylation reaction” (R6). Thus, the activity **reaction** is reconstrued as an itemized activity, which Hao (2020) classifies as an activity entity. A qualitative property **coloration** is another element of the field that was reconstrued as an item (itemized property), since in R5, “the mixture turned slightly red”. The student, when grading the intensity of the property, places it within a classification taxonomy. The reconstrual of property as an item was also identified in the synthetic text from group 1, in which the partial charge property occurs as positive, suggesting that the charge can exist in another form, the negative. No other types of reconstruals were evidenced in the texts from this group.

Also in group 1, a single complete example of interdependence by interrelation was evidenced. In R3, the student relates three items by extension: 2-hydroxybenzoic acid, acetic anhydride, and sulfuric acid to an addition-elimination reaction, which the student later names as an esterification reaction. In Figure 12, we represent this interrelation using the symbology of Halliday and Mattiessen (2014): + for extending, x for enhancing, and = for elaborating.

Figure 12

Interrelation between items evidenced in the student's text



Other examples do not include the elaborating step, for example, in R2, the items salicylic acid and acetic anhydride are interrelated by extending in an acid medium (a circumstance). This interrelating has as enhancing the formation of acetylsalicylic acid, another item that does not go through the elaborating stage.

In the texts from groups 2 and 3, among the types of reconstrual suggested by Doran and Martin (2021), activities reconstruing as items were evidenced, the Activity entities, for example, **hydrogen bonding** and **intermolecular interaction**. Property such as the strength of a base is reconstrued as an item (itemized property), in the form of an ordering, classified as **strong bases** and **weak bases**, and constructing classification taxonomies. The interdependency system by interrelating was not observed in the texts from group 2, but one example was identified in a text (R30) from group 3, when the student relates items from the submicroscopic world, H⁺ ions and OH⁻ ions, to a neutralization reaction. The elaborating as the last step of the interrelating process was not evidenced for this group of texts.

Entity Mapping

Items and activities as elements of the field are carried out in the semantic discourse through entities. Items are performed through thing, instrumental, occurrence, or inferable entities, and activities, when reconstrued as items, are performed by the reported entities if they occur in the macroscopic world, and inferable entities when they occur in the submicroscopic world.

The entities categorized as thing entities were predominant in these texts from group 1 (Table 1). Regarding the thing entities used by the students, most refer to linguistically defined instrumental thing entities (ECIId), as the names of the substances

involved in the esterification reaction. The difference occurs in R3 and in the synthetic text (S8), because in these texts, there is a more detailed explanation of the mechanism of reaction, including figures showing the structures. Explaining in detail how the reaction occurs requires the use of inferable thing entities (ECInf), such as nucleophiles, which are found only in the submicroscopic world. Despite presenting a greater number of linguistically defined instrumental thing entities (ECIld), the text R6 also presents a large number of inferable thing entities (ECInf) because, like R3 and S8, it also presents the reaction mechanism in detail through a figure. When there is no explanation of the mechanism of reaction or the identification test, ECInf are absent, as in R4, or have relatively lower percentages, as in the rest of the reports.

Table 1

Data from the mapping of the texts in the Results and Discussion section of the esterification reaction experiment reports (group 1), in percentage

Report	ECIod	ECIld	ECO	ECInf	EAR	EAIInf	% terms
R1	1,00	10,03	1,67	4,35	1,00	2,34	20,39
R2	1,73	9,38	0,69	5,56	0,35	3,47	21,18
R3	2,64	8,30	0,38	9,81	1,13	6,79	29,05
R4	2,06	7,35	1,47	0	0	0,88	11,76
R5	2,25	8,39	0,97	2,58	0,65	2,25	17,09
R6	2,63	16,84	0,53	7,89	1,05	7,89	36,83
R7	0,98	5,15	0,74	5,39	0	2,45	14,71
S8	0,75	6,20	0,75	9,02	0,75	2,63	20,10

It was expected that, in the explanation of the reaction, a greater number of inferable activity entities (EAIInf) would also be used, such as protonation, but this was only observed in R3 and R6. Most students did not even identify the type of reaction involved in the experiment; only R3 and R6 made this identification, as well as the synthetic text S8. In R6, the number of EAIInf is greater because this student repeats a few times the type of reaction being performed and uses other activity entities that do not appear in the other texts, such as **dehydration**, referring to an aspect that is not directly related to the experiment, for example: “acetic anhydride is produced from the dehydration of acetic acid, and the reverse reaction occurs when adding water to acetic anhydride”. In text R3, the greater number of inferable activity entities (EAIInf) is because the student presents a detailed figure of the mechanism and a detailed description of this figure.

The activity entities categorized as reported (EAR) are limited to the terms **identification test**, **test**, and **experiment**, and appear minimally in the analyzed texts, which is within the expected, since most students in this group used the text to describe and explain the reaction mechanism, and to describe and explain the test to confirm the formation of the product. Likewise, the ostensibly defined instrumental thing entities (ECIod) were also little evidenced, being restricted to the term “**test tube**” except for

text R4, which introduces the analyzed section with a summary of the experimental procedure. The occurrence thing entities (ECO) in this group were limited to the terms **product**, **precipitate**, but mainly **crystal**, and is expected since the experiment provides, at the end of the process, a product of reaction.

When we look at group 2, the caffeine extraction experiment, we observe a greater number of entities classified as activity entity reported (EAR), (Table 2), which is expected since these entities are associated with this type of experiment. The lower number of ECInf and EAInf is due to the fact that students keep writing the text in the macroscopic world, and when inferable entities appear, they are restricted to the complex salt formed between caffeine and tannin as an inferable thing entity, and the hydrolysis reaction and the term reaction, as activity entities that occur in the submicroscopic world. In all the texts from group 2, the students mention glassware and equipment necessary for the extraction technique (ECIod), and the occurrence entities (ECO) are mostly limited to the term “crystal”, the product of the extraction technique. The number of mapped entities in group 2 was lower than in group 1, but similar to this group, most entities are of the linguistically defined instrumental thing entities (ECIId) type.

Table 2

Data from the mapping of the texts in the Results and Discussion section of the caffeine extraction experiment (Group 2) reports, in percentage

Report	ECIod	ECIId	ECO	ECInf	EAR	EAInf	% terms
R9	2,63	3,38	1,13	0	1,13	0	8,44
R10	1,06	5,82	1,58	0	1,58	0	10,04
R11	2,02	9,09	0,34	0	2,69	0	14,14
R12	2,05	9,09	0,58	0,29	1,76	0,29	14,06
R13	1,52	9,16	0,76	0	2,29	0	13,73
R14	2,16	3,88	0,86	0	4,31	0	11,64
R15	0,74	11,85	0,37	0	4,07	0,37	17,40
R16	0,41	8,26	0,82	0,82	2,89	0,41	13,61
R17	2,70	8,91	0,27	0,27	2,16	0,27	14,58
S18	0,52	11,81	0,52	0,79	2,62	0,79	17,05

Except for the synthetic text (S18), only two students identified the technique used in this experiment. This identification enriches the meaning of the reported activity entities (EAR) **process** and **method**, for example, “liquid-liquid separation process” (R13) and “The method used in caffeine extraction was solid-liquid, but mainly liquid-liquid” (R15). The terms classified as activity entity reported (EAR) were, for the most part, extraction, procedure, heating, and cooling. We also identified in most of the students’ texts and in the synthetic text, the activity **evaporation**, but this was not classified as EAR, because, according to Hao (2020), EAR is an activity carried out by a person entity and the evaporation process described in the students’ texts refers to a spontaneous phenomenon that occurs due to the action of a thing entity, heating.

In the texts from this group, some linguistically defined instrumental entities (ECIld) are also measured with quantitative measures, for example, “10 mL of dichloromethane” (R11). This happens when the student inserts in the Results and Discussion section some aspects of the experimental procedure performed, which can be observed in 5 of the analyzed texts, or when the student explains some characteristic of the entity. In the synthetic text (S18), the measured dimension of an entity is little evidenced, dealing only with income values, an entity not considered in this study. According to Hao (2020, p. 86), “Measured dimensions correspond to the quantified qualities of an entity”, for example, “[dichloromethane] has a boiling point of around 40°C” (R12). Although in the field of chemistry, “boiling point” is understood more as a property than as a quality, it can be quantified in a specific way, 40°C. Other properties associated with entities are quantified using what Hao (2020, p. 87) calls **general quantifications**, for example, “dichloromethane, which is a highly polar solvent” (R13).

Another characteristic in the texts from this group is that the students describe functions for some entities of the instrumental thing entity linguistically defined (ECIld) type as a way of explaining the reason why this substance was used in the experiment, for example, “calcium carbonate in the solution, it has the function of reducing this emulsion and facilitating the extraction of caffeine” (R17). Other times, the students describe characteristics of the entities, for example, “in the fourth process, anhydrous sodium sulfate, a drying substance, is added” (R15).

In all the texts from this group, the students describe the circumstances in which the processes occur. **Water**, for example, was classified as ECIld in most cases, but when measured, the term **water** was classified as a circumstance, for example, “The heating is also due to the fact that caffeine is more soluble in **hot** water than in **cold** water” (R17).

Basically, the entities mapped in the texts from group 3 (Table 3) are of three types: linguistically defined thing instrumental entities (ECIld), which, as in groups 1 and 2, are the ones with the largest number; and, also, inferred thing entities (ECInf) and inferred activity entities (EAIInf). Only 3 of the texts include names of glassware, for example, “test tube” (R19 and R29), and in 6 texts, the reported activity entities (EAR) identified are restricted to “solubility test” (R31), “procedure” (R24 and R27), “laboratory experiments” (R22), and “miscibility test” (R26). Since in this experiment substances are only mixed to verify their solubility, occurrence entities, such as precipitate and gas formation, are not evidenced.

Table 3

Data from the mapping of the texts in the Results and Discussion section of the solubility experiment reports (Group 3), in percentage

Report	ECIod	ECIld	ECo	ECInf	EAR	EAIInf	%terms
19	0,83	11,24	0	3,62	0	1,40	17,09
20	0	23,03	0	2,68	0	1,73	27,44
21	0	13,61	0	4,66	0	2,98	21,25
22	0,11	11,28	0	3,25	0,46	2,44	17,54
23	0	16,76	0	3,23	0	0,61	20,60
24	0	10,03	0	2,76	0,87	1,31	14,97
25	0	13,07	0	5,50	0	3,21	21,78
26	0	21,49	0	2,03	1,36	1,36	26,24
27	0	15,17	0	2,47	0,88	1,06	19,58
28	0	19,07	0	2,60	0	0,88	22,55
29	0,38	13,58	0	0,89	0,13	1,78	16,76
30	0	15,44	0	2,97	0	2,17	20,58
31	0	14,03	0	5,79	0,39	2,89	23,10

During the explanation of the solubility phenomenon, many taxonomic relations are established, particularly compositional relations (Doran & Martin, 2021), for example, “in its molecule there is a large hydrophobic part” (R22) and “acetanilide [...] has carbonyl in its structure” (R23). The terms **hydrophobic part** and **carbonyl** as well as other similar terms that are described as a part of a whole, have not been classified as entities. Although the term **carbonyl** could be considered, in principle, an entity that maintains a co-elaboration relation with the entity **acetanilide** (Hao, 2020), the term **hydrophobic part** does not designate an entity. However, both terms designate parts of a whole, enriching the meaning of the entities with which they relate, in these examples, molecule and acetanilide.

The entities instrumental thing entity linguistically defined (ECIld), thing entity inferable (ECInf), and activity entity inferable (EAIInf) are used by students to explain the solubility property through the polarity of substances, intermolecular interactions, and molecular structure. Sometimes students order these entities by adding words that confer them gradation, for example, “methylene chloride makes an induced dipole-dipole interaction (EAIInf), which are **weaker** bonds and cannot break the hydrogen bonds (EAIInf) in water because it is **stronger**” (R19).

In other cases, the entity is dimensioned using “general quantifications, e.g. small, large, long, short” (Hao, 2020, p. 87), within the measured dimension, for example, “biphthalate has a hydrocarbon chain *much larger* than that of the solvent” (R19).

The ordering of items (thing entity), as the case of “hydrocarbon chain” (R19) and of activity entity reconstrued as item (activity entity), in this case, **induced dipole-dipole interaction**, is not foreseen in the model proposed by Doran and Martin (2021),

in which only properties can be arrayed and gauged (see Figure 3). Hao's (2020) model allows this ordering within the concept of dimensionality. The measured dimension, therefore, allows the dimensioning of not only the properties, but also includes the other elements of the field when they are seen from the perspective of entities in semantic discourse.

Comparing the mapping of the entities in the three texts analyzed, we found that in group 1, which presents a description of the experiment at the macroscopic level and an explanation at the submicroscopic level, it was possible to map all types of entities. Group 3 was the one with the lowest diversity of entities. Even though the macroscopic and submicroscopic worlds were related, the experiment did not involve technique or product production, which contributed to the lower diversity.

The average number of terms referring to entities used in the texts was about 21% for groups 1 and 3 and 13% for group 2. The rest of the terms present in the texts are words that indicate properties, dimensions, and intensities of these properties, activities performed by verbal groups that are not classified as reported activity entities (EAR) and inferable activity entities (EAInf). In addition, words that indicate circumstances, other entities not mapped here, such as semiotic entities and other words with a function in grammar, for example, prepositions, definite and indefinite pronouns, were not classified.

In the students' texts, the activities, whether as a description of the phenomenon in the macroscopic world or an explanation of the phenomenon in the submicroscopic world, are carried out by verbs, such as "the hydroxyl of salicylic acid attacks the polarized carbon of acetic anhydride" (R6). This may explain the few activities found in the nominalized form, although from a technical point of view, the scientific text should be more abstract, with the activities being described in the nominalized form (Halliday & Martin, 1993). Another consequence of students not using nominalization processes in their writing is the absence of live grammatical metaphors in their texts. Grammatical metaphor is an important element of scientific language, as it has great potential to build knowledge and compose texts (Hao, 2021).

The use of a grammatical metaphor was only evidenced in the synthetic text of group 1, when the mechanism of reaction is explained. In both cases, the noun (in bold) in the second clause encapsulates the information contained in the preceding clause:

In this step, a carbonyl group of acetic anhydride is attacked by the sulfuric acid present in the reaction medium. In this **attack**, a pair of free oxygen electrons from the carbonyl group forms a bond with a proton donated by the acid, as shown below (S8).

A proton is transferred between oxygen atoms within the molecule. This **transfer** makes part of the molecule a good leaving group (S8).

Figure 13 presents a systematization of the main findings from the students' reports in articulation with the field analysis model proposed by Doran and Martin (2021) and the entities and dimension model proposed by Hao (2020).

Figure 13*Systematization of the analysis of the reports*

Aspect analyzed		Group 1 (Chemical reaction)	Group 2 (technique)	Group 3 (property)
The field in the context of chemistry	Static perspective	Some reports provide examples of compositional taxonomy, but no classification taxonomy was evidenced.	Classification taxonomy is only suggested; there is no evidence of compositional taxonomy.	Both the taxonomy of classification and composition were evidenced.
	Dynamic perspective	Most of the activities described are momented and implication. Unmomented activities are restricted to titles that encapsulate activities.	Most activities described are momented expectancy because students bring part of the experimental procedure to the analyzed section. There is evidence of momented implication activities, but no unmomented activities were found.	Only a few examples of momented implication activities were found. The activities are described as a product of other elements, such as chemical property and structure, not described in the evaluated model.
	Property	Students use qualitative properties of color and shape to describe the macroscopic world and spatiotemporal properties to describe the submicroscopic world.	Only the qualitative properties were evidenced, and some were arrayed and gauged. Some properties seemed to depend on a circumstance, a fact not described in the model evaluated.	Only the qualitative properties were evidenced, and some were arrayed and gauged. Some properties were just gauged and were not previously arrayed with the model evaluated.
	Interdependency	Few cases of reconstruing of activities as items were evidenced, and only one case of complete interrelating and two others that relate the dependence of an activity to the interrelating between items.	Few cases of reconstruing of activities as items and no cases of interrelating were evidenced.	Few cases of reconstruing of activities as items and no cases of interrelating were evidenced.

Figure 13*Systematization of the analysis of the reports (continuation)*

Aspect analyzed	Group 1 (Chemical reaction)	Group 2 (technique)	Group 3 (property)
Mapping entities in the semantic discourse that realizes the field	All categories of mapped entities describing the macroscopic and submicroscopic worlds were highlighted.	Most of the mapped entities were entities describing the macroscopic world. The inferable thing entities and activity entities were little evidenced.	Most of the mapped entities are grouped into three categories, with the entities that describe the submicroscopic world being the most evidenced.

Conclusion

In this study, we mapped different elements of the field (the taxonomic relations between items, the types of activities, and the types of properties) and their realizations within the stratum of the semantics of language discourse in the form of entities and dimensions of these entities, using the field analysis models of Doran and Martin (2021) and the entities and dimensions analysis models of Hao (2020).

Considering the texts of the three groups of reports analyzed here, we observed that the students, despite the few examples identified, construe taxonomies used to do the classification and part-whole composition items. This was more evident in group 3 when students associated a property with an item or thing entity, as a way to explain the phenomenon experienced, whether at the macroscopic or submicroscopic level. In the texts from groups 1 and 2, we found only a potential for taxonomy construction in which the students described items associated with properties, but did not explain any order of classification. The classification taxonomy seems to be independent of the type of experiment, while the composition taxonomy was only evidenced when the students explained the phenomenon at the submicroscopic level, which was the case of the texts from groups 1 and 3.

In terms of activity, the analysis of the students' texts showed that the type of activity depends on the type of experiment. The texts from group 1, involving chemical reaction, present a greater number of momented implication activities than the texts from groups 2 and 3, due to the fact that in this group of reports, the students describe in detail the mechanism of the reaction in stages in which one depends/implies another, in addition to associating evidence observed in the macroscopic world as a result of an activity that occurs in the submicroscopic world. As the students from group 2 described part of the experimental procedure within the section analyzed, most of the activities found in the texts from this group were momented expectancy activities. In group 3, not many sequences of activities were observed, but isolated activities were identified that did not fit the Doran and Martin (2021) classification of activities, requiring further analysis in a subsequent study on this type of activity.

Although the students did not attribute properties to items or activities, both types of property — qualitative and spatio-temporal — were identified in the analyzed texts, with the latter being found only in the texts from group 1. The interdependency relation among the elements was little observed in the reports and was basically limited to the texts from group 1. In most cases, interdependency occurs through the reconstruing of an activity as an item (activity item) or a property as an item, and when it occurs through interrelating, students relate the items by extending and enhancing, but do not elaborating.

The entity mapping showed that the diversity of entities used in the texts depends on the type of experiment. The texts from group 1 present a greater diversity of entities, while in the texts from group 3, basically three types of entities were identified. The entities in greater quantity in all the groups of texts analyzed referred to the substances used in the experiments, the linguistically defined instrumental entities, due to the fact that the students repeated the names of the substances several times in the text, either in the description or in the explanation of the phenomenon, and associated with these entities. The students described measured dimensions and properties, and sometimes the function of the entity within the experiment.

In this analysis, we found that students have difficulty describing the submicroscopic world. The texts from group 2, for example, described and explained the extraction technique only at the macroscopic level, which limited the diversity of inferable entities. Another observation concerns the nominalization of activities. Few nominalized terms were identified, and, consequently, there was no construction of a live grammatical metaphor, an important resource of scientific language, in the students' texts.

Considering only the students' texts analyzed in this work, we can suggest that Doran and Martin's (2021) field model and Hao's (2020) entity and dimensionality model, with some modifications, can be used to analyze the writing used by students to describe and explain phenomena in the field of chemistry. However, the texts used here refer to a specific area of chemistry, organic chemistry. It is therefore necessary to extend this study to the areas of analytical, inorganic, and physical chemistry to examine how the different contents of chemistry reflect different types of use of field models and the model of entities and dimensionality.

Authors' Contribution

Data curation: Feitosa, E. M. A.; **Investigation:** Feitosa, E. M. A.; **Methodology:** Feitosa, E. M. A., Mortimer, E. F.; **Project administration:** Mortimer, E. F.; **Supervision:** Mortimer, E. F.; **Writing – original draft:** Feitosa, E. M. A.; **Writing – review & editing:** Feitosa, E. M. A., Mortimer, E. F.

Data Availability Statement

The data will be provided upon request.

References

- Bortnik, B., Stozhko, N., Pervukhina, I., Tchernysheva, A., & Belysheva, G. (2017). Effect of virtual analytical chemistry laboratory on enhancing student research skills and practices. *Research in Learning Technology*, 25, 1–20. <https://doi.org/10.25304/rlt.v25.1968>
- Carvalho, A. M. P. C. (2018). Fundamentos Teóricos e Metodológicos do Ensino por Investigação. *Revista Brasileira de Pesquisa em Educação em Ciências*, 18(3), 765–794. <https://doi.org/10.28976/1984-2686rbpec2018183765>
- Clary-Lemon, J., Gervacio, R., & Latimer, D. (2019). Writing as a mode of learning: staged approaches to chromatography and writing in the undergraduate organic lab. *Journal Chemical Education*, 96(5), 965–969. <https://doi.org/10.1021/acs.jchemed.9b00072>
- Crawford, G. L., & Kloepper K. D. (2019). Exit Interviews: Laboratory Assessment Incorporating Written and Oral Communication. *Journal Chemical Education*, 96(5), 880–887. <https://doi.org/10.1021/acs.jchemed.8b00950>
- Doran, Y. J., & Martin, J. R. (2021). Field relations: Understanding scientific explanations. In K. Maton, J. R. Martin, & Y. J. Doran (Orgs.), *Teaching science: Knowledge, language, pedagogy* (pp. 105–133). Routledge.
- Dreyfus, S., & Hao, J. (2020) A multi-stratal perspective on circumstantial meaning. A multi-stratal perspective on circumstantial meaning. *Journal of World Languages*, 6(1), 27–45. <https://doi.org/10.1080/21698252.2020.1720159>
- Fuzer, C., & Cabral, S. R. S. (2014). *Introdução a gramática sistêmico-funcional em língua portuguesa*. Mercado de Letras.
- Halliday, M. A. K. (1978). *Language as social semiotic: The social interpretation of language and meaning*. Edward Arnold.
- Halliday, M. A. K. (1993). Some Grammatical Problems in Scientific English. In M. A. K. Halliday, & J. R. Martin (eds.), *Writing Science: literacy and discursive power*. University of Pittsburgh Press.
- Halliday, M. A. K. (1998). Things and relations. In J. R. Martin, & R. Veel (Orgs.), *Reading science: critical and functional perspectives on discourses of science* (pp. 185–235). Routledge.
- Halliday, M. A. K., & Martin, J. R. (2005). *Writing Science: literacy and discursive power*. Taylor & Francis e-Library.
- Halliday, M. A. K., & Hasan, R. (1985). *Language, context, and text: Aspects of language in a social-semiotic perspective*. Oxford University Press.

Hao, J. (2018). Reconsidering “cause inside the clause” in scientific discourse: From a discourse semantic perspective in systemic functional linguistics. *Text & Talk*, 38(5), 525–550. <https://doi.org/10.1515/text-2018-0013>.

Hao, J. (2020). *Analysing scientific discourse from a systemic functional linguistic perspective: A framework for exploring knowledge-building in biology*. Routledge.

Hao, J. (2021). Building taxonomies: A discourse semantic model of entities and dimensions in biology. In K. Maton, J. R. Martin, & Y. J. Doran (Eds.), *Teaching science: Knowledge, Language, Pedagogy* (pp. 134–161). Routledge.

Hoffa, D., & Freeman, S. A. (2008). An Evaluation of Student Performance on Traditional vs. Synopsis Laboratory Reports in Industrial Technology. *International Journal for the Scholarship of Teaching and Learning*, 2(1), 1–13. <https://doi.org/10.20429/ijstl.2008.020111>

Johnstone, A. H. (2000). Teaching of chemistry - logical or psychological? *Chemistry Education: Research and Practice in Europe*, 1(1), 9–15.

Malinowski, B. (1923). The problem of meaning in primitive languages. Supplement 1. In C. K. Ogden, & I. A. Richards (eds.), *The Meaning of Meaning: A Study of the Influence of Language upon Thought and of the Science of Symbolism* (pp. 297–336). Harcourt, Brace & World, Inc.

Martin, J. R. (1993). Literacy in science: Learning to handle text as technology. In M. A. K. Halliday, & J. R. Martin (eds.), *Writing science: Literacy and discursive power* (pp. 166–202). University of Pittsburgh Press.

Paton-Walsh, C. (2015). Can Report Templates Aid Student Learning in Undergraduate Chemistry Laboratory Classes? *Journal of University Teaching & Learning Practice*, 12(2), 1–12.

Sampson, V. & Walker, J. P. (2012) Argument-Driven Inquiry as a Way to Help Undergraduate Students Write to Learn by Learning to Write in Chemistry. *International Journal of Science Education*, 34(10), 1443–1485. <https://doi.org/10.1080/09500693.2012.667581>

 **Edinilza Maria Anastácio Feitosa**

Universidade Estadual do Ceará
Fortaleza, Ceará, Brasil
edinilza.feitosa@uece.br

 **Eduardo Fleury Mortimer**

Universidade Federal de Minas Gerais
Belo Horizonte, Minas Gerais, Brasil
efmortimer@gmail.com

 **Editor in charge:** Suzani Cassiani

Translated by: Sofia Bocca

Journal financed by Associação Brasileira de Pesquisa em Educação em Ciências — ABRAPEC



Manifestation of Attention to Good Scientific Practices and Exemption from Interest and Responsibility

The authors declare that they are responsible for complying with the ethical procedures provided by law and that no competing or personal interests could influence the work reported in the text. They assume responsibility for the content and originality, as a whole or in part.

Copyright (c) 2025 Edinilza Maria Anastácio Feitosa, Eduardo Fleury Mortimer



This text is under a **Creative Commons CC BY 4.0 License**

You are free to Share (copy and redistribute the material in any medium or format for any purpose, even commercially) e Adapt (remix, transform, and build upon the material for any purpose, even commercially). Under the following license terms:

Attribution: You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

No additional restrictions: You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.
