Original Article

LEXICAL CHOICE IN PAIRED WRITING: ORAL ERASURE, UNFOLDED COMMENTING AND METALINGUISTIC ACTIVITY IN CHILDREN'S TEXTUALIZATION

Eduardo CALIL[*]
Cristina FELIPETO[**]
Maria Hozanete Alves de LIMA[***]

▪ **ABSTRACT:** This article investigates how metalinguistic activity in the semantic-lexical domain emerges during children's collaborative writing. Since research on spontaneous processes of lexical reflection remains limited – largely confined to elicitation protocols that displace the student from the position of "author in action" – this study follows a dyad of second-grade primary school students across ten processes of producing invented etiological tales. The approach integrates Textual Genetics and Benveniste's theory of enunciation, drawing on the Ramos System protocol, which synchronizes audio and graphic records, making it possible to grasp the manuscript *in statu nascendi*. Sixteen semantic-lexical textual objects and thirty unfolded comments were identified, organized along a *continuum* of metalinguistic explicitness ranging from lexical definition and comparison to description of the referent. Self-glossing stood out as a gesture of lexical retention that constitutes the most explicit form of oral erasure. The two students displayed complementary modes of functioning: one through autonomous definition, the other through dialogical confrontation. We conclude that lexical reflection is constitutive of the act of writing, even when it leaves no graphic traces.

▪ **KEYWORDS:** Textual Genetics; Real-time text production; Enunciation; School manuscript; Literacy.

Introduction

Over the past few decades, research on school writing has undergone a significant theoretical and methodological shift. Under the influence of the founding contributions

[*] Federal University of Alagoas (UFAL), Graduate Program in Education (PPGE); Graduate Program in Linguistics and Literature (PPGLL), Maceió, AL, Brazil. calil@cedu.ufal.br. ORCID: https://orcid.org/0000-0002-8696-3697.

[**] Federal University of Alagoas (UFAL), Graduate Program in Linguistics and Literature (PPGLL), Maceió, AL, Brazil. crisfelipeto@fale.ufal.br. ORCID: https://orcid.org/0000-0003-3729-0796.

[***] Federal University of Rio Grande do Norte (UFRN), Graduate Program in Language Studies (PPGEL), Natal, RN, Brazil. hozanetelima@gmail.com. ORCID: https://orcid.org/0000-0001-8916-2740.

This work is licensed under a Creative Commons Attribution 4.0 International License.

of Genetic Criticism (Hay, 1979; Grésillon, 1994), the drafts and manuscripts produced by students have ceased to be conceived as ephemeral residues of the teaching-learning process and have come to be recognized as privileged objects of investigation. In the field of writing didactics, work such as that of Fabre (1990) and Boré (2014) contributed decisively to this change of status by showing that the draft of a text written in the classroom constitutes a space of elaboration, of reflexive return and of the progressive construction of the writing subject. From this perspective, the text written by the student comes to be understood not merely as an assessable product but as a process, one in which cognitive, linguistic and discursive operations that are fundamental to learning to write are inscribed.

This object of investigation, treated as a "school manuscript" (Calil, 2008) by virtue of its intricate and constraining didactic-pedagogical implications, has afforded an innovative approach to the writing process in the here and now of the classroom: it has come to be understood not only as a space for learning and for building knowledge about written language, but also from the standpoint of its genetic, creative and authorial dimension. In seeking to understand the dynamics of the writing process among young learners and adolescents, particularly in didactic contexts of collaborative production (Daiute; Dalton, 1993; Calil, 1998; Barbeiro, 1999; Camps, 1995; De Gaulmyn, 1994; Vass *et al.*, 2008), a methodological approach centered on the genesis of the school manuscript as it is built in real time emerges as a privileged field for observing how students think about their own writing: a space in which cognitive operations, metalinguistic activities, gestures of revision and emergent forms of authorship are manifested. Paired writing intensifies the enunciative dimension of the process, making audible debates and hesitations that, in individual writing situations, would remain silent or merely hinted at by written erasures.

From a psycholinguistic standpoint, this capacity to take language as an object of reflection has been described in terms of metalinguistic activity (Gombert, 1990), whose development is heightened by the formal learning of reading and writing. In children's writing, such activity may be manifested implicitly, through actions performed on the text, or explicitly, through verbalizations bearing on different domains of written language. Following Gombert (1990), the notion of metalinguistic activity is taken here in a broad sense, covering everything from epilinguistic behaviors, that is, effective adjustments to language carried out without deliberate control, to metalinguistic behaviors proper, which are conscious and intentional. We reserve metalinguistic awareness for the capacity corresponding to this second pole, and semantic-lexical awareness for its specification within the domain focused on here, taking as its empirical indicators those verbalizations in which language is explicitly thematized. From this follows a caution that guides the analysis throughout: the data give direct access to degrees of explicitness in the verbalizations and only inferential access to states of awareness. We likewise use emergence not in a developmental sense, as though it were a matter of the first ontogenetic appearance of this capacity, but in an enunciative

sense: the situated appearance, in the very course of textualization, of manifestations of reflection on language.

When these oral manifestations bear on the lexicon, they still occupy a place that has been little explored in the literature. The importance of vocabulary for the construction of textual meaning is widely acknowledged in more than one research tradition: measures of lexical diversity and lexical sophistication in the text are associated with writing quality as early as the second year of schooling (Olinghouse; Leaird, 2008); the didactics of the lexicon has consolidated itself as a field of investigation in its own right (Grossmann, 2011); and intervention studies show that the structured teaching of lexical units has an effect on the quality of written production (Duin; Graves, 1987; Garcia-Debanc *et al.*, 2009; Roubaud; Moussu, 2012). This acknowledgment, however, operates either on the finished product or on the teaching pole: research on the semantic-lexical metalinguistic activities that emerge spontaneously during the writing of newly literate students remains limited. A significant proportion of the existing studies resorts to verbal elicitation protocols or retrospective interviews, which displace students from their position as authors in the act of writing and capture *a posteriori* rationalizations of the process (Nisbett; Wilson, 1977; Gufoni, 1996; Smagorinsky, 2001). Even research that analyzes spontaneous comments in collaborative writing situations (Daiute; Dalton, 1993; Barbeiro, 1999; Camps *et al.*, 2021) faces instrumental limitations, above all when the records are restricted to orality, dissociating the comments from the exact moment of inscription and linearization of the manuscript.

The present article proposes to advance this field by analyzing metalinguistic activity in the semantic-lexical domain from a genetic and enunciative perspective, centered on following the school manuscript as it is being built *in statu nascendi*. Drawing on the methodological protocol of the Ramos System (Calil, 2020), which articulates synchronized records of orality, of the writing gesture and of the graphic trace, we investigate how a dyad of second-year elementary school students recognizes, comments on and justifies lexical choices during real collaborative writing situations. This approach makes it possible to grasp both what was actually inscribed in the final text and the set of lexical possibilities that competed to occupy the same position in the syntagmatic chain.

The general aim of the study is therefore to understand how semantic-lexical metalinguistic activity is manifested during children's collaborative writing, articulating orality and writing in the very course of textualization. More specifically, it seeks to: (i) identify the semantic-lexical textual objects recognized by the dyad across different production processes; (ii) describe when and how these recognitions occur, distinguishing profiles of metalinguistic functioning; and (iii) typify the unfolded comments produced by the students, characterizing the different degrees of metalinguistic explicitness and the types of argument mobilized in the negotiation of lexical choices.

Metalinguistic activity on the lexicon in children's writing: methodological limits and the search for the process *in statu nascendi*

The specialized literature on the acquisition of the lexicon in children's writing has largely privileged the analysis of the effectiveness of didactic devices and explicit teaching strategies (Grossmann, 2011; Roubaud; Moussu, 2012; Roubaud; Sardier, 2020). By contrast, the investigation of the spontaneous metalinguistic activities that emerge during the writing process itself, particularly in the semantic-lexical domain, remains relatively restricted. Although the existing studies offer relevant contributions, many of them rest on methodological designs which, in seeking access to the child's cognitive processes, end up modifying them, thereby instituting a well-known methodological paradox. This section offers a critical analysis of these approaches, with the aim of delimiting the theoretical-methodological gap that the present article sets out to address.

A first line of investigation resorts to verbal elicitation protocols, in which the researcher assumes a central role in activating and orienting students' verbalizations (Rojo, 1989; Hall, 2005; Dávalos; Ferreiro, 2019; Vieira, 2012; Spinillo, 2015). Prominent among these are retrospection methodologies (Ericsson; Simon, 1993) or clinical interviews conducted after text production, to which the French-language tradition of research with children in the early stages of literacy adds the entretiens métagraphiques (Jaffré, 2003; David, 2003; Mauroux; Morin, 2018). In Spinillo's (2015) study, for example, third- and fifth-grade students were invited to revise texts they had produced previously and to justify the written erasures identified by the researcher. Although such justifications reveal aspects of the linguistic knowledge mobilized by the students, they are verbalizations produced a posteriori, strongly conditioned by the interview situation and by the presence of the adult researcher. The explanations tend to take the form of retrospective rationalizations (Nisbett; Wilson, 1977), frequently anchored in world knowledge or in the recall of instructional directives, and thus removed from the moment of hesitation, negotiation and decision that characterizes the writing process as it unfolds. It should be noted that even the proponents of verbal protocol analysis restrict the validity of retrospective reports to strict conditions, such as the immediate retrieval of information still available in working memory and the absence of directed probing (Ericsson; Simon, 1993), conditions that are hardly met in interviews conducted after the text has been completed. In the case of newly literate students, the limitation is compounded: the discourse genre of the explicitation interview is unfamiliar to them (Vermersch, 2019; Roussel, 2017), and managing the asymmetry between the adult interviewer and the young writer requires specific interviewing techniques (Mauroux; Morin, 2018). Moreover, the exclusive focus on written erasures that have already been consolidated restricts the analysis to the final marks left on the manuscript, effacing the rejected alternatives and the lexical conflicts that never came to be materialized graphically.

In an attempt to move temporally closer to the act of writing, a second methodological approach employs concurrent verbalization protocols, such as adaptations of the think-aloud technique (Reinhart *et al.*, 2022; Newton *et al.*, 2023). In these situations, children are asked to verbalize their thoughts while writing or immediately afterwards, under the explicit guidance of a researcher. Although this technique makes it possible to capture reflections closer to the moment of production, it introduces a decisive enunciative limitation: in speaking to the researcher, the child shifts from her position as an author immersed in the writing process to that of a subject explaining her own activity to an external and hierarchically asymmetrical interlocutor. This reconfiguration of the enunciative framework is no minor side effect: the verbal protocol itself constitutes a discourse genre addressed to an interlocutor, even an absent one (Smagorinsky, 2001), and there is empirical evidence that participants instructed to "talk to themselves" filter their verbalizations in relation to a presumed addressee (Barkaoui, 2011). Consequently, the verbalizations frequently take on a generic metacognitive character concerning the task, at the expense of fine-grained negotiations about specific lexical choices and their effects of meaning in the manuscript under construction. External demand thus shapes the type of reflection produced, privileging global, decontextualized justifications.

A third approach, more directly aligned with the object of the present study, began to emerge in the 1990s. These studies concentrate on the analysis of spontaneous comments produced in collaborative writing situations (Daiute; Dalton, 1993; Camps, 1995; Calil, 1996; Barbeiro, 1999). This represents a significant methodological advance, since metalinguistic activity emerges within the flow of peer interaction itself, without direct elicitation by the researcher. As systematized by Camps (2025), these verbalizations may be manifested either through statements formulated in ordinary language or through statements that mobilize specific grammatical terminology, or "technical terms".

Despite this advance, important instrumental and analytical limitations persist, particularly as regards the identification and description of metalinguistic activities of a semantic-lexical nature. In the studies by Camps and colleagues, such comments appear infrequently and are analyzed on the basis of audio records alone. In privileging oral reformulation, this methodological option loses the material anchoring in the manuscript under construction, making it impossible to establish with precision the relationship between the term commented on, what was actually linearized on the sheet of paper and the enunciative context of the negotiation. As a consequence, the analysis remains at the level of the verbal accompaniment of reformulation: the scale proposed by the authors orders reformulations along four degrees of explicitness, from simple reformulations, with no apparent intermediate step, to those accompanied by reflections employing metalinguistic terms, but it does not discriminate the content of the reflection. This is shown by the episode presented as an example of the highest degree on the scale: in formulating the price of a product, a student proposes replacing alto ['high'] with elevado ['elevated'], justifying the proposal as the use of "scientific words" (Camps *et al.*, 2021, p. 162-163). The classification records the presence of the metaterm but does not make explicit that alto and elevado function, in this context, as near-synonyms:

words very close in meaning but not fully equivalent, since they differ in register of use (more common in the case of alto, more formal or specialized in the case of elevado). Nor does it describe the type of reason mobilized by the student, which is one of register and not of meaning. Applied to this episode, the instrumental limitation also becomes visible: the audio record does not make it possible to determine whether alto was ever linearized; that is, whether the substitution constitutes an oral erasure, prior to inscription, or the correction of a segment already written. The authors themselves, moreover, acknowledge that these data make it impossible to infer the level of the underlying cognitive activity when no explicitation occurs (Camps *et al.*, 2021, p. 164).

It must be emphasized that these verbalizations emerge spontaneously, in the course of the interaction between the students and without researcher intervention, which lends them considerable enunciative value. Disanchored from the concrete writing gesture, however, the comment loses analytical density: the classification describes what was said, but not how, why and under what conditions that saying constitutes itself as metalinguistic activity.

These significant limitations point to the need for a methodological approach capable of articulating, in an integrated way, spontaneous verbalization, the writing gesture and the process of linearization of writing, while preserving the complexity of the manuscript under construction. It is precisely this requirement that guides the proposal adopted in the present study, presented in the following section, in which we mobilize a genetic approach to the school manuscript, anchored in multimodal records, in order to analyze the emergence and functioning of semantic-lexical comments in statu nascendi.

Manuscript under construction, enunciation and unfolded comments: foundations for a genetic analysis of the lexicon

The analysis of school writing in an authentic production situation requires theoretical and methodological devices capable of grasping the text as process, that is, as an ongoing activity marked by hesitations, negotiations and progressive choices. Genetic approaches to the manuscript provide access to this processual dimension by shifting the focus from the finished product to the movements that lead to its constitution.

The Ramos System belongs to this framework, articulating audiovisual records of the verbal interactions between the students with the synchronized capture of the graphic trace of the manuscript by means of digital pens. This device makes it possible to correlate, with temporal precision, what is said, what is written and the moment at which particular linguistic units are actually linearized on the sheet of paper. By preserving the ecology of the classroom and minimizing researcher intervention, the methodology favors the emergence of spontaneous verbalizations produced between peers.

From a theoretical standpoint, this approach is anchored in the convergence between genetic criticism and Benveniste's theory of enunciation. Both conceive of writing and

speaking as event: in genetic criticism, as an object in continuous transformation; in the theory of enunciation, as a singular act in which the subject constitutes himself as speaker by appropriating language. From this perspective, the manuscript under construction is not reducible to a material support but takes shape as an enunciative space in which the writer positions himself in relation to the other, to the manuscript and to the linguistic possibilities that present themselves (Fenoglio, 2014). Enunciation thus becomes the theoretical axis that articulates the linguistic plane of lexical choices with the discursive plane of the negotiation of meanings.

At this point it becomes necessary to make explicit how the terms writing, speech and orality are understood in this study, since they constitute central concepts of its theoretical and methodological construction. We assume, with Benveniste (1974; 2014), that there are not two languages, one spoken and one written, but one and the same language realized in two modes of enunciation: spoken enunciation and written enunciation. Writing is therefore conceived neither as an autonomous code nor as a mere transcription of the spoken, but as the process by which language is semiotized in another order of manifestation (Benveniste, 2014). Within this framework, we reserve the term speech for the singular, situated oral enunciative acts actually produced by the students during the task; orality for the enunciative modality in which those acts are realized; and writing for the graphic modality of enunciation, whose process of inscription and linearization we call written textualization. The established expression written language, when used, for instance, in the context of acquisition, refers to the written modality of this single language and not to a distinct linguistic system. Turn at talk, in turn, is used in its technical sense, derived from Conversation Analysis (Sacks, 1992), to designate the unit of sequential organization of talk-in-interaction. From this delimitation follows a distinction between two planes that the methodological device must not conflate: what is synchronized are the records, namely the audiovisual streams and the graphic trace captured by the digital pen; what is articulated, or interwoven, are the enunciations, namely oral enunciation and graphic inscription, which co-determine one another in the course of textualization. Far from presupposing a dichotomy between speech and writing, the analysis rests precisely on their imbrication (Marcuschi, 2010): it is because the same language is at work in both enunciative modes that the students' oral comments can bear, at the very instant of linearization, on what is being written.

In the course of the writing process, students are confronted with problems that cut across different levels of language: phonographic, orthographic, punctuational, morphosyntactic, semantic-lexical and discursive. These levels may become interwoven within a single unit in the process of inscription, as demonstrated by collaborative writing processes in which discussion of a word's meaning and discussion of its spelling occur successively within the same episode (Lima; Fernandes; Calil, 2026). In the semantic-lexical domain, these problems frequently take the form of competing choices between possible words or expressions, accompanied by verbalizations that seek to explain, justify or describe why a given term should or should not be used. Such verbalizations may constitute explicit manifestations of metalinguistic activity,

in which the subject momentarily interrupts the flow of writing in order to reflect on the functioning of language and on the effects of meaning of the choices at stake, even when these spontaneous utterances leave no written marks of erasure (Calil, 2017).

The concept of deferred recognition[1] proposed by Calil and Myhill (2020) makes it possible to describe this specific moment at which a potential textual object is orally identified as problematic. This is a recognition that is not confined to the identification of a linguistic unit but institutes a relation of estrangement with the manuscript under construction, opening a point of tension in the syntagmatic chain. Within this space, different possibilities of textual objects come to compete with one another, even though only one will actually be inscribed in the manuscript. Even in the absence of a visible written erasure, deferred recognition makes audible what could have been written, constituting what Calil (2017) terms oral erasure. When competing textual objects are distinguished by their meaning, this erasure operates within the semantic domain (Calil; Felipeto, 2014).

It is within this framework that unfolded comments are situated (Calil, 2016), understood as verbalizations that expand the initial recognition of a deferred textual object through the elaboration of arguments. Unlike statements that merely accept or reject a proposed formulation, unfolded comments make visible the procedures by which students construct and legitimize their linguistic choices in the course of writing. As already indicated, deferred textual objects may bear on different units of language; in this study, however, the analysis concentrates on semantic-lexical units, with the aim of describing the types of argument mobilized and their effects on the process of textualization. These arguments may be of various kinds: semantic (relating to the meaning of words), discursive (appropriateness to the genre or to the narrative context), pragmatic (effects projected onto the reader or onto the communicative situation), school-normative (reference to models legitimized within the school setting) or subjective (personal preferences, beliefs or evaluations). Considered in their enunciative dimension, such comments also make it possible to observe different modes of the subject's implication in the activity of writing, which will subsequently be described analytically in terms of the relations that the students establish with their own formulation and with the enunciation of the other.[2]

---

[1] The term deferred recognition is rendered in Portuguese as "reconhecimento diferido". The notion derives from Calil and Myhill (2020), where the recognition of a textual object figures among the competences involved in revision: the student identifies a textual object, recognizes that it could or should differ from what was written, and modifies it. Here the term names an operation on the plane of orality whose defining trait is recursivity: while dictating, the writer (or the interlocutor within the dyad) returns to a term already uttered. It is this return, rather than the naming or the use of the term, that defers recognition: in returning, the utterance suspends the term from the syntagmatic chain and constitutes it as a possibility of non-inscription, converting an already-given syntagmatic position into a still-unresolved paradigmatic question. This suspension opens a space of competition that is at first merely virtual: no rival term need be posited for deferral to occur. Because it operates in orality rather than in the graphic flow, deferred recognition is not indexed to inscription: it may precede, accompany, or follow the moment at which the textual object is inscribed and linearized. For the full characterization of the notion and of its relation to the types of erasure with which it operates, the reader is referred to that study.

[2] When the comment bears on the speaker's own formulation, it mobilizes an operation that Authier-Revuz (1995) describes as autonymic modalization: in the thread of discourse, the subject momentarily interrupts the saying in order

The identification of these different types of argument makes it possible to understand metalinguistic activity not as a homogeneous phenomenon but as a set of situated discursive operations that articulate linguistic knowledge, school experience and the enunciative positioning of the subject. When analyzed longitudinally (that is, across different text production situations involving the same pair of students), unfolded comments offer relevant evidence about the recurrence, stability or transformation of the criteria mobilized in lexical choices.

Corpus and procedures for the analysis of semantic-lexical comments

Research design

Our investigation is characterized as a qualitative, longitudinal case study. The case study design (Yin, 2001) was chosen because it allows for an in-depth, contextualized investigation of metalinguistic activity during paired text production situations in the real environment of the classroom. Its longitudinal character consists in following the same dyad of students over a period of three months, which makes it possible to observe developmental processes and the recurrence of interactional patterns.[3]

Research context

The data were collected in a second-year class of elementary school at a private school in the city of Maceió, located in the central district of the city and serving families of middle socioeconomic status (class B). The school's pedagogical and curricular project followed the guidelines of the Brazilian curriculum documents (Brasil, 1997), adopting didactic practices based on socio-constructivism and on project-based pedagogy (Hernandez, 2000), combined with the use of textbooks for the various curricular subjects.

---

to return to one of his own words, which thereby comes to be simultaneously used and mentioned. Such reflexive returns manifest, according to the author, the non-coincidences of saying — in particular the non-coincidence between words and things, when the chosen form is felt to be inadequate, insufficient or not self-evident for designating what it refers to. As discussed in Felipeto (2005) and Calil (2006), this apparatus offers a productive framework for the analysis of the verbalizations that accompany the recognition of semantic-lexical textual objects, insofar as these frequently make explicit, in the very course of textualization, the relation between a lexical unit and its meaning. It is not our aim, however, to develop here the fruitfulness of this articulation, which would require separate treatment; we confine ourselves to signaling it as the theoretical horizon of those comments which, in the analysis, bear on the writer's own formulation.

[3] The corpus used in this study was drawn from the ECr2012 Dossier. It belongs to the database Práticas de Textualização na Escola, held at the School Manuscript Lab, based at the Federal University of Alagoas. This material was collected with the support of CNPq (process 305312/2011-1) and authorization for its collection was approved by the Research Ethics Committee (CAAE 17963513.6.0000.5013).

In the teaching of the mother tongue, the school takes the text, in its diversity of genres, as the central unit of teaching and learning, linking reading and writing practices to communicative purposes and to contexts of social circulation.

Participants

Owing to the complexity of this type of filmed record, only two pairs were filmed. Our investigation focuses on one of these dyads of students, identified here by the pseudonyms Ca (7 years and 4 months) and Ig (6 years and 10 months). The grouping of all the students was proposed by the teacher, who took into account the pairings that had formed spontaneously and the relations of empathy among them. The decision to focus this study on a single dyad is consistent with the analytical depth sought in a case study. The participating students were selected by the teacher, who took into consideration the fact that they were friends and engaged productively in collaborative work.[4]

Data generation and collection procedures

The data were collected during the implementation of a reading and writing didactic project entitled "Contos do como e do por quê" ['Tales of How and Why'] (Calil, 2012), whose aim was to work with the textual genre of "etiological tales". In the course of this project, the teacher and/or the students read approximately 30 etiological tales. The teacher also proposed 10 collaborative text production tasks in which her students, working in pairs, invented an etiological tale. The tasks alternated between "writing an etiological tale on a free topic" (students could choose freely what to write about) and "writing an etiological tale on a suggested topic" (the teacher presented a topic on the basis of which the students were to invent the tale, for example, "por que o cachorro faz au-au e o gato miau-miau" ['why the dog goes bow-wow and the cat goes meow'], "por que as bruxas são más" ['why witches are wicked'] or "como surgiram as palavras" ['how words came about']). At the end of the didactic project, the students selected some of the tales, compiled them and published them in book form.[5][6]

Given the limits of this study, which is ethnographic, qualitative and microgenetic in character, we adopted convenience criteria in our methodological choices, video-recording the tasks carried out by the focal dyad (Ca and Ig) and following the protocol

[4] These criteria were determined by the researchers and form part of the procedures for collaborative work.

[5] The didactic project lasted three months, from April to June 2012, as set out in the didactic project Contos do como e do por que (Calil, 2012).

[6] Alongside the development of this didactic project, the school was also running the project "Escola de Leitores". This project was carried out with all classes from Year 1 to Year 5. Each classroom had a small library of about 100 children's literature books pre-selected by the class teacher. The second-year teacher held a discussion circle once a week. The students chose a book and took it home. They filled in a short form about what they thought of the book and commented on it in the discussion circle.

proposed by the Ramos System. The procedure for each task therefore followed a systematic protocol led by the teacher, based on the suggestion made by the researchers:[7][8]

1. Presentation: The students are already organized in pairs. The teacher presents the text production task, asking each pair to plan and orally create an invented etiological tale, with or without a prompt topic.
2. Joint creation: The pairs talk to one another in order to determine what the invented tale will be like and to plan the plot, the characters and the structure of the narrative.
3. Inscription and linearization: After the collaborative oral production, the teacher hands each pair a sheet of paper and writing materials (a black pen), indicating which student is responsible for linearizing the writing and which for dictating what is to be written. This interactional phase is characterized by written textualization and constitutes the main locus for the occurrence of the semantic-lexical comments that we shall analyze.
4. Reading and revision: after finishing the writing of the invented etiological tale, the students could read what they had written, checking whether it accorded with what had been agreed. At this point there may or may not be direct intervention by the teacher.

After data collection, the material was prepared for analysis on the basis of the assumptions of ethnographically grounded Conversation Analysis (Garfinkel, 1967; Sacks, 1992), adapted to the objectives of our study. The methodological process involved the synchronization of the data collected in the classroom by means of different media (camera, audio recorder and smart pen), generating a "synchronized film" that served as the basis for the full transcription of the dialogue between the students, including verbal, prosodic (such as pauses and intonation), interactional (such as overlapping turns) and relevant multimodal elements (gestures, facial expressions, erasures and pen position); on the basis of this transcription, we identified the dialogal texts through the points of tension marked orally in relation to the recognition of lexical units during the inscription and linearization stage, which were then segmented and analyzed qualitatively, with a focus on those that gave rise to unfolded semantic-lexical comments; finally, an interpretive analysis was carried out of the argumentative value of these comments, seeking to understand how justifications, explanations, examples and descriptions influenced the pair's textual decisions and the final version of the narrative.

[7] In order to preserve the interactional ecology of the class, the researchers left the classroom once the equipment had been installed.

[8] The material collected with the participation of this pair also served as the basis for the studies by Santos, Calil, 2018, Calil, Amorim, Lira, 2015, Lira, Calil, 2017.

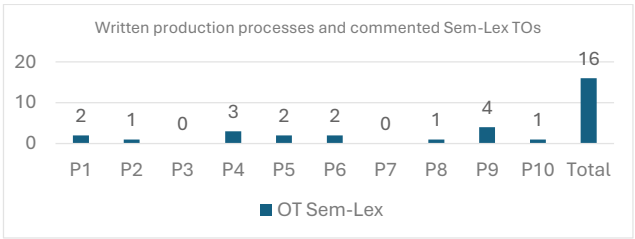
Mapping semantic-lexical recognitions and comments

The analysis that follows focuses on the moments of deferred recognition, in which the students orally identify a potential textual object and treat it as a problem, instituting zones of tension in the syntagmatic chain. It is at these points that unfolded comments are produced: spontaneous verbalizations in which the students mobilize arguments to justify, explain or describe their lexical choices. The focus therefore falls not only on the terms in dispute but on the types of argument brought into play and their recurrence across the production processes.

Semantic-lexical textual objects (Sem-Lex TOs) recognized per process

In the results of the analyses of the 10 synchronized films that recorded the processes of producing the invented etiological tales, it is possible to identify the recognition, by the pair of students, of 16 semantic-lexical textual objects, distributed across eight processes.

Graph 1 – Recognitions of Sem-Lex TOs accompanied by unfolded comments



Source: Prepared by the authors.

In only two processes (P3 and P7) was there a complete absence of Sem-Lex TO recognitions with unfolded comments. The highest concentrations occur in P4 and P9 (3 and 4 recognitions respectively), while P1, P5 and P6 each show 2 recognitions. The processes with a single, isolated recognition (P2, P8, P10) appear to correspond to moments in which attention to the lexicon was punctual and instrumental. This distribution indicates that attention to lexical units accompanied by comments is intermittent and situated: the students do not maintain continuous metalinguistic control throughout the entire course of production but turn to lexical questions at particular moments, probably when the narrative sequence and content call for a lexical decision (the choice of terms for characterization, the reformulation of stretches of text, adjustments of temporal or causal coherence, and so on).

The absence of recognitions in processes 3 and 7 warrants an interpretive note: this is not necessarily a matter of a loss of lexical competence, but of the absence of

discursive triggers calling for reflection on the lexicon at those moments — for example, phases in which the pair advanced the linearization without recognizing problematic terms, or in which the focus shifted to overall organization, narrative transitions or other domains of writing (spelling, punctuation, among others).

Description of the Sem-Lex TOs recognized and of the unfolded comments

Chart 1 summarizes, for each Sem-Lex TO recognized, who recognized it, how that recognition took place and which comments it gave rise to.

Chart 1 – TOs recognized and unfolded comments by Ca and Ig

EE[9]	•[10]O homem queria lugares e mais seres vivos *[The man wanted places and more living beings]*
01P1[11]	01P1C1Ca*[12]: (Saying aloud what he is about to write) O homem**... porque 'o homem' quer dizer 'o homem e a mulher'.**[13] *[The man… because 'the man' means 'the man and the woman'.]*
	01P1C2Ca*: Mas só que... **falando assim, em todos...** fala... 'o homem'. *[But it's just that… speaking like that, about everyone… you say… 'the man'.]*
	01P1C3Ig: É. **Antigamente falava assim mesmo.** *[Yeah. That's just how people talked in the old days.]*
EE	Os animais estavam morrendo / porque as pessoas eles • maltratavam *[The animals were dying / because people mistreated them]*

[9] EE corresponds to the Written Statement (Enunciado Escrito) linearized on the sheet of paper, with the lexical unit inscribed on the sheet, related to the recognized TO, highlighted in a green block.

[10] The sign • indicates the moment at which the point of tension arose in relation to what is being linearized. In this case, for example, the point of tension (the moment at which the lexical unit was recognized) occurred before Ca (the writing student) linearized the beginning of the sentence "O homem...".

[11] The code 01P1 refers to the number of the semantic-lexical object identified (01) in the process analyzed (P1).

[12] In the code 01P1C1 the number of the comment has been added. Thus, the superscript number accompanying the student's initials gives the number of the TO (01), the number of the process (P1) and the number of the comment relating to that TO (C1). The asterisk (*) indicates which student is the writing student in that process. The yellow block over the student's name indicates that it was he who identified the TO.

[13] The recognized TO is identified by the use of red font. This TO may have been recognized on the basis of the previous turn or within the same turn of one of the students; it may also be recognized on the basis of what was written or of what is anticipated in relation to what is to be written. In blue font we have the comment, with the explicit argument highlighted in bold blue. In this first occurrence we have 'O homem' in red, meaning that this phrase was recognized as a problem even before being linearized, and in bold blue the continuation of his comment, "porque 'o homem' quer dizer 'o homem e a mulher'", indicating the verbalized argument and characterizing the statement as an "unfolded comment".

02P1	Ig: Eles não sabiam, mas eles jogavam o lixo... do outro lado sem saber o que era. *[They didn't know, but they threw the rubbish… over the other side without knowing what it was.]*
	02P1C1Ca*: (Saying aloud what he is about to write) Eles... eles os maltratavam, né? **Porque jogar lixo... faziam um montão de coisas ruins...** *[They… they mistreated them, right? Because throwing rubbish… they did a whole load of bad things…]*
	02P1C2Ca*: **Aí tem uma palavra que junta tudo isso, que é 'maltratavam'** *[Then there's a word that puts all that together, which is 'mistreated'.]*
EE	Numa casa vivia um gato / e um cachorro. *[In a house there lived a cat / and a dog.]*
03P2	Ig*: (At the end of the linearization of the sentence, interrupting his inscription and addressing the teacher.) Pode colocar cão... ao invés de cachorro? *[Can I put cão [hound] instead of cachorro [dog]?]* Teacher: (Answering, but addressing the whole class) Olhe, não é para colocar 'cão', tá! A palavrinha que a tia pôs no quadro foi 'cachorro'. *[Look, you're not to put 'cão', all right! The word Miss put on the board was 'cachorro'.]*
	03P2C1Ig*: (Speaking to the teacher) Mas 'cachorro' **é a mesma coisa de 'cão'**! *[But 'cachorro' is the same thing as 'cão'!]*
EE	'Há vários anos atrás não existia arco-íris. *[Several years ago there were no rainbows.]*
04P4	Ca: (Dictating) Há muito tempo atrás... *[A long time ago…]* Ig*: (Objecting to the expression uttered by Ca) Há um século... não é melhor[14]? *[A century ago… wouldn't that be better?]*
	04P4C1Ca: Não. 'Há muito tempo' é mais que 'um século'. Há muito, muito, muito, muito tempo atrás... **é mais que 'um século'.** *[No. 'A long time ago' is more than 'a century'. A very, very, very, very long time ago… is more than 'a century'.]*
	04P4C2Ig*: 'Há vários'... **é mais do que 'um século'**. 'Há vários anos atrás' **é mais do que 'um século'.** *['Several'… is more than 'a century'. 'Several years ago' is more than 'a century'.*
	04P4C3Ig*: **E mais do que 'muitos anos atrás'**. Sabia? [...] *And more than 'many years ago'. Did you know that? […]]*
	04P4C4Ig*: 'Muito tempo, muito, muito... muito, muito... **representa 'há vários anos atrás'**. *['A long time, very, very… very, very…' stands for 'several years ago'.]*

[14] This form of recognition may occur through the repetition of a term or through the proposal of another lexical unit. In some occurrences, the recognition of the textual object may be accompanied by brief evaluative expressions such as "não concordo" ['I don't agree'], "fica ruim" ['it sounds bad'] or "não gosto" ['I don't like it'], among others. These statements should not be understood as argumentative comments but as immediate evaluative manifestations of the subject in the face of the formulation under way. Although every enunciation is, strictly speaking, subjective — insofar as it implies a situated speaker — the subjective character referred to here concerns the fact that such comments do not make explicit, within the enunciation itself, a linguistic reasoning that would sustain the evaluation formulated. In these cases, the saying does not unfold into justifications, explanations, examples or comparisons that would make the textual object a theme of shareable reflection. For this reason, such occurrences are classified as simple comments,

EE	Há vários anos atrás não existia arco-íris. *[Several years ago there were no rainbows.]*
05P4	Ca: (Reading 'Há vários anos') Ah, 'anos'! Eu pensei que você ia colocar há séculos. *[Ah, 'years'! I thought you were going to put 'centuries ago'.]*
	05P4C1Ig*: 'Há vários anos' **é mais do que 'um século'**. *['Several years ago' is more than 'a century'.]*
	05P4C2Ig*: **Se eu colocar 'infinito' na palavra não tem ninguém que vença**. 'Há vários infinitos anos atrás'. 'Há infinitos muitos'. 'Há muitos infinitos anos'. 'Há infinitos séculos atrás'. **Não tem ninguém que vença.** *[If I put 'infinite' in the word, nobody can beat it. 'Several infinite years ago'. 'Infinite many ago'. 'Many infinite years'. 'Infinite centuries ago'. Nobody can beat it.]*
	05P4C3Ig*: 'Infinito' **é a mesma coisa de infinito e de infinito.** *['Infinite' is the same thing as infinite and as infinite.]*
EE	Uma noite uma estrela cadente caiu do céu. *[One night a shooting star fell from the sky.]*
06P4	06P4C1Ca: Uma estrela cadente caiu do céu. 'Caiu do céu' **porque ela é 'cadente'**. *[A shooting star fell from the sky. 'Fell from the sky' because it is 'cadente' [falling].]*
	06P4C2Ig*: É. 'Estrela cadente' **sempre cai.** *[Yeah. A 'shooting star' always falls.]*
EE	O vulcão não espirrava fogo. *[The volcano did not sneeze out fire.]*
07P5	Ig: Espirrava era? *[Sneezed out, was it?]*
	07P5C1Ca*: (Interrupting the inscription of 'espirrava') Espirrava o fogo. **Tipo assim, ó! O vulcão não explode fogo? Então, espirrava. A gente num espirra o catarro? Imagine se a lava fosse o catarro! O vulcão não explode? O nosso nariz também explode. Mas só que não o fogo. O catarro.** *[It sneezed out the fire. Like this, look! Doesn't the volcano explode fire? So, it sneezed. Don't we sneeze out snot? Imagine if the lava were the snot! Doesn't the volcano explode? Our nose explodes too. Only not fire. Snot.]*
EE	O dragão foi morar em um vulcão gigante. *[The dragon went to live in a giant volcano.]*
08P5	Ig: (Rereading) Um dragão... foi morar em um vulcão... (Dictating) ...gigante [...] mas... o 'vulcão' que você falou num era 'grande'? Falou 'grande', mas é melhor 'gigante'. *[A dragon… went to live in a volcano… …giant […] but… wasn't the 'volcano' you said 'big'? You said 'big', but 'giant' is better.]*
	08P5C1Ca*: 'Gigante' **é o maior [que 'grande']**... o 'vulcão gigante'. *['Giant' is the biggest [compared with 'big']… the 'giant volcano'.]*
EE	Ele começou a soltar fogo pela boca e pelo nariz, quando está muito bravo. *[He began to breathe fire from his mouth and his nose, when he is very angry.]*

as opposed to unfolded comments, in which the initial evaluation is accompanied by the explicitation of arguments guiding the maintenance, substitution or rejection of the recognized textual object. Simple comments occurred in all the processes, including P3 and P7, but, being "simple", they were not included in Graph 1.

09P6	Ca: (Dictating) Quando está furioso. *[When he is furious.]* Ig*: Quando está bravo, não é melhor? [...] Não. Eu não sei como é a palavra. Eu não sei qual é o significado da palavra 'furioso'. *[When he is angry, wouldn't that be better? […] No. I don't know what the word is like. I don't know what the meaning of the word 'furious' is.]*
	09P6C1Ca: ['furioso' significa] **Raiva. Tá com muita raiva.** *[['furioso' means] Anger. He's very angry.]*
	09P6C2Ca: **Quando [alguém] está muito bravo**. *[When [someone] is very angry.]*
EE	~~Nes~~ • Desde esse dia todos os dragões soltavam fogo. *[Nes • From that day on all dragons breathed fire.]*
10P6	Ca: (Dictating) Desde esse dia... Desde esse dia... Desde... *[From that day on… From that day on… From…]* Ig*: (Objecting to the suggestion 'Desde esse dia') Nesse... [Nes] *[On that…]*
	10P6C1Ca: Ig, 'Desde esse dia' **porque 'desde esse dia' são os outros dias.** *[Ig, 'from that day on' because 'from that day on' means the other days.]*
	10P6C2Ig*: E 'nesse dia' **só é esse dia mesmo.** *[And 'on that day' is only that very day.]*
EE	Há • muito [anos] atrás, o leão não tinha juba. *[A long [years] ago, the lion had no mane.]*
11P8	Ca: (Dictating) Há muito tempo atrás... *[A long time ago…]* Ig*: (Writing [A] for 'Há') Há vários séculos? *[Several centuries ago?]* Ca: Não. ´Há muito tempo atrás´. 'Há vários séculos' fica ruim. *[No. 'A long time ago'. 'Several centuries ago' sounds bad.]*
	11P8C1Ig*: Séculos... **é cem.** *[Centuries… is a hundred.]*
	11P8C2Ca: (Speaking with an expression of annoyance, irony) 'Há séculos tempo atrás'. 'Há séculos o leão não tem juba'? Não. 'Há muito tempo atrás' **é muito 'séculos'.** *['Centuries time ago'. 'For centuries the lion has had no mane'? No. 'A long time ago' is a lot of 'centuries'.]*
EE	E ouviu dizer que tinha um mago na • montanha mais alta. *[And he heard that there was a wizard on the highest mountain.]*
12P9	Ca*: (Stopping linearizing) Colina ou montanha? *[Hill or mountain?]*
	12P9C1Ig: 'Montanha' **é mais alta [que 'colina']** *['Mountain' is higher [than 'hill']]*
EE	Na floresta, ele encontrou tigres, cobras, aranhas, morcegos e • escorpiões. *[In the forest, he came across tigers, snakes, spiders, bats and scorpions.]*
13P9	Ca*: ...encontrou tigres, cobras, aranhas, morcegos e escorpiões. *[…came across tigers, snakes, spiders, bats and scorpions.]* Ig: Hum? Ca*: Escorpiões. Escorpiões, sabe? *[Scorpions. Scorpions, you know?]*
	13P9C1Ig: Eu sei. **Aquele bichinho pequeno.** *[I know. That little creature.]*
	13P9C2Ca*: (Referring to the part of the scorpion's body where the sting is) **De curva.** *[Curved.]*

EE	No rio, ele encontrou co·bras d'água. *[In the river, he came across water snakes.]*
14P9	Ca*: ...cobras (writing [co])... cobras d'água. *[…snakes (writing [co])… water snakes.]* Ig: (objecting to the term 'cobras d'água') 'cobras aquáticas'. *['aquatic snakes'.]*
	^{14P9C1}Ca*: 'Cobras d'água' **é a mesma coisa que 'cobras aquáticas'**. *['Water snakes' is the same thing as 'aquatic snakes'.]*
	^{14P9C2}Ig: **Aí vão pensar que ela é feita de água. Aí vão pensar que a cobra é feita de água.** *[Then they'll think it's made of water. Then they'll think the snake is made of water.]*
EE	No rio, ele encontrou cobras d'água, croco·dilos. *[In the river, he came across water snakes, crocodiles.]*
15P9	Ig: (Dictating) Crocodilos, jacarés... *[Crocodiles, alligators…]*
	^{15P9C1}Ca*: Eu acho melhor 'crocodilos'... **que é mais ferozes.** *[I think 'crocodiles' is better… which are more ferocious.]*
EE	Um dia os peixes encontraram ·um grupo de gaviões. *[One day the fish came across a group of hawks.]*
16P10	^{16P10C1}Ca: (Dictating.) Um... um grupo... (Ig writing [um]) (Ca with emphasis) '**Um grupo' é muitos! Entendeu?** *[A… a group… 'A group' is many! Got it?]* Ig*: (Agreeing) Isso. *[That's it.]*

Source: Prepared by the authors.

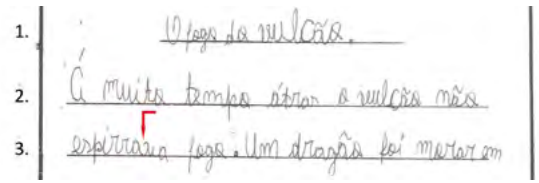
Finally, it is worth noting that only EE10P6 presents an explicit mark of written erasure in the manuscript, indicating a metalinguistic activity carried out through a material return to a textual object previously inscribed ("Nes[se]"). In the remaining episodes analyzed, no written marks of erasure are observed, which, in an approach restricted to the product (the finished school manuscript), might lead to the erroneous conclusion that there was no (meta)linguistic activity at those points of the written manuscript. In the light of the multimodal data generated by the Ramos System, however, these episodes can be characterized as occurrences of oral erasure accompanied by unfolded comments. They are situations in which the recognition of a textual object and the associated metalinguistic activity are realized exclusively on the plane of orality, leaving no traces on the graphic surface of the manuscript.[15]

This functioning becomes particularly visible in the fragment of the school manuscript entitled O fogo do vulcão ['The Volcano's Fire'], at the moment of linearization of the term "espirrava" ['sneezed out'] (EE07P5). With the support of the synchronized film, it can be observed that, as Ca dictates and linearizes the word, Ig recognizes the term as a problematic textual object, instituting a point of tension in

[15] In research based exclusively on the final manuscript (Abaurre, 1997; Capistrano, 2013), the attribution of meanings to written erasures must be treated with caution, since such interpretations tend to remain on the inferential plane: without process data, there is no sufficient empirical basis for reconstructing the operations that gave rise to the graphic mark.

the writing process. This recognition prompts an unfolded comment from Ca, in which he justifies the choice of "espirrava" as against "explodia" ['exploded'], mobilizing arguments of a semantic-descriptive and imaginative nature. This entire movement (recognition, evaluation and justification) takes place in the course of writing, in a situated and interactional manner, without the finished school manuscript presenting any graphic trace of this metalinguistic activity.

Figure 1 – Fragment of the manuscript "O fogo do vulcão": espirrava



Source: School Manuscript Lab.

This example shows that the absence of a written erasure does not imply the absence of reflection on language. On the contrary, oral erasure proves to be a central device for the identification of metalinguistic activities in collaborative writing situations, giving access to competing textual objects and to lexical selection operations that remain invisible when the analysis is confined to the product. The synchronized recording of speech and graphic trace thus constitutes a decisive methodological condition for understanding the genesis of the school manuscript and the ways in which students think about language in the very act of writing.

Lexical recognitions and unfolded comments: manifestations of metalinguistic activity

In this section we examine how Ca and Ig mobilize semantic-lexical reflection during collaborative writing, considering when and how the recognitions occur, as well as the metalinguistic operations they trigger: at times proactive, when the student detects and comments on the lexical problem in the very act of formulation; at times reactive, when the recognition emerges in response to the partner's linguistic usage.

Sem-Lex TOs recognized and moments of recognition

As we have shown, 16 semantic-lexical textual objects (Sem-Lex TOs) associated with unfolded comments were recognized across the ten production processes. These objects encompass a wide variety of lexical units, from terms common to the children's

repertoire, such as “o homem” [‘the man’], “cachorro” [‘dog’], “gigante” [‘giant’] and “montanha” [‘mountain’], to denser or compound expressions such as “há vários anos atrás” [‘several years ago’], “estrela cadente” [‘shooting star’], “cobras d’água” [‘water snakes’] and “um grupo” [‘a group’].

The nature of these TOs indicates that recognition is not restricted to unknown or difficult words but to moments at which lexical usage interferes with the meaning of the manuscript under construction. In several cases, the point of tension (marked by the sign [•] in the summary chart) emerges before or during the linearization of the word, suggesting that recognition occurs in the act of production, when the student has to decide how to formulate a narrative content linguistically. The recognition of Sem-Lex TOs is thus less a gesture of decoding or an act of textual revision after the story has been completed than a movement of semantic decision in the course of linearization, oriented towards the coherence and expressiveness of the manuscript from the point of view of the one who is writing/dictating it.

Who recognizes, how they recognize and when they recognize

A reading of Chart 1 reveals a relatively balanced distribution between Ca and Ig in the role of recognizer. Ig recognized 9 (03P2, 04P4, 07P5, 08P5, 09P6, 10P6, 11P8, 13P9, 14P9); Ca recognized 7 TOs (01P1, 02P1, 05P4, 06P4, 12P9, 15P9, 16P10). The attribution of the recognition in 13P9 requires a note: it is accomplished through an open-class repair initiator (Drew, 1997), Ig’s “Hum?”, a format that signals a problem without locating its source; the data do not make it possible to decide whether the student did not understand the word “escorpiões” [‘scorpions’], did not hear it properly, or was merely surprised by his partner’s choice. Its semantic-lexical nature is, however, interactionally ratified in what follows: Ca treats Ig’s interrogative enunciation as a problem of lexical knowledge (“Escorpiões, sabe?” [‘Scorpions, you know?’]) and Ig responds by displaying understanding and describing the referent. The classification therefore rests on the sequential treatment given by the participants themselves, and not on an inference about the state that gave rise to the interjection.

There are nonetheless differences in the way recognition is carried out. Ca tends to recognize the TO while dictating or writing the lexical unit, signaling simultaneous attention to formulation and evaluation. His recognition is frequently anticipatory, occurring even before the expression has been fully written, as in “o homem” (01P1), “maltratavam” [‘mistreated’] (02P1), “estrela cadente” (06P4) or “um grupo” (16P10). This reveals a stance of internal monitoring of language, focused on semantic appropriateness and lexical precision during the writing process itself.

Among Ca’s proactive recognitions, four (01P1, 02P1, 06P4, 16P10) display a specific enunciative configuration that is worth naming. In these cases, the student does not merely recognize the textual object that he himself formulates for inscription; he glosses it in the same movement, making explicit what the chosen form means (“porque

‘o homem’ quer dizer ‘o homem e a mulher’” [‘because “the man” means “the man and the woman”’]; “‘Um grupo’ é muitos” [“‘A group” is many’]). We designate this recognition-comment pair a self-gloss (Calil, 2014), that is, a gloss bearing on the speaker’s own formulation which, unlike substitutive comparisons, operates retentively — not in order to replace the term but in order to keep it, by justifying it. In Authier-Revuz’s (1995) terms, this is an autonymic modalization: in saying “‘o homem’ quer dizer…”, the student simultaneously uses and mentions the word, suspending for an instant the flow of textualization in order to comment on the unit being inscribed within it.[16]

This anchoring in the manuscript under construction marks the difference from analyses based on audio records alone. In the episode from Camps *et al.* (2021) discussed above, the substitution of “alto” by “elevado” is described in terms of the presence of the metaterm, without its being possible to determine whether “alto” was ever linearized; in the comment “Aí tem uma palavra que junta tudo isso, que é ‘maltratavam’” [‘Then there’s a word that puts all that together, which is “mistreated”’] (02P1C2), by contrast, the synchronized film makes it possible to situate the self-gloss at the exact instant at which Ca subsumes, under a single lexical item, the series of actions he had just described (“jogar lixo… faziam um montão de coisas ruins” [‘throwing rubbish… they did a whole load of bad things’]), thereby making visible not only that metalinguistic reflection is taking place, but on what object and at what point of the inscription it bears.

The self-gloss does not constitute a functioning separate from oral erasure; it is its most explicit form. In all four cases there is competition between elements that could occupy the same position in the syntagmatic chain, which is what defines oral erasure, with the particularity that here the competing term does not remain merely hinted at but is named by the gloss itself in the act of being rejected. In 01P1C1, “o homem” enters into competition with “o homem e a mulher”, a form that the gloss enunciates precisely in order to prevent it from being inscribed; in 02P1C2, “maltratavam” replaces the descriptive periphrasis Ca had just formulated (“jogar lixo… um montão de coisas ruins”), condensing into a single lexical item what had been said as a sequence of actions; in 16P10C1, “um grupo” takes the place of “muitos” [‘many’], equally inscribable in that position (“os peixes encontraram muitos gaviões” [‘the fish came across many hawks’]), preferred and justified by the gloss “‘Um grupo’ é muitos”;

[16] The relation between the recognition of a textual object and the comment bearing on it was established in Calil (2016), a study that analyzes the semantic comments produced by a newly literate dyad during real-time writing and defines, as a type of commented oral erasure, the relation between the textual objects recognized and the comments associated with them. It was already shown there that the writer’s comment not only makes explicit what he knows about the recognized term but suggests the relation he establishes between that term and its relevance for the ongoing text, within the points of tension that organize paired writing. The self-gloss, as delimited here, specifies a subset within that field: the one in which recognition and comment come from the same speaker and bear, retentively, on his own formulation as it is about to be inscribed. Studies such as Calil (2016), Calil and Felipeto (2014) and Lima *et al.* (2026), in taking the comment on the lexicon as a route of access to the writer’s thinking about language in the very act of writing, constitute the analytical matrix of which the notion of self-gloss is an outgrowth; their fruitfulness lies in having shifted the analysis from the product to the situated, co-enunciative relation between the recognized word and the text being constructed.

finally, in 06P4C1, “cadente” and “caiu do céu” [‘fell from the sky’] compete for the same descriptive function[17], and the student retains both forms, so that the redundancy inscribed in the manuscript is the graphic trace of a choice by retention where the inscription of a single form was possible. What the self-gloss adds is therefore not the absence of competition but its explicitation: rather than making audible only the rejected alternative, it enunciates the competing term and the reason for keeping it, raising to the plane of saying the associative series (Suenaga, 2005) that the form condenses (cf. Calil, 2014), and thereby making audible the non-coincidence between the chosen word and what it designates (Authier-Revuz, 1995).

It should be stressed that the prefix self- designates the object of the gloss, namely the speaker’s own formulation, and not its addressee, since the utterances remain co-enunciative. In 01P1 and 06P4, Ca’s self-gloss is followed by an equally glossed ratification from Ig, opened by the same “É.” [‘Yeah.’] (“Antigamente falava assim mesmo” [‘That’s just how people talked in the old days’]; “‘Estrela cadente’ sempre cai” [‘A “shooting star” always falls’]). The explicitation of meaning is therefore not a monologue of the writer about language; it constitutes rather a movement that projects and obtains the partner’s agreement, becoming part of the collaborative negotiation of the lexical choice.

Ig, on the other hand, performs recognitions that generally occur after his partner’s enunciation, configuring a reactive and evaluative movement. He recognizes the TO on hearing the term in use, questioning it, recognizing it or proposing alternatives, as when he tried to replace “cachorro” [‘dog’] with “cão” [‘hound’] (03P2), recognized the difference in meaning between writing “nesse dia” [‘on that day’] or “desde esse dia” [‘from that day on’] (10P6), or when he asked about “espirrava?” (07P5), “furioso?” [‘furious’] (09P6), “escorpião?” [‘scorpion’] (13P9). This dynamic suggests that Ig takes part in the writing process with an auditory and comparative attention, oriented more towards the validation of the lexicon than towards autonomous formulation.

In this way, Ca performs a recognition of a proactive nature, characterized by the self-detection of a problem of meaning, of usage or of form in the course of his own formulation, constituting a procedure of linguistic self-monitoring. Ig, for his part, performs a recognition of a reactive nature, understood as a dialogic and responsive operation of comparative attention, triggered by a linguistic form previously enunciated by his partner.

At this point it is worth considering a variable of the device that might be confused with the distinction proposed here: the division of roles between dictating and linearizing. If proactive recognition depended solely on the student who formulates aloud, and

[17] The redundancy at issue is visible in Portuguese but not in English. Cadente is a participial adjective derived from the verb cair (‘to fall’), so that estrela cadente means, literally, ‘falling star’: the property of falling is already predicated of the star by the very adjective that qualifies it. In adding caiu do céu (‘fell from the sky’), Ca therefore inscribes twice, within the same stretch of text, a descriptive feature that the adjective alone would have supplied, and his gloss, “Caiu do céu” porque ela é “cadente”, makes exactly that relation explicit. The usual English equivalent, shooting star, is lexically opaque: it does not contain the verb ‘to fall’, so a reader working from the English rendering of the utterance would not see the redundancy the analysis describes.

reactive recognition on the one who transcribes, the profiles described would correspond to an effect of the organization of the task and not to distinct modes of metalinguistic activity. The data, however, do not support this reduction. Each student's recognitions appear in both roles: Ca recognizes TOs both as the writing student (01P1, 02P1, 12P9, 15P9) and as the dictating student (05P4, 06P4, 16P10); Ig, in turn, recognizes TOs both as the writing student (03P2, 04P4, 09P6, 10P6, 11P8) and as the dictating student (07P5, 08P5, 13P9, 14P9). The proactive/reactive distinction therefore does not refer to the role assumed in the task but to the locus of detection of the problem (one's own formulation or that of the other), a dimension that cuts across the division between dictating and linearizing. In this pair's practice, the writing student himself verbalizes what he is about to inscribe ("saying aloud what he is about to write"), so that both students have, in either role, linguistic material to monitor. A caveat nonetheless remains: since this is a single dyad, it cannot be ruled out that the distribution of roles influences the frequency, though not the existence, of each type of recognition. This point recommends that future research adopt designs including the controlled rotation of the writing and dictating roles.

Distribution of unfolded comments by student

Analyzing the number of unfolded comments produced by each student in relation to the Sem-Lex TOs recognized makes it possible to observe significant differences in the way Ca and Ig engage in lexical reflection during collaborative writing.

Table 1 – Number of unfolded comments by each student per TO recognized

	01 P1	02 P1	03 P2	04 P4	05 P4	06 P4	07 P5	08 P5	09 P6	10 P6	11 P8	12 P9	13 P9	14 P9	15 P9	16 P10	**Total**
Ca	2	2	0	1	0	1	1	1	2	1	1	0	1	1	1	1	**16**
Ig	1	0	1	3	3	1	0	0	0	1	1	1	1	1	0	0	**14**
Total	3	2	1	4	3	2	1	1	2	2	2	1	2	2	1	1	30

Source: Prepared by the authors.

In all, 30 unfolded comments were identified, 16 produced by Ca and 14 by Ig, distributed across the 16 textual objects analyzed. Although both students show involvement in the semantic elaboration of the text, each one's pattern of participation reveals distinct levels of metalinguistic control and continuity.

Categorization and characterization of the unfolded comments

The linguistic-enunciative analysis of the unfolded comments made evident the need for a more refined classification of the forms of metalinguistic argumentation (sequences in which the student, besides signaling a lexical problem, formulates reasons to sustain, contest or negotiate a choice) mobilized by the students during the collaborative writing process. The qualifier *metalinguistic* derives from the object of the argumentation, a unit of language in the process of being inscribed in the manuscript, and not from the nature of the reasons mobilized: these may take as their support language itself, the text under construction or the referent, and the way in which they thematize language is precisely what the continuum of explicitness, presented below, orders. Although all the comments are associated with prior lexical recognitions, they vary widely as regards the type of semantic reasoning, the nature of the relation established between the words and the degree of explicitness of meaning. This diversity reveals that, in identifying problems of lexical usage, the students construct different ways of thinking about the lexicon, sometimes explaining meanings, sometimes comparing terms, describing properties or justifying choices in relation to the narrative context.

We therefore propose four categories (in addition to a limiting case, discussed below) on the basis of which we classified the 30 unfolded comments described in the summary chart. Each category is defined by the type of relation the comment focuses on: the relation of the word to its meaning (lexical definition), to other units of the lexicon (lexical comparison), to the text under construction (contextual-argumentative) or to the referent it designates (description).

Chart 2 – Types of unfolded comment

	Type of comment	Description
1.	Lexical definition	Seeks to make explicit the meaning of the recognized term, by definition, paraphrase or synonymic substitution.
2.	Lexical comparison	Establishes relations between the recognized term and other lexical units: by difference (semantic opposition or gradation) or by similarity (equivalence or closeness of meaning).
3.	Contextual-argumentative	Relates the recognized term to the narrative context, justifying its use in terms of coherence, temporality, causality or textual progression.
4.	Description	Describes physical, behavioral or functional characteristics of the referent designated by the recognized term.

Source: Prepared by the authors.

Chart 3 below reorganizes the 30 unfolded comments identified in the data, grouping them by argumentative category.

Chart 3 – Unfolded comments grouped by category

	TO	Comment	Student
Lexical definition			
1.	01P1C1	Porque ‘o homem’ quer dizer ‘o homem e a mulher’. [‘Because “the man” means “the man and the woman”.’]	Ca*
2.	02P1C2	Aí tem uma palavra que junta tudo isso [fazer um montão de coisas ruins], que é ‘maltratavam’. [‘Then there’s a word that puts all that together [doing a whole load of bad things], which is “mistreated”.’]	Ca*
3.	04P4C4	‘Muito tempo, muito, muito... representa ‘há vários anos atrás’. [‘“A long time, very, very…” stands for “several years ago”.’]	Ig*
4.	06P4C1	‘Caiu do céu’ porque ela é ‘cadente’. [‘“Fell from the sky” because it is “shooting”.’]	Ca
5.	09P6C1	‘Furioso’ significa raiva. [‘“Furious” means anger.’]	Ca
6.	09P6C2	Quando [alguém] está muito bravo. [‘When [someone] is very angry.’]	Ca
7.	11P8C1	‘Séculos’... é cem. [‘“Centuries”… is a hundred.’]	Ig*
8.	16P10C1	‘Um grupo’ é muitos. [‘“A group” is many.’]	Ca
Lexical comparison – by difference			
1.	04P4C1	‘Há muito tempo’ é mais que ‘um século’. [‘“A long time ago” is more than “a century”.’]	Ca
2.	04P4C2	‘Há vários’... é mais do que ‘um século’. [‘“Several”… is more than “a century”.’]	Ig*
3.	04P4C3	E mais do que ‘muitos anos atrás’. [‘And more than “many years ago”.’]	Ig*
4.	05P4C1	‘Há vários anos’ é mais do que ‘um século’. [‘“Several years ago” is more than “a century”.’]	Ig*
5.	05P4C2	‘Infinito’... não tem ninguém que vença... ‘Há infinitos anos atrás’. [‘“Infinite”… nobody can beat it… “Infinite years ago”.’]	Ig*
6.	08P5C1	‘Gigante’ é o maior [que ‘grande’]. [‘“Giant” is the biggest [compared with “big”].’]	Ca*
7.	12P9C1	‘Montanha’ é mais alta que ‘colina’. [‘“Mountain” is higher than “hill”.’]	Ig
Contextual-argumentative			
1.	01P1C2	Mas só que... falando assim, em todos... fala... ‘o homem’. [‘But it’s just that… speaking like that, about everyone… you say… “the man”.’]	Ca*
2.	01P1C3	É. Antigamente falava assim mesmo. [‘Yeah. That’s just how people talked in the old days.’]	Ig

3.	10P6C1	‘Desde esse dia’ porque ‘desde esse dia’ são os outros dias. [‘“From that day on” because “from that day on” means the other days.’]	Ca
4.	10P6C2	E ‘nesse dia’ só é esse dia mesmo. [‘And “on that day” is only that very day.’]	Ig*
5.	11P8C2	‘Há séculos tempo atrás’... não. ‘Há muito tempo atrás’ é muito séculos. [‘“Centuries time ago”… no. “A long time ago” is a lot of centuries.’]	Ca
6.	14P9C2	Aí vão pensar que ela [cobra d´água] é feita de água. [‘Then they’ll think it [water snake] is made of water.’]	Ig*
Description			
1.	02P1C1	[maltratavam] Porque jogar lixo... faziam um montão de coisas ruins... [‘[mistreated] Because throwing rubbish… they did a whole load of bad things…’]	Ca*
2.	06P4C2	‘Estrela cadente’ sempre cai. [‘A “shooting star” always falls.’]	Ig*
3.	13P9C1	[escorpião] Aquele bichinho pequeno. [‘[scorpion] That little creature.’]	Ig
4.	13P9C2	De curva. [‘Curved.’]	Ca*
5.	15P9C1	‘Crocodilos’... que é mais ferozes. [‘“Crocodiles”… which are more ferocious.’]	Ca*
Lexical comparison – by similarity			
1.	03P2C1	Mas ‘cachorro’ é a mesma coisa de ‘cão’! [‘But “cachorro” is the same thing as “cão”!’]	Ig*
2.	07P5C1	O vulcão não explode fogo? Então, espirrava. [...] Imagine se a lava fosse o catarro! [‘Doesn’t the volcano explode fire? So, it sneezed. […] Imagine if the lava were the snot!’]	Ca*
3.	14P9C1	‘Cobras d’água’ é a mesma coisa que ‘cobras aquáticas’. [‘“Water snakes” is the same thing as “aquatic snakes”.’]	Ca*
Limiting case: echoic comment			
1.	05P4C3	‘Infinito’ é a mesma coisa de infinito e de infinito. [‘“Infinite” is the same thing as infinite and as infinite.’]	Ig*

Source: Prepared by the authors.

Comments of the “lexical definition” type, with 8 occurrences, represent the most explicit level of metalinguistic reflection. In them, the students dwell on the construction or reformulation of meanings, mobilizing strategies of explanation, paraphrase and synonymic equivalence. There is a clear attempt to stabilize the meaning of words, as in “‘Furioso’ significa raiva” [‘“Furious” means anger’] or “‘Um grupo’ é muitos” [‘“A group” is many’].

“Lexical comparison” (10 occurrences: 7 by difference and 3 by similarity) likewise indicates metalinguistic activity, but with a different degree of abstraction. In

“comparison by difference” (“‘Gigante’ é o maior que ‘grande’” [‘“Giant” is bigger than “big”’]), the reasoning is relational: the student does not define the term but establishes semantic boundaries between closely related lexical units, which points to an understanding of lexical gradation and contrast. In “comparison by similarity” the movement is the reverse: what is sought is the integration of meanings belonging to the same semantic field, “‘Cachorro’ é a mesma coisa de ‘cão’” [‘“Cachorro” is the same thing as “cão”’]; or to an equivalent semantic field, “O vulcão não explode fogo? Então, espirrava […] Imagine se a lava fosse o catarro!” [‘Doesn’t the volcano explode fire? So, it sneezed […] Imagine if the lava were the snot!’]. In both cases there is reflection on the relations of meaning between words, although the focus lies more on comparison than on conceptual explanation. 07P5C1, however, lies at the boundary of the classification. The justification of “espirrava” constructs an analogy between referential domains (volcano/lava, nose/snot): the similarity may be read either as a correlation between the meanings of “espirrar” [‘to sneeze’] and “explodir” [‘to explode’] or as a contrast, by similarity, of physical phenomenon between the two. But the utterance also allows an interpretation in terms of described scenarios: here there would be a description of the phenomenon serving to justify the use of the word “espirrava”. For cases such as this we adopt a rule of precedence: since argumentativity is constitutive of every unfolded comment, the categories discriminate the type of reason mobilized; when form and focused relation diverge, the relation prevails, identifiable through the scope of the justification, that is, what exactly the reason presented bears upon: the volcano explodes fire/lava, just as the nose sneezes out snot.

“Contextual-argumentative” comments, with 6 occurrences, reveal a more discursive type of activity. In them, the student evaluates the appropriateness of the term to the narrative context, justifying choices in terms of coherence, temporality or textual progression (“‘Desde esse dia’ porque ‘desde esse dia’ são os outros dias” [‘“From that day on” because “from that day on” means the other days’]). This category demonstrates a sensitivity to the textual and pragmatic dimension of the lexicon, going beyond isolated meaning to consider how the word functions within the fictional narrative being written.

“Descriptive” comments, for their part, come closer to the referential plane. The student describes the referent, whether in physical (“Aquele bichinho pequeno” [‘That little creature’]), behavioral (“‘Crocodilos’… que é mais ferozes” [‘“Crocodiles”… which are more ferocious’]) or attitudinal terms (“Porque jogar lixo… faziam um montão de coisas ruins…” [‘Because throwing rubbish… they did a whole load of bad things…’]). Although they contain semantic inferences, they are less metalinguistic, since they concentrate on the represented content and not on the linguistic form itself. One item in this category constitutes a further borderline case that it is worth making explicit, since in it the form of the utterance and the focused relation diverge. In “eu acho melhor ‘crocodilos’… que é mais ferozes” [‘I think “crocodiles” is better… which are more ferocious’] (15P9C1), the form is argumentative and comparative; the reason

mobilized, however, is a property of the referent (the animal's ferocity) and not a relation between the meanings of lexical units, as in "'gigante' é o maior [que 'grande']".

Finally, a single comment ("'Infinito' é a mesma coisa de infinito e de infinito" ['"Infinite" is the same thing as infinite and as infinite']) resists classification within the four categories: although it mimics the form of comparison by similarity ("é a mesma coisa de" ['is the same thing as']), it does not place two distinct lexical units in relation, confining itself to echoing the term itself. Taken within the turn in which it occurs (05P4C2–C3), however, the echo is not discursively empty: it closes the argumentative sequence by reaffirming, through emphatic repetition, the feature of unsurpassability that the student attributes to "infinito"; the very iteration of the word enacts the meaning it designates. The distinction that organizes the continuum is therefore one of explicitness and not of function: the comment presents a zero degree of metalinguistic explicitness, since it predicates nothing of the term beyond the tautology, without this cancelling its argumentative value in the interaction. Rather than elevating it into a category, we treat it as a limiting case, since it marks the zero degree of the continuum of explicitness and, at the same time, delimits by contrast the criterion of lexical comparison, which requires a second term effectively placed in relation.

In sum, the categories are distributed along a continuum of metalinguistic explicitness:

- Greater reflection:

Lexical definition → Lexical comparison → Contextual-argumentative

- Lesser reflection:

Description → echo (zero degree)

This gradient supports a general tendency: 18 of the 30 comments are concentrated in the categories of greater explicitness (lexical definition and lexical comparison), in which language itself is taken as object; adding the contextual-argumentative comments, 24 of the 30 focus on language or on the text, and not solely on the referent.

The self-gloss, described in the previous section, occupies a notable position within the continuum of metalinguistic explicitness proposed above: the four occurrences identified are all classified as lexical definition (01P1, 02P1, 06P4, 16P10), the most explicit category of the gradient, the one in which the word is taken as object and its meaning made explicit. This convergence is not an effect of the definition of the self-gloss, which is established by an enunciative criterion, namely the retentive way in which the subject returns to his own formulation, and which stipulates nothing as to the type of relation the comment focuses on. Since these two axes, the enunciative one and that of the focused relation, are independent, the fact that all four self-glosses fall within a single category, and precisely the most explicit one, calls for explanation. That explanation lies, as the data suggest, in the very structure of the gesture: retaining a form by naming what it means, or the term with which it competes, requires constituting the lexical unit as an object of reflection and making its meaning explicit, an operation that defines the most explicit end of the continuum. The self-gloss therefore does not occupy

just any point on the gradient; it tends to produce its most explicit pole, raising to the plane of saying what, in other recognitions, remains implicit in the lexical choice alone.

Individual differences in semantic-lexical comments during collaborative writing

Broadly speaking, Ca displays a more analytical and systematic metalinguistic functioning, oriented towards conceptual explicitation and the semantic justification of lexical choices. His sixteen unfolded comments are distributed across lexical definition (6 occurrences), comparison (4), description (3) and the contextual-argumentative type (3), a configuration that indicates a continuous effort to control meaning and to monitor textual coherence. More than the frequency of the categories, however, it is the way they are realized that singles this student out. Ca tends to take the word as object and to make its meaning explicit on his own initiative, in the very act of formulating what is to be inscribed: it is from him that the four self-glosses of the corpus originate (01P1, 02P1, 06P4, 16P10), in which recognition and justification come from the same speaker and bear on his own formulation. This disposition to constitute the lexical unit as a theme of reflection, stabilizing its meaning by definition or paraphrase before any objection has been raised, reveals a functioning oriented towards the analysis of language as object, in which the writer momentarily distances himself from the manuscript in order to think about it.

Ig, by contrast, displays a more reactive and exploratory metalinguistic functioning. His fourteen unfolded comments are distributed across lexical comparison (6 occurrences: 5 by difference and 1 by similarity), the contextual-argumentative type (3), lexical definition (2), description (2) and the echoic comment (1), the only one in the corpus, which fixes the zero degree of the continuum of explicitness. The distribution is close to Ca's in several categories, and the difference between the two students therefore lies not in frequency but in the quality of the metalinguistic gesture. Unlike Ca's autonomous definitions, Ig's comments generally emerge in counterpoint to a formulation previously enunciated by his partner: he sizes up the lexical object less by defining its meaning than by confronting it with other terms, as the recurrent comparisons by gradation and equivalence show ("é mais do que" ['is more than'], "é a mesma coisa de" ['is the same thing as'], "é mais alta que" ['is higher than']). His descriptions too, and to a large extent his definitions, unfold within this dialogic and responsive regime, and not as a spontaneous thematization of the lexicon. The echoic comment, in emphatically reiterating "infinito" without placing it in relation to a second, distinct term, functions as the limit of this same comparative dynamic, deriving from it through the emptying out of the contrasting term. What takes shape, then, is a reflection situated in interaction, in which meaning emerges from contrast and from the confrontation of ideas.

Taken together, the two profiles trace a gradient of metalinguistic explicitness between the students, the criterion for which is not the frequency of the categories, similar in both, but each one's mode of implication with the lexicon: in Ca, the explicit

and autonomous thematization of the word, taken as object on his own initiative; in Ig, a situated and responsive reflection that sizes up the term by contrast with what his partner enunciates. These are not hierarchical levels of competence but two ways of constituting the lexicon as an object of reflection, one by definition and the other by confrontation, which complement one another in the joint construction of the manuscript.

Conclusion

This study investigated the manifestations of metalinguistic activity in the semantic-lexical domain during children's collaborative writing, focusing on the deferred recognitions and unfolded comments produced by a dyad of second-year elementary school students. Anchored in textual genetics (Grésillon, 1994; Doquet, 2011) and in the theory of enunciation (Benveniste, 1974), the research proceeded from a critique of approaches centered on the product or on elicited verbalizations, which tend to displace students from their position as authors in the act of writing and to impoverish access to the actual processes of linguistic reflection.

The results show that semantic-lexical reflection emerges in a situated and functional manner, tied to points of tension instituted in the course of the linearization of the text (Calil, 2020). The sixteen textual objects recognized do not necessarily correspond to rare or unknown words but to lexical choices that directly affect the construction of meanings, narrative coherence and the expressiveness of the manuscript. Each recognition institutes a zone of competition between lexical possibilities and makes audible alternatives that could have been inscribed but remain solely on the plane of orality, in the form of oral erasure. That a single episode (EE10P6) presents a written mark of erasure, whereas the remainder are realized exclusively in orality, confirms that metalinguistic activity is inherent in the act of writing (Tolchinsky, 2001) and that an analysis restricted to the product would render it invisible.

The typology of unfolded comments revealed that this activity is not manifested homogeneously: it is distributed along a continuum of metalinguistic explicitness running from lexical definition and lexical comparison, in which language itself is taken as object, to description of the referent, with the echoic comment as its zero degree. Twenty-four of the thirty comments focus on language or on the text, and not solely on the referent, which indicates that thinking about the lexicon, in children's writing, mobilizes a range of discursive operations in which linguistic knowledge, school experience, imagination and enunciative positioning are articulated.

Among these operations, the self-gloss stood out: the gesture whereby the writer, in the act of inscribing a form, names the term with which it competes and makes explicit what it means, retaining it rather than replacing it. Described, in Authier-Revuz's (1995) terms, as an autonymic modalization anchored in the manuscript under construction, the self-gloss constitutes the most explicit form of oral erasure: all four occurrences in the corpus fell within the most explicit category of the continuum, lexical definition.

Their identification was made possible only by the synchronization of speech and graphic trace, which situates the reflection at the exact instant of inscription; precisely what escapes analyses based on audio records alone.

The comparison between the two students revealed distinct and complementary modes of functioning that are not reducible to the frequency of the categories, similar in both, but to their mode of implication with the lexicon. Ca operates through proactive recognition: he takes the word as object and makes its meaning explicit on his own initiative, in the act of formulating what is to be inscribed, and it is from him that the four self-glosses of the corpus originate. Ig operates through reactive recognition: he sizes up the lexical object by confronting it with what his partner enunciates, privileging comparison by gradation and equivalence. Rather than hierarchical profiles of competence, these are two ways of constituting the lexicon as an object of reflection (one by definition, the other by confrontation) which complement one another in the joint construction of the manuscript. It should be noted, as a caveat, that since this is a single dyad it cannot be ruled out that the distribution of the dictating and writing roles influences the frequency of each type of recognition.

From a theoretical and methodological standpoint, the study reinforces the centrality of oral erasure (Calil, 2003), of the unfolded comment (Calil, 2016), and of deferred recognition (Calil; Myhill, 2020) to the analysis of the genesis of the school manuscript, and adds to them the self-gloss as an enunciative operator specific to lexical retention. By making audible what could have been written, the genetic approach shifts metalinguistic activity from a normative perspective to a processual, enunciative and situated conception, within which unfolded comments prove to be not merely indicators of metalinguistic awareness but fundamental discursive operators of collaborative writing.

The implications for the teaching and learning of writing are significant. Reflection on the lexicon emerged spontaneously in a collaborative writing situation that was meaningful to the students, without direct elicitation by the adult, which confers on orality a constitutive role in the elaboration of writing, as a space for the negotiation of meanings. For teaching practice, this entails recognizing the pedagogical value of students' verbalizations, even when they leave no written mark in the final text, and rethinking assessment practices centered exclusively on the finished product.

This is, finally, a qualitative case study, circumscribed to a single dyad and to a specific age range, which restricts the statistical generalization of the results and calls for research that increases the number of participants, encompasses other stages of schooling and articulates the semantic-lexical domain with other levels of writing. These limits do not, however, compromise the heuristic value of the work, whose generalization is analytical in nature — directed at the categories and mechanisms described, potentially transferable to other collaborative writing contexts — and not statistical (Yin, 2001). By situating semantic-lexical recognitions and unfolded comments within the interweaving of the spoken and the written, this research contributes to a finer understanding of how metalinguistic activity emerges, in the act, in children's writing, and offers a basis for

teacher education attentive to the ways in which students think about language in the very act of writing.

Acknowledgments

This study received twofold support from the Brazilian National Council for Scientific and Technological Development (CNPq): a Research Productivity Fellowship (PQ) linked to the project Produção Textual e Atividades Metalinguísticas (MetaWriting II) (Process no. 307248/2023-2), and a research grant awarded under CNPq/MCTI Call no. 10/2023 – Consolidated Groups, relating to the project Produção Textual e Atividades Metalinguísticas – Apoio (Process no. 404398/2023-5). We are likewise grateful to the school and the teacher who welcomed the research into their classroom and, in particular, to the pair of participating students and their families, whose consent and generous collaboration made it possible to record the writing processes analyzed here.

CALIL, Eduardo; FELIPETO, Cristina; LIMA, Maria Hozanete Alves de. Escolha lexical na escrita a dois: rasura oral, comentário desdobrado e atividade metalinguística na textualização infantil. **Alfa**, São Paulo, v. 70, 2026.

- ***RESUMO:*** *Este artigo investiga como a atividade metalinguística no domínio semântico-lexical emerge durante a escrita colaborativa infantil. Considerando que a investigação dos processos espontâneos de reflexão sobre o léxico permanece limitada, restrita a protocolos de eliciação que deslocam o aluno de sua posição de autor em ato, o estudo acompanha uma díade de alunos do 2º ano do Ensino Fundamental ao longo de dez processos de produção de contos etiológicos inventados. A abordagem articula a Genética Textual e a teoria da enunciação de Benveniste, com base no protocolo do Sistema Ramos, que sincroniza os registros da oralidade e do traçado gráfico e permite apreender o manuscrito* in statu nascendi*. Foram identificados dezesseis objetos textuais semântico-lexicais e trinta comentários desdobrados, organizados num* continuum *de explicitação metalinguística que vai da definição e da comparação lexicais à descrição do referente. Destacou-se a autoglosa, gesto de retenção lexical que constitui a forma mais explícita da rasura oral. Os dois alunos revelaram funcionamentos complementares: um por definição autônoma, outro por confrontação dialógica. Conclui-se que a reflexão lexical é constitutiva do ato de escrever, ainda que não deixe marcas gráficas.*

- ***PALAVRAS-CHAVE:*** *Genética Textual; Produção textual em tempo real; Enunciação; Manuscrito escolar; Alfabetização.*

Author contributions (CRediT taxonomy)

Eduardo Calil: Conceptualization, Data curation, Formal analysis, Funding acquisition, Methodology, Supervision, Writing – original draft.

Cristina Felipeto: Investigation, Writing – review & editing.

Maria Hozanete Alves de Lima: Writing – review & editing.

Data Availability Statement

All datasets supporting the findings of this study is available in a personal file and can be requested from the author(s).

References

ABAURRE, M. B. M. (Re)escrevendo: o que muda? *In*: ABAURRE, M. B. M.; FIAD, R. S.; MAYRINK-SABINSON, M. L. (org.). **Cenas de Aquisição da escrita**: o sujeito e o trabalho com o texto. Campinas: Associação de Letras do Brasil (ALB); Mercado de Letras, 1997. p. 61-69.

AUTHIER-REVUZ, J. **Ces mots qui ne vont pas de soi:** boucles réflexives et non-coïncidences du dire. Paris: Larousse, 1995.

BARBEIRO, L. F. **Os alunos e a expressão escrita**. Lisboa: Fundação Calouste Gulbenkian, 1999.

BARKAOUI, K. Think-aloud protocols in research on essay rating: an empirical study of their veridicality and reactivity. **Language Testing**, v. 28, n. 1, p. 51-75, 2011.

BENVENISTE, É. **Problèmes de linguistique générale, II**. Paris: Gallimard, 1974.

BENVENISTE, É. Últimas aulas no **Collège de France (1968 e 1969)**. São Paulo: Editora Unesp, 2014.

BORÉ, C. O rascunho, inencontrável objeto de estudo? **Debates em Educação**, [*S. l.*], v. 5, n. 10, p. 66, 2014. DOI: 10.28998/2175-6600.2013v5n10p66. Disponível em: https://seer.ufal.br/index.php/debateseducacao/article/view/1135. Acesso em: 4 abr. 2025.

BRASIL, Ministério da Educação. **Parâmetros Curriculares Nacionais. Língua Portuguesa.** Volume 2. 1ª a 4ª séries. Brasília: Secretaria de Ensino Fundamental. Ministério da Educação, Brasília (DF), 1997/1998.

CALIL, E. Título, Fim, Família F atrapalhada: uma análise do processo de rasuramento em textos escritos por crianças. **Cadernos de Estudos Linguísticos**, Campinas: Unicamp, v. 1, p. 35-48, 1996. Disponível em: https://revistas.iel.unicamp.br/index.php/sinteses/article/view/6007. Acesso em: 22 nov. 2023.

CALIL, E. **Autoria:** a criança e a escrita de histórias inventadas. Maceió: Editora da Universidade Federal de Alagoas, 1998.

CALIL, E. Processus de création et ratures: analyses d'un processus d'écriture dans un texte rédige par deux écoliers. **Langages & Société,** 103: Ecriture en acte (Irène Fenoglio, éd.), p. 31-55, 2003. Disponível em: https://www.cairn.info/revue-langage-et-societe-2003-1-page-31.htm. Acesso em: 22 nov. 2023.

CALIL, E. Modalizações autonímicas como marcas de subjetividade em processos de criação. **Revista Intercâmbio**, v. XV, 2006. Disponível em: https://revistas.pucsp.br/index.php/intercambio/article/view/3659. Acesso em: 28 dez. 2020.

CALIL, E. **Escutar o invisível:** escritura & poesia na sala de aula. São Paulo: Editora da UNESP; FUNARTE, 2008.

CALIL, E. Contos do como e do por que: Projeto didático 2012. DOI: https://doi.org/10.5281/zenodo.21416198

CALIL, E. Escritura, Memória e Associação: processos de criação de poemas por alunos recém-alfabetizados. **Revista da Anpoll**, n. 36, p. 371-402, 2014. DOI: https://doi.org/10.18309/anp.v1i36.681.

CALIL, E. O sentido das palavras e como eles se relacionam com o texto em curso: estudo sobre comentários semânticos feitos por uma díade de alunas de 7 anos de idade. **Alfa**, v. 60, n. 3, p. 531-555, 2016. DOI: http://dx.doi.org/10.1590/1981-5794-1612-4.

CALIL, E. Rasura oral comentada: definição, funcionamento e tipos em processos de escritura a dois. *In*: SILVA, A. Del Ré; CAVALCANTE, M. (org.). **A criança na/com a linguagem:** saberes em contraponto. Porto Alegre: UFRGS, 2017. p. 161-192.

CALIL, E. Sistema Ramos: método para captura multimodal de processos de escritura a dois no tempo e no espaço real da sala de aula. **Alfa**, v. 64, p. 1-28, 2020. Disponível em: https://doi.org/10.1590/1981-5794-e11705. Acesso em: 15 jan. 2024.

CALIL, E.; AMORIM, K. A.; LIRA, L. E. A criação de títulos para contos de origem inventados por escreventes novatos. *In*: CALIL, E. BORÉ, C. (org.). **Criação textual na sala de aula**. Maceió: Editora da Universidade Federal de Alagoas, 2015. p. 15-42.

CALIL, E.; FELIPETO, C. Rasuras orais semânticas na escritura a dois: a metaenunciação em histórias inventadas. **Intersecções**, ed 13, Ano 7, n. 2, p. 188-202, maio/2014. Disponível em: https://revistas.anchieta.br/index.php/RevistaInterseccoes/article/view/1193. Acesso em: 15 jan. 2024.

CALIL, E.; MYHILL, D. Dialogue, erasure and spontaneous comments during textual composition: What students' metalinguistic talk reveals about newly-literate writers' understanding of revision. **Linguistics and Education**, v. 60, p. 1-15, 2020. DOI: https://doi.org/10.1016/j.linged.2020.100875.

CAMPS, A. Production de textes en situation de groupe. **Bulletin suisse de linguistique appliquée**, n. 61, p. 119-136, 1995. Disponível em: https://libra.unine.ch/server/api/core/bitstreams/4f2e5112-48f1-4136-aad0-ac99864b4625/content. Acesso em: 15 jan. 2024.

CAMPS, A.; GUASCH, O.; MILIAN, M.; RIBAS, T. Actividad metalingüística durante el proceso de producción de un texto argumentativo. *In:* CAMPS, A., FONTICH, X. (ed.). **La actividad metalingüística como espacio de encuentro de la escritura y la gramática:** un itinerario de enseñanza e investigación en educación lingüística, San Juan: Editorial UNSJ, 2021. p. 155-169. Disponível em: https://ddd.uab.cat/record/311393. Acesso em: 30 mar. 2024.

CAMPS, A. Atividade metalinguística na aprendizagem da língua. *In*: CALIL, E.; FELIPETO, C.; BORÉ, C. (org.). **Escrita colaborativa na sala de aula**. Campinas: Pontes, 2025. p. 238-260.

CAPISTRANO, C. C. Um entre outros: a emergência da rasura na aquisição da escrita. **Linguagem em (Dis)curso**, v. 13, n. 3, p. 667-694, 2013. Disponível em: https://www.scielo.br/j/ld/a/x8yY5HDLjsSdZdNLszYRMBv/?format=pdf&lang=pt. Acesso em: 22 set 2020.

DAIUTE, C.; DALTON, B. Collaboration between children learning to write: can novices be masters? **Cognition and Instruction**, v. 10, n. 4, p. 281-333, 1993. DOI: 10.1207/s1532690xci1004_1.

DÁVALOS, A.; FERREIRO, E. Comprender la puntuación desde una perspectiva evolutiva. *In*: FERREIRO, E. *et al*. (ed.). **Reflexiones metalingüísticas de los niños sobre ciertos aspectos de la escritura del español**. Rosario: HyA Ediciones, 2019. p. 13-58.

DAVID, J. Les procédures orthographiques dans les productions écrites des jeunes enfants. **Revue des Sciences de l'Éducation**, v. 29, n. 1, p. 137-158, 2003. DOI: 10.7202/009495ar.

DE GAULMYN, M.-M. La rédaction conversationnelle: parler pour écrire. **Le Français Aujourd'hui**, n. 108, p. 73-81, 1994.

DOQUET, C. **L'écriture débutante**: pratiques scripturales à l'école élémentaire. Rennes: Presses Universitaires de Rennes, 2011.

DREW, P. 'Open' class repair initiators in response to sequential sources of troubles in conversation. **Journal of Pragmatics**, v. 28, n. 1, p. 69-101, 1997.

DUIN, A. H.; GRAVES, M. F. Intensive vocabulary instruction as a prewriting technique. **Reading Research Quarterly**, v. 22, n. 3, p. 311-330, 1987.

ERICSSON, K. A.; SIMON, H. A. **Protocol analysis:** verbal reports as data. Cambridge: MIT Press, 1993.

FABRE, C. **Les brouillons d'écoliers ou l'entrée dans l'écriture**. Grenoble: CEDITEL; L'Atelier du Texte, 1990.

FELIPETO, C. Autonímia e modalização autonímica no processo de produção em ato na escola. **Revista do GELNE**, [*S. l.*], v. 7, n. 1/2, p. 141-152, 2005. Disponível em: https://periodicos.ufrn.br/gelne/article/view/11562. Acesso em: 25 jun. 2025.

FENOGLIO, I. Énonciation et genèse d'un texte : élaboration sémantique et textualisation. *In*: MONTE, M.; PHILIPPE, G. (éd.). **Genres & textes**. Lyon: Presses universitaires de Lyon, 2014. DOI: https://doi.org/10.4000/books.pul.3114.

GARCIA-DEBANC, C.; DUVIGNAU, K.; DUTRAIT, C.; GANGNEUX, M. Enseignement du lexique et production écrite: un travail sur les verbes de déplacement à la fin de l'école primaire. **Pratiques**, n. 141-142, p. 208-232, 2009. DOI: 10.4000/pratiques.1361.

GARFINKEL, H. **Studies in ethnomethodology**. Englewood Cliffs, NJ: Prentice-Hall, 1967.

GOMBERT, J.-É. **Le développement métalinguistique**. Paris: Presses Universitaires de France, 1990.

GRÉSILLON, A. **Éléments de critique génétique:** lire les manuscrits modernes. Paris: Presses Universitaires de France, 1994.

GROSSMANN, F. Didactique du lexique: état de lieux et nouvelles orientations. **Pratiques**, n. 149-150, p. 163-183, 2011. Disponível em: https://journals.openedition.org/pratiques/1732. Acesso em: 12 jan. 2024.

GUFONI, V. Les protocoles verbaux comme méthode d'étude de la production écrite: approche critique. Études **de Linguistique Appliquée**, n. 101, p. 20-31, 1996.

HALL, N. **The development of punctuation knowledge in children aged seven to eleven**. Swindon: ESRC, Full Research Report, R000238348. 2005.

HAY, L. La critique génétique: origines et perspectives. *In*: HAY, L. **Essais de critique génétique**. Paris: Flammarion, 1979.

HERNANDEZ, F. **Cultura visual, mudança na educação e projetos de trabalho**. Porto Alegre: Artmed, 2000.

JAFFRÉ, J.-P. Les commentaires métagraphiques. **Faits de Langues**, n. 22, p. 67-76, 2003.

LIMA, M. H. A. de; CALIL, E.; FERNANDES, J. A. R.; FELIPETO, C.; COSTA, A. L. Parágrafos, reconhecimentos e comentários: atividade metalinguística durante

a gênese textual de narrativas ficcionais. **Letras de Hoje**, v. 60, n. 1, e48221, p. 1-16, 2025. DOI: https://doi.org/10.15448/1984-7726.2025.1.48221

LIMA, M. H. A. de; FERNANDES, J. A. R.; CALIL, E. O sentido das palavras e a ortografia no processo de escrita colaborativa. **Revista do GELNE**, v. 28, n. 1, p. e41401, 2026. Disponível em: https://periodicos.ufrn.br/gelne/article/view/41401. Acesso em: 19 jun. 2026.

LIRA, L. E.; CALIL, E. Gênese da fala do personagem em contos etiológicos inventados por dois alunos recentemente alfabetizados: o jogo entre o oral e o escrito. **Signum**, v.20, p. 122-150, 2017. Disponível em: https://ojs.uel.br/revistas/uel/index.php/signum/article/view/27553/21334. Acesso em: 19 jun. 2026.

MARCUSCHI, L. A. **Da fala para a escrita**: atividades de retextualização. 10. ed. São Paulo: Cortez, 2010.

MAUROUX, F.; MORIN, M.-F. Soutenir le travail des jeunes scripteurs par la conduite de l'entretien métagraphique en milieu scolaire: pourquoi et comment? **Repères**, n. 57, p. 123-140, 2018.

NEWTON, M.; JESSON, R.; PARR, J. Word choice for writing: the nature and types of children's meta-lexical awareness. **The Australian Journal of Language and Literacy**, v. 46, p. 145-160, 2023. DOI: 10.1007/s44020-023-00035-5. Disponível em: https://link.springer.com/article/10.1007/s44020-023-00035-5. Acesso em: 22 jun. 2025.

NISBETT, R. E.; WILSON, T. D. Telling more than we can know: verbal reports on mental processes. **Psychological Review**, v. 84, n. 3, p. 231-259, 1977. Disponível em: https://home.csulb.edu/~cwallis/382/readings/482/nisbett%20saying%20more.pdf. Acesso em: 01 jul. 2026.

OLINGHOUSE, N. G.; LEAIRD, J. T. The relationship between measures of vocabulary and narrative writing quality in second- and fourth-grade students. **Reading and Writing**, v. 22, n. 5, p. 545-565, 2008. DOI: 10.1007/s11145-008-9124-z.

PESSÔA, M.; CALIL, E. Relações associativas visuais na criação do título e personagens de uma história inventada. **Educação & Pesquisa**, São Paulo, v. 50, e264650, 2024, p. 1-23. DOI: https://doi.org/10.1590/S1678-4634202450264650por

REINHART, A. *et al.* Think-aloud interviews: a tool for exploring student statistical reasoning. **Journal of Statistics and Data Science Education**, v. 30, n. 2, p. 100-113, 2022. DOI: 10.1080/26939169.2022.2063209.

ROJO, R. H. R. Metacognição e produção de textos: o que as crianças sabem sobre os textos que escrevem? **Boletim da Associação Brasileira de Psicopedagogia**, São Paulo, n. 17, p. 39-56, 1989.

ROUBAUD, M.-N.; MOUSSU, M.-J. Un enseignement structuré du lexique dès la maternelle au service de l'écriture. **Pratiques**, n. 155-156, 2012. Disponível em: https://journals.openedition.org/pratiques/3484. Acesso em: 15 jan. 2024.

ROUBAUD, M.-N.; SARDIER, A. Déplier le mot: deux dispositifs d'apprentissage du lexique en faveur du réemploi. **Repères**, n. 61, 2020. Disponível em: https://journals.openedition.org/reperes/2552. Acesso em: 15 jan. 2024.

ROUSSEL, K. Les protocoles verbaux (think-aloud protocols): enjeux méthodologiques de validité pour la recherche en contexte scolaire. **Canadian Journal for New Scholars in Education**, v. 8, n. 1, p. 160-167, 2017.

SACKS, H. **Lectures on conversation**. Oxford: Blackwell, 1992.

SANTOS, J. P. L. de S.; CALIL, E. **Aprendizagem da ortografia durante a produção textual**: reflexões de uma díade recém-alfabetizada. Mauritius: Novas Edições Acadêmicas, 2018.

SMAGORINSKY, P. Rethinking protocol analysis from a cultural perspective. **Annual Review of Applied Linguistics**, v. 21, p. 233-245, 2001.

SPINILLO, A. G. "Por que você alterou isso aqui?" As razões que as crianças adotam quando fazem alterações ao revisar seus textos. **Letras de Hoje**, Porto Alegre, v. 59, n. 1, p. 32-39, jan./mar. 2015. DOI: 10.15448/1984-7726.2015.1.18348.

SUENAGA, A. **Saussure, un système de paradoxes**: langue, parole, arbitraire et inconscient. Limoges: Lambert-Lucas, 2005.

TOLCHINSKY, L. (ed.). **Developmental aspects of learning to write**. Dordrecht: Kluwer Academic Publishers, 2001.

VASS, E.; LITTLETON, K.; MIELL, D.; JONES, A. The discourse of collaborative creative writing. **Thinking Skills and Creativity**, v. 3, n. 3, p. 192-202, 2008.

VERMERSCH, P. **L'entretien d'explicitation**. Paris: ESF Sciences humaines, 2019.

VIEIRA, I. L. Aquisição da escrita: o que as crianças sabem sobre a pontuação que usam. **Revista de Letras**, Fortaleza, v. 31, n. 1/2, p. 17-24, jan./dez. 2012. Disponível em: http://www.revistadeletras.ufc.br/revista30_arquivos/04_Artigo%202_Rev_Letras_31_1_2_20123.pdf. Acesso em: 16 set. 2022.

YIN, R. K. **Estudo de caso:** planejamento e métodos. 2. ed. Porto Alegre: Bookman, 2001.

Received on January 26, 2026

Approved on July 18, 2026

Editor responsible: Gisele Cássia de Sousa