DOI: 10.5433/2237-4876.2025v28n2p82-96

Textual Complexity in Oral and Written Narratives

Juliano Desiderato **ANTONIO***

* Doctorate in Linguistics and Portuguese Language (UNESP/ Araraquara). Associate Professor – State University of Maringá (UEM/ Maringá). Contact: jdantonio@uem.br.

Abstract:
Automatic analysis of text complexity can bring several benefits, such as selecting texts with an appropriate complexity level for students, or writing documents, recipes, and medical prescriptions in a way that makes them more accessible to the general public. This study investigates the textual complexity of oral and written narratives produced by students at three educational levels: elementary, secondary, and higher education. The main goal is to determine whether textual complexity varies according to schooling level and linguistic modality (oral or written). Complexity was automatically analyzed by a computational tool that processes 200 metrics. The results indicate that written narratives tend to exhibit greater textual complexity than oral narratives, particularly among students in basic education. Among university students, the difference between modalities is less pronounced. Across various metrics, textual complexity increases with educational level. Some narratives are considered superior due to their elaboration, the use of refined language, authorial interpretations, more complex presentation of conflict, greater variety of adjectives and verb tenses, and the establishment of diverse logical-semantic relations.

Keywords:
Narratives; Textual complexity; Language modalities.

Resumo:
A análise automática da complexidade de textos pode trazer benefícios como a escolha de textos de complexidade adequada ao nível dos alunos, a redação de documentos, de receitas e de procedimentos médicos de forma mais acessível

para o público geral. Neste trabalho, investiga-se a complexidade textual de narrativas orais e de narrativas escritas produzidas por alunos de três níveis de escolaridade: fundamental, médio e superior. O objetivo principal do trabalho é verificar se a complexidade textual varia conforme o nível de escolaridade e as modalidades oral e escrita. A complexidade foi analisada automaticamente por uma ferramenta computacional que processa 200 métricas. Os resultados obtidos indicam que: as narrativas escritas tendem a apresentar maior complexidade textual em comparação com as narrativas orais, principalmente nos textos dos alunos da educação básica; nos textos de alunos do ensino superior, a diferença entre as modalidades não é tão marcada; em diversas métricas, a complexidade textual aumenta conforme aumenta o grau de escolaridade; algumas narrativas são melhores do que outras por serem mais elaboradas, utilizarem linguagem mais requintada, trazerem interpretações feitas pelos autores, apresentarem o conflito de forma mais complexa e maior variedade de adjetivos e de tempos verbais, estabelecerem relações lógico-semânticas diversas.

Palavras-chave:
Narrativas; Complexidade Textual; Modalidades de língua.

Signum: Estudos da Linguagem, Londrina, v.28, n.2, August, 2025.
Recebido em: 30/04/2025
Aceito em: 21//09/2025

Textual complexity in oral and written narratives

Juliano Desiderato ANTONIO

INTRODUCTION

With recent advances in the field of Natural Language Processing (NLP), linguistically complex tasks have increasingly been performed by computational tools. Examples include speech recognition, machine translation, automatic summarization, sentiment analysis, and, more recently, interactive dialogues with artificial intelligence chatbots such as ChatGPT. An interesting emerging application, according to Santucci *et al*. (2020), is the automatic analysis of text complexity from a linguistic perspective. Branco *et al*. (2014b), for instance, present a system that automatically classifies texts written in Portuguese based on the Common European Framework of Reference for Languages. The classification includes five levels of difficulty: A1 (easiest), A2, B1, B2, and C1 (most difficult), following the proficiency certification criteria established by Instituto Camões. The parameters used for classification, according to Branco *et al*. (2014a), include readability, lexical density, number of syllables per word, and number of words per sentence. By reducing a large set of textual properties into a smaller number of dimensions of textual complexity (Goldman; Lee, 2014), these tools can also assist in selecting texts appropriate for different learning levels (Sheehan; Flor; Napolitano, 2013).

Jensen (2009) warns that difficulty and textual complexity are distinct concepts that should not be conflated. While difficulty is subjective and can vary from reader to reader, complexity is more objective and can be measured through factual criteria such as readability indexes, word frequency (the more common a word, the lower the cognitive effort required for processing it), and non-literal expressions (metaphors, metonymies, and idioms tend to increase textual complexity because they demand greater cognitive effort to interpret).

McNamara *et al*. (2014) identify several benefits of using tools that automatically analyze text complexity: teachers can select texts with appropriate complexity for their students' level; public documents can be written in a less complex manner to ensure greater comprehension by the general public; and the same applies to medical prescriptions and procedural guidelines.

According to Leal and Aluísio (2024), the main NLP tasks related to textual complexity are text adaptation—which includes text simplification (syntactic and lexical simplification), automatic summarization, and text generation - and textual complexity assessment, which includes both text-level and sentence-level complexity evaluation. Text adaptation "allows changing the content of a text without altering its meaning, in most cases" (Leal; Aluísio, 2024, p. 539). Although there are criticisms of some adaptation tasks, such as simplification - since teaching reading is often seen as preferable to simplifying texts - certain situations justify simplification. First, teaching requires time, while simplification may help an individual access a particular text when needed, regardless of their reading proficiency. Another justification concerns the ideal level of complexity: assigning overly complex texts to learners in the early stages of reading development may discourage them. The ideal practice is to provide texts compatible with the learner's proficiency level.

In this study, we investigate the textual complexity of oral and written narratives produced by students at three educational levels: elementary, secondary and higher education. The narratives were collected after the students watched a silent film. First, the narratives were retold orally, and then, without listening to their oral versions, students produced written versions. The corpus comprises 80 texts (10 oral and 10 written narratives for each group): 10 from elementary school, 10 from secondary school, and 20 from higher education (10 from a Language and Literature

program and 10 from a Communication program). The main objective is to verify whether textual complexity varies according to educational level and whether there are differences between oral and written modalities. Textual complexity was automatically analyzed using the NILC-Metrix tool, which processes 200 metrics.

Describing these potential differences can provide insights into the linguistic features that characterize more complex texts. Such features may be used in classroom contexts to help teachers guide students in producing more complex narrative texts, whether oral or written. The research questions presented in the Methodology section aim to offer pedagogical insights in this regard.

In terms of organization, this paper includes, besides this introduction, four additional sections. The Theoretical Framework section presents the metrics used for automatic analysis of textual complexity. The Methodology section describes the data collection and transcription procedures, the analytical methods, and the research questions. Next, the Analysis section examines the results obtained from the metrics processed by the tool. Finally, the Conclusion section provides answers to the research questions outlined in Methodology.

THEORETICAL FRAMEWORK

Among the various tasks related to textual complexity, those relevant to this study concern complexity assessment. The first formulas for evaluating textual complexity emerged in the early twentieth century, and, according to Leal and Aluísio (2024), by around 1980, approximately 200 formulas had already been developed. With the advancement of NLP, tools capable of automatically analyzing text complexity were created, and many traditional metrics were incorporated into these computational systems.

One of the most well-known tools is Coh-Metrix, developed by McNamara *et al*. (2014), which analyzes 106 metrics for English across several categories: Descriptive Indices, Text Easability, Referential Cohesion, Latent Semantic Analysis (LSA), Lexical Diversity, Connectives, Syntactic Complexity, Syntactic Pattern Density, Word Information and Readability.

For Portuguese, the most notable tool is NILC-Metrix (Leal *et al*., 2021), which evaluates 200 metrics grouped into fourteen categories: Descriptive indices, Text easability metrics, Referential cohesion, LSA-semantic cohesion, Psycholinguistic measures, Lexical diversity, Connectives, Temporal lexicon, Syntactic complexity, Syntactic pattern density, Morphosyntactic word information, Semantic word information, Word frequency and Readability formulas.

According to Leal *et al*. (2021, p. 5), the NILC-Metrix tool can assist researchers in the following ways:

> (i) how text characteristics correlate with reading comprehension; (ii) which are the most challenging characteristics of a given text, that is, which characteristics make a text or corpus more complex; (iii) which texts have the most adequate characteristics to develop target learners' skills; and (iv) which parts of a text are disproportionately complex and should be simplified to meet a given audience.

Sheehan, Flor and Napolitano (2013) explain that automatic text complexity assessment systems employ readability metrics to estimate the likelihood that texts with certain linguistic characteristics will be more or less accessible to readers with varying levels of proficiency or reading ability.

Studies on readability are not recent. Flesch (1949) proposed a formula to quantify the readability of texts - essentially, a mathematical way to measure how easily a text can be read. Leal *et al*. (2023) adapted this formula for Portuguese

248,835 – [1,015 × (average words per sentence)] – [84,6 × (average syllables per word)]

The higher the resulting score, the lower the textual complexity - that is, a high metric value indicates lower complexity, and vice versa.

METHODOLOGY

Several criteria were adopted for data collection to minimize the risk of discrepancies in the corpus that might affect the results. To ensure that all participants produced texts on the same topic and with comparable features such as length and content, it was decided that data collection would be based on the viewing of a video presenting a story to be retold by the participants. The choice of the narrative genre was due to the fact that, for this type of text, the film could serve as a script to be followed by the informants, allowing for the construction of a fairly homogeneous corpus. To avoid the influence of a narrator's or characters' speech on how participants would linguistically formulate their stories, a silent film was selected, in which the sequence of scenes was sufficient for full comprehension of the plot.

The chosen video was "O pavão misterioso" ("The Mysterious Peacock"), based on a story from Northeastern Brazilian folklore of the same name, featuring puppets representing human characters. The film is 9 minutes and 20 seconds long. The story begins when the protagonist arrives in a city where the events unfold. After exploring the place and checking into a hotel, the young man attends a folk festival and falls in love with a young woman. However, her father forbids the relationship. The young man then goes to a workshop and builds a peacock-shaped aircraft, which he uses to flee the city with his beloved.

The research corpus consists of 80 narratives - 40 oral and 40 written. These narratives were produced by four groups of ten participants each, belonging to different educational levels: higher education, secondary education, and elementary education. The first part of the corpus was collected in 1996 from ten first-year students in a Language and Literature program and ten first-year students in a Communication program at a public university in the state of Paraná. The remaining narratives were collected in 2001 from a state school in the city of Sarandi, adjacent to Maringá, Paraná. The participants included ten 5th-grade students (elementary education) and ten 1st-year students (secondary education). The absence of significant linguistic differences between the regions of Londrina and Maringá (approximately ninety kilometers apart), as well as the five-year gap between the two data collection periods, ensures that neither geographical nor diachronic factors influenced the study's results.

Each text was processed using the NILC-Metrix tool, and the results were grouped by educational level (elementary, secondary, higher education – Language and Literature, higher education – Communication) and by language modality (spoken or written). Due to space limitations, this paper analyzes four key metrics: clauses per sentence, words per sentence, and the Flesch readability index.

To better guide the analysis of these metrics, five research questions were formulated. The answers to these questions are expected to provide pedagogical insights to help teachers guide their students in producing more complex narratives. The questions are as follows:

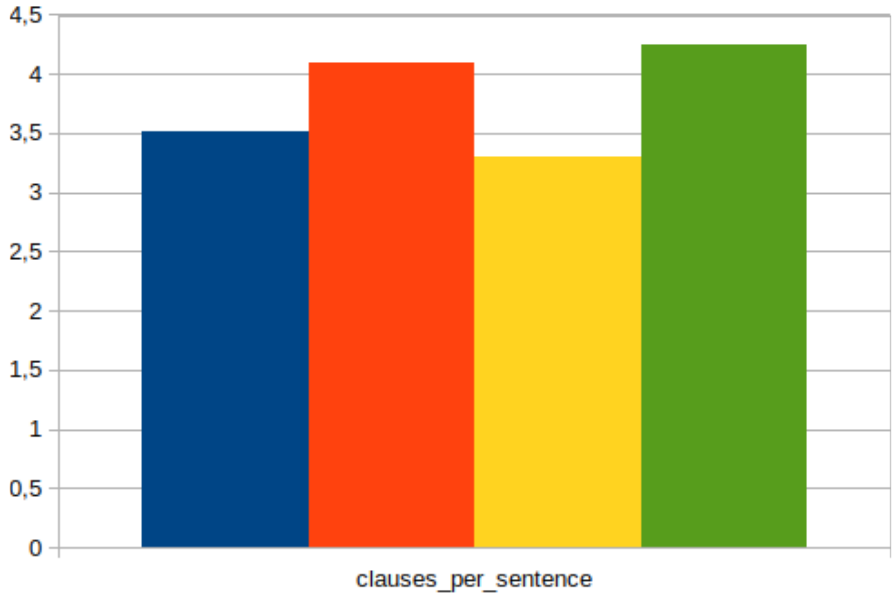
1) Is there a correlation between oral/written modalities and higher/lower textual complexity?
2) What properties characterize oral and written narratives?
3) Is there a correlation between educational level and higher/lower textual complexity?
4) What properties characterize narratives produced by students at different educational levels?
5) Are some narratives better than others? What are their distinguishing features?

ANALYSIS

To compare the textual complexity results among the student groups, the data are presented through graphs that compare each of the selected metrics. The legend for the graphs is as follows:

In Graph 1, the data correspond to the clauses per sentence metric in the oral narratives of the four groups of participants. The higher the average number of clauses per sentence, the greater the textual complexity. The number of clauses per sentence tends to increase with the educational level, except in the oral narratives of students in the Communication program, who attempt to engage the listener in the storyline. To do so, they use techniques such as repeating events and offering detailed descriptions of subtopics within the narrative, as illustrated in example (1). The use of these strategies gives greater fluidity to the plot and emphasizes the sequence of actions.

Graph 1 - Clauses per sentence in oral narratives across the four groups of participants



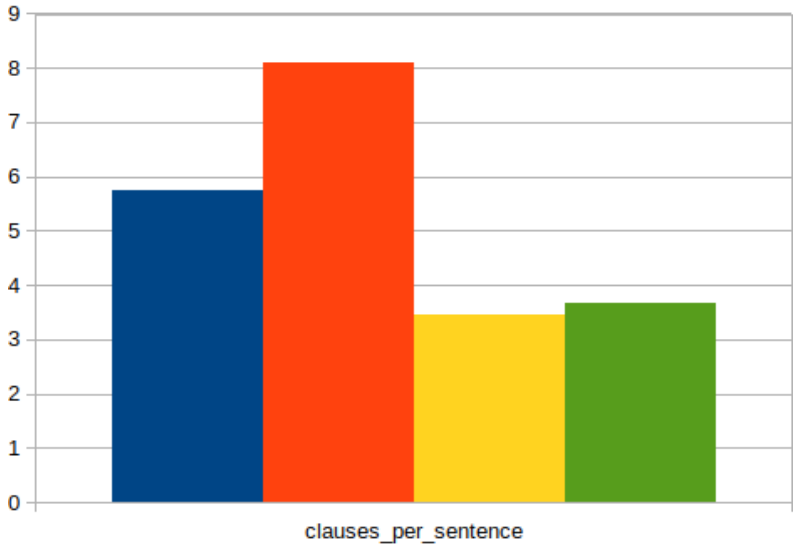
Source: Prepared by the author.

(1)

Correu, correu, desceu as escadas, então saiu pela porta, os guardas já haviam percebido o que estava acontecendo e então também o queriam pegar. E ele então conseguiu escapar e correu e correu pelos quarteirões da cidade tão pacata, que silenciada pela noite estava.

In Graph 2, the data correspond to the clauses per sentence metric in the written narratives of the four groups. As observed, in the written narratives of higher education students, there are fewer clauses per sentence. Students from both the Language and Literature and Communication programs aimed to create a more literary, almost poetic style, with a slower and more dramatic rhythm, as shown in example (2). In this example, the author juxtaposes three adjectives in three short sentences. At the end, an unnatached clause (Decat, 2011) is used to create suspense regarding the reason why the calm rhythm of the small town has been interrupted.

Graph 2 - Clauses per sentence in written narratives across the four groups of participants



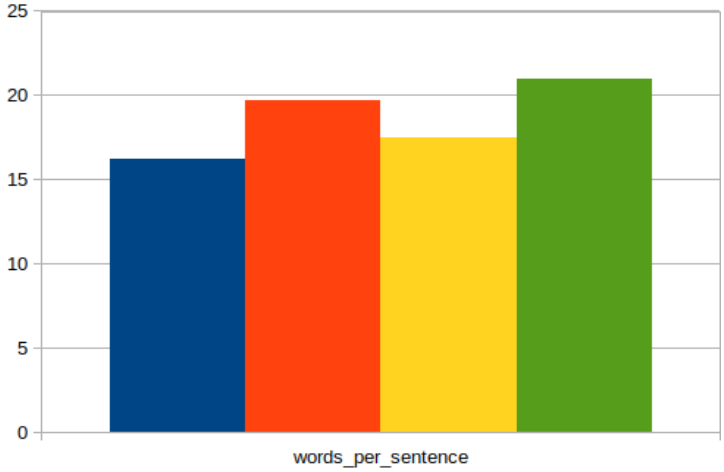
Source: Prepared by the author.

(2)

Silencioso. Observador. Compenetrado. Um estranho viajante chega, no porto de uma pequena cidade. A monotonia, o ritmo quieto e lento da pacata cidade parece ter sido quebrado. Momentaneamente mobilizada com o falecimento de uma pessoa importante.

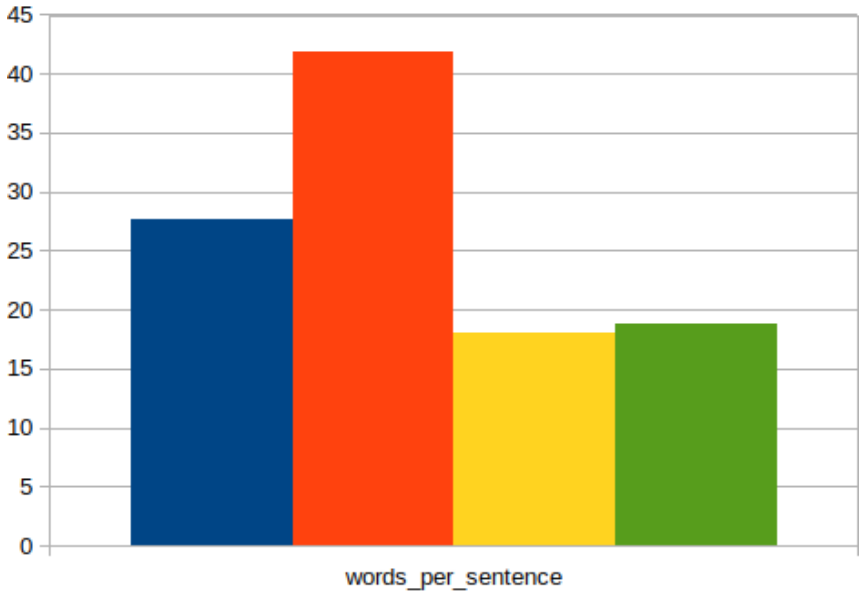
Graphs 3 and 4 present the data for the words per sentence metric, which measures the average number of words per sentence - the higher the number, the greater the textual complexity. A distinct pattern can be observed between the narratives of higher education students and those of basic education students. In the elementary group, the written narratives contained, on average, ten more points than the oral ones; in the secondary group, the difference was twenty-two points higher in favor of the written texts. However, in the Communication students' narratives, the difference was only one point higher for the written versions. Conversely, in the Language and Literature group, the oral narratives actually presented a two-point higher average than the written ones. This greater sophistication in written narratives relates, as previously mentioned, to the intent of higher education students to engage listeners in the oral retelling of the story, increasing their interest. To achieve this, higher education participants produced highly detailed narratives, as illustrated in example (3).

Graph 3 - Words per sentence in oral narratives across the four groups of participants



Source: Prepared by the author.

Graph 4 - Words per sentence in written narratives across the four groups of participants



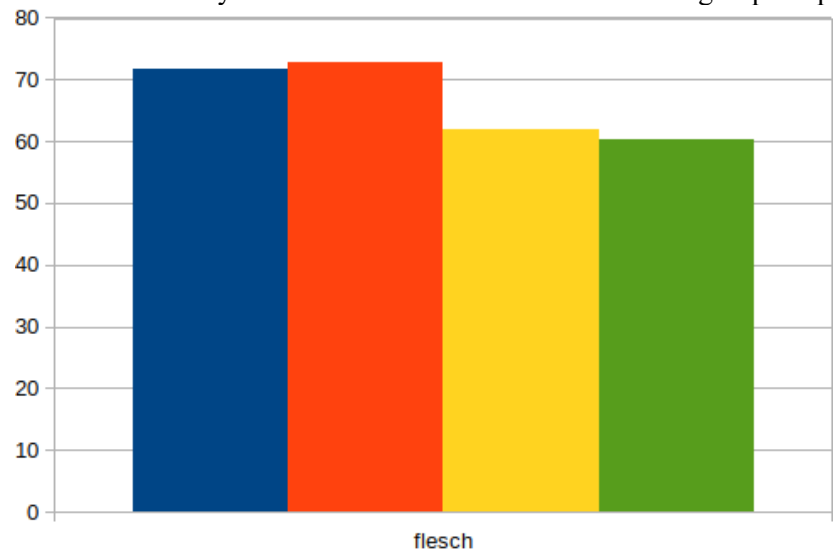
Source: Prepared by the author.

(3)

Logo ao sair escutou ao longe um barulho, um som, muita alegria, muito festejo, era uma música, era uma festa folclórica em que todos dançavam, se alegravam, comiam porque estavam muito felizes, e aquele som entrava pelos seus ouvidos, só que não tocava em seu coração.

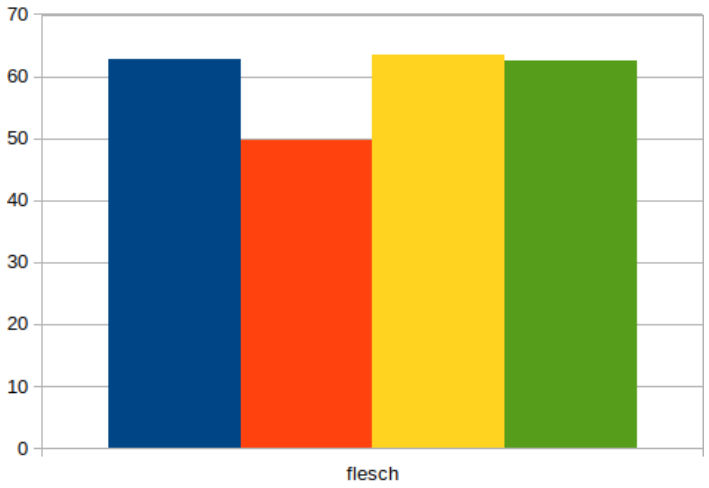
The final metric analyzed in this study is the Flesch readability index, in which lower scores indicate greater textual complexity. The results for this metric are presented in Graphs 5 and 6.

Graph 5 - Flesch readability index in oral narratives across the four groups of participants



Source: Prepared by the author.

Graph 6 - Flesch readability index in written narratives across the four groups of participants



Source: Prepared by the author.

In the oral narratives of higher education students (Graph 5), the index is lower than that of students in basic education, indicating greater complexity in the texts produced by the Language and Literature and Communication groups. In contrast, for the written narratives (Graph 6), the texts produced by secondary education students show greater complexity compared to those from the other three groups (elementary, Language and Literature and Communication). Since this metric is strongly influenced by sentence length, the results may be explained by the long sentences used by secondary students in their written narratives.

Beyond these metrics, additional indicators of complexity can be found when comparing texts from basic and higher education students. Higher education narratives tend to be more elaborated and detailed, as in example (4), taken from an oral narrative by a Communication student, compared to example (5), from an oral narrative by an elementary student.

(4)

Ele observa também danças folclóricas, a banda da cidade tocando, devem ser marchas conhecidas que compõem o folclore. Todos os tipos e níveis sociais se confraternizam, dançam alegremente, espalhando felicidade no ar durante essa noite.

(5)

Lá tinha muitas danças mas no bairro mais sociedade alta eram danças mais modernas e na sociedade baixa eles dançavam forró.

Higher education students' narratives include more references to the characters' subjectivity, as in example (6), from a Communication student's oral narrative, whereas students in basic education focus primarily on describing visible actions, as in example (7), from a secondary student's oral narrative.

(6)

A alma estava tão triste e desolada, mas ele não desistia do seu amor e pela força desse sentimento ele corria…

(7)

Aí ele sai correndo pelo portão, aí os guardas saem junto correndo atrás, mas não conseguem, ele se esconde atrás de uma parede.

In higher education narratives, the language is more refined, even incorporating figures of

speech. Examples (8) and (9), from Communication and Language and Literature students respectively, contrast with example (10), from a secondary student.

(8)
Como num ritual, tudo está pronto, e ele voa rumo à felicidade dele.

(9)
Ele planejou tudo antes e, quando menos Juliana esperava, ele entrou no seu quarto por uma corda. Os dois subiram juntos e começaram a voar pelo céu da cidade. Voavam, voavam, voavam, voavam. Juliana e Manoel estavam muito felizes e com certeza o pai de Juliana não iria impedir esse amor.

(10)
Ele pegou e foi até a moça e saiu com ela sobrevoando sobre a cidade.

In higher education students' narratives, interpretive elements are also present, as in example (11), from a Language and Literature student:

(11)
Então ele reagiu a todo obstáculo que tinha, enfrentou qualquer tipo de preconceito pra ficar com ela, os dois viveram juntos, terminaram juntos e conseguiu fugir junto com ela.

Conflict is also presented more complexly in higher education students' texts as they attempt to create greater tension, as in example (12). In contrast, basic education students present conflict more directly, as in example (13).

(12)
Quando então se aproximaram e quando iam se tocar, chega o pai, os guardas, e tudo se transforma, tudo vira um alvoroço, e a afastam dele. O jovem amante ficou desesperado, e o pai da moça, com muita braveza, dizia que não era para que ele se aproximasse dela.

(13)
O pai dela viu ela com o rapaz e mandou ela para o quarto pra ela ficar de castigo.

Differences also appear in adjective use. Basic education students use a smaller and more concrete set of adjectives - often related to visible, perceptual qualities. In contrast, higher education narratives display greater lexical diversity and include more abstract adjectives. For instance, compare example (14), from an elementary student, with the characterizations found in university students' texts: "senhor carrancudo", "pai sério e autoritário", "incansável patriarca", "poderoso coronel", "pai da moça, muito bravo, muito rígido", "pai tão bravo, tão violento", "homem rígido, obscuro", "coronel frio e rancoroso".. This difference in characterization does not concern only the young woman's father, but applies to all the characters in the narrative.

(14)
Maria tinha um pai muito bravo.

Higher education narratives also show chains of adjectives describing the same entity, as in "numa pequena cidade bem pacata, quieta, cotidiana", "longa e cansativa viagem", "linda moça de

cabelo loiro", "pequeno vilarejo nordestino", "Silencioso. Observador. Compenetrado. Um estranho viajante…", among others.

Another difference lies in the variety of verb tenses. Higher education students use a broader range, including the "pretérito mais que perfeito" ("ela também havia gostado muito dele", "sua mulher havia morrido", "haviam percebido"), the "futuro do pretérito" ("local onde se hospedaria", "Sua alma e seu coração só achariam alegria nos braços de seu amor", "ele iria fugir com sua amada"), and the "pretérito imperfeito do subjuntivo" ("como quem não acreditasse que estavam se vendo", "queria que ela pensasse", "que o ajudasse numa fuga").

Similarly, they employ a greater range of adverbial clauses, which establish diverse logical-semantic relations. Unlike basic education students, whose temporal relations are mostly marked by "quando" ("when"), higher education students use more varied connectors, such as: "Logo que eles vão se tocar, o pai dela, um coronel frio rancoroso, com dois guardas, seus capangas, se coloca à frente dela e a impede de dançar com esse homem", "Enquanto ele anda nas ruas, encontra uma moça loira".

Concessive constructions, which appear in higher education narratives, are absent from those of basic education students. For example: "parece que não se encontra muito feliz, mesmo apesar da festa estar a princípio sendo boa", "Mesmo sabendo que o pai de Juliana era um senhor muito bravo, ele não desistia".

Causal clauses also occur frequently in the higher education group, providing explanations for the characters' actions: "João de imediato não pôde fazer nada pois o pai da moça chamara dois guardas para protegê-la", "como ele tava sozinho, ele continua a andar", "procura imediatamente uma maneira de poder ficar com ela, já que o pai era tão bravo", "Mas isso de nada adianta porque o forasteiro vai atrás da moça e entra à às escondidas no seu quarto".

Although it may seem contradictory that, among higher education students, oral narratives are sometimes as or more complex than written ones - given that schooling typically enhances written proficiency - these results may suggest something different. Our hypothesis is that higher education students receive more training in oral communication skills. At university, students often present seminars orally. Furthermore, their programs prepare them for professional communication: Language and Literature students train to become teachers and engage in teaching internships, while Communication students must master public speaking for journalistic or commercial contexts. Conversely, in basic education, classroom instruction in language tends to prioritize the written modality. Few schools or textbooks emphasize developing students' oral communicative skills, despite the current educational guidelines (such as Brazil's BNCC) recommending equal attention to spoken and written language.

CONCLUSION

To conclude this study, the answers to the research questions presented in the Methodology section are summarized below.

1) Is there a correlation between oral/written modalities and higher/lower textual complexity?

Written narratives tend to show greater textual complexity than oral narratives, especially in the texts produced by basic education students. In written texts, sentences tend to be longer and contain a higher average number of words per sentence, likely due to the possibility of global planning during writing. Moreover, written narratives present greater syntactic depth (higher average dependency distance) and a larger proportion of subordinate clauses relative to the total number of clauses.

Similarly, the Flesch readability index also indicates greater complexity in written narratives compared to oral ones. An exception occurs at the higher education level: oral narratives by

university students are often as complex - or even more complex - than their written ones. This may be due to the fact that oral communication is more frequently practiced in higher education settings, especially in programs where students are trained for professions requiring strong speaking skills.

In the texts of higher education students, particularly those in Communication, the difference between modalities is not as pronounced. These students employ stylistic devices to create a more literary, almost poetic style, seeking to increase the listener's engagement.

2) What properties characterize oral and written narratives?

In oral narratives, since planning occurs almost simultaneously with production, traces of the production process can often be found in the text. In contrast, written narratives allow for prior and global planning, enabling the writer to make more deliberate choices. The editable nature of written texts also helps remove such production marks.

In oral narratives, elements such as "aí" and "daí" (roughly equivalent to "then" or "so") are frequently used to organize discourse. In written narratives, however, students use other strategies - especially adverbial subordination - to express temporal relations that, in oral texts, are primarily conveyed by "aí" and "daí." The discourse verb "pegar" (literally "to grab"), often used colloquially to mark event transitions (e.g., "ele pega e sai correndo" / "he grabs and runs"), appears frequently in oral productions.

Another difference, likely related to production conditions, concerns justifications for characters' actions, which are far more common in written narratives. Moreover, written texts tend to include richer character descriptions, even for secondary characters, due to the greater possibility of global textual planning.

3) Is there a correlation between educational level and higher/lower textual complexity?

Yes. Textual complexity increases with educational level across several metrics.

For the clauses per sentence metric, in the oral narratives of all four groups, the number of clauses per sentence rises with education level, except in the Communication students' group, who - as noted earlier - emphasize dramatic pacing and listener engagement. In the written narratives of higher education students, the number of clauses per sentence is lower than in the texts produced by basic education students, since the former tend to adopt a more literary, rhythmic style with shorter sentences.

Regarding the words per sentence metric, it tends to increase with the level of education, with the exception of texts written by Communication students who, as previously discussed, sought to create a more literary, almost poetic style, with a slower, more dramatic rhythm. To achieve this, they made use of shorter sentences. The use of shorter sentences to create a more literary style thus explains the lower average number of words per sentence in the narratives written by higher education students.

The final metric is the Flesch readability index. The lower the index, the higher the complexity. In the oral narratives of higher education students, greater complexity is observed in the texts produced by students in Language and Literature and Communication programs. In written narratives, however, it is the texts of secondary education students that exhibit higher complexity compared to those of the other three groups (elementary education, higher education - Language and Literature, and higher education - Communication). Since this metric is strongly influenced by sentence length, these results can be explained by the long sentences used by secondary education students in their written narratives.

Other qualitative differences include:

- Higher education students' narratives are more elaborated.

- They present more details related to the characters' subjectivity, while basic education students focus on describing observable actions.

- Their language is more refined, often employing figures of speech.

- Their texts contain interpretive commentary by the narrators.

- Conflicts are depicted more complexly, creating greater tension.

- They use a greater variety of adjectives—many abstract—and sometimes chain multiple adjectives for a single referent.

- They employ a wider range of verb tenses.

- They establish diverse logical-semantic relations through adverbial clauses.

4) What properties characterize narratives produced by students at different educational levels?

The following properties distinguish narratives according to educational level:

- Higher education students' narratives tend to be more elaborate.

- Higher education students' narratives include more details regarding the characters' subjectivity, whereas in narratives by basic education students the focus is on describing what is visible in the film.

- In higher education students' texts, the language is more refined, with the use of rhetorical figures.

- Interpretations made by the informants are also found in higher education students' texts.

- The conflict is presented in a more complex way in higher education students' texts, in an attempt to create greater tension. In narratives by basic education students, the conflict is presented more directly.

- Higher education students' narratives exhibit a greater variety of adjectives, many of them more abstract, whereas texts by basic education students show a much smaller range of adjectives, generally limited to qualities easily perceived visually.

- Higher education students' texts also display a greater variety of verb tenses compared to texts by basic education students, both in oral and written narratives.

- Higher education students' texts also show a greater variety of adverbial clauses, establishing diverse logical-semantic relations.

5) Are some narratives better than others? What are their distinguishing features?

Yes. Some narratives are qualitatively superior to others. The main characteristics of the best narratives include:

- Attempt to increase the interlocutor's engagement by creating a greater sense of anticipation regarding the events and more frequently introducing hypotheses about the events that are part of the story.

- Relate events through explanations and reflections, rather than merely juxtaposing them.

- Refer to the setting using rhetorical figures.

- Explore the characters' emotional states.

- Intertwine the setting with the characters' emotions.

- Create tension, suspense, and a dramatic atmosphere.

- Use a more refined vocabulary.

- Provide more details.

- Justify the events.

- Develop interpretative hypotheses.

- Present the conflict in a more complex manner.

- Use a variety of adjectives, including combining multiple adjectives to qualify a referent.

- Employ a greater variety of verb tenses.

- Establish logical-semantic relationships through adverbial clauses.

In terms of future research, several directions emerge, both in teaching and in studies of spoken and written discourse. Pedagogically, writing activities can be designed based on the characteristics identified in this research. Teachers may encourage students, during text revision exercises, to incorporate some of the features observed in more complex narratives. For linguistic studies, further research on spoken and written discourse may explore phenomena such as discourse markers, topic introduction and maintenance mechanisms, and structural components of narrative organization, among others.

REFERENCES

BRANCO, António; RODRIGUES, João; COSTA, Francisco; SILVA, João; VAZ, Rui. Assessing automatic text classification for interactive language learning. *In*: INTERNATIONAL CONFERENCE ON INFORMATION SOCIETY (i-Society), 2014, London. [*Anais…*]. London: IEEE**,** 2014a. p. 70-78, DOI: 10.1109/i-Society.2014.7009014.

BRANCO, António; RODRIGUES, João; COSTA, Francisco; SILVA, João; VAZ, Rui. Rolling out text categorization for language learning assessment supported by language technology. *In*: BAPTISTA, Jorge; MAMEDE, Nuno; CANDEIAS, Sara; PARABONI, Ivandré; PARDO, Thiago Alexandre Salgueiro; NUNES, Maria das Graças Volpe (ed.). *Computational processing of the portuguese language*. Cham: Springer, 2014b. p. 256–261. DOI: 10.1007/978-3-319-09761-9_29.

DECAT, Maria Beatriz Nascimento. *Estruturas desgarradas em língua portuguesa*. Campinas: Pontes, 2011.

FLESCH, Rudolf. *The art of readable writing*. New York; Evanston: Harper & Row Publishers, 1949. Disponível em: https://dc135.wordpress.com/wp-content/uploads/2012/11/flesch-the-art-of-readable-writing.pdf. Acesso em: 1 out. 2025.

GOLDMAN, Susan Rotter; LEE, Carol Diane. Text complexity: state of the art and the conundrums it raises. *Elementary School Journal*, Chicago, v. 115, n. 2, p. 290–300, 2014. Disponivel em: https://www.journals.uchicago.edu/doi/10.1086/678298. Acesso em: 1 out. 2025.

JENSEN, Kristian Tangsgaard Hvelplund. Indicators of Text Complexity. *Copenhagen Studies in Language*, Copenhagen**,** n. 37, p. 61-80, 2009. Disponível em: https://research.cbs.dk/en/publications/indicators-of-text-complexity. Acesso em: 1 out. 2025.

LEAL, Sidney Evaldo; ALUÍSIO, Sandra Maria. Complexidade textual e suas tarefas relacionadas. *In*: CASELI, Helena de Medeiros; Nunes, Maria das Graças Volpe (org.) *Processamento de linguagem natural*: conceitos, técnicas e aplicações em português. 2. ed. São Carlos: BPLN, 2024. p. 537-562. Disponível em: https://brasileiraspln.com/livro-pln/2a-edicao/Livro_PLN_BPLN_2ed.pdf. Acesso em: 8 out. 2025.

LEAL, Sidney Eevaldo; DURAN, Magali Sanches; SCARTON, Carolina Evaristo; HARTMANN, Nathan Siegle; ALUÍSIO, Sandra Maria. NILC-Metrix: assessing the complexity of written and spoken language in Brazilian Portuguese. *Language Resources and Evaluation*, Dordrect, v.8, p. 73–110, 2021. Disponível em: https://doi.org/10.1007/s10579-023-09693-w. Acesso em: 1 out. 2025.

MCNAMARA, Danielle Siobhan; GRAESSER, Arthur C.; MCCARTHY, Philip M.; CAI, Zhiqiang. *Automated evaluation of text and discourse with Coh-Metrix*. Cambridge: Cambridge University Press, 2014.

SANTUCCI, Valentino; SANTARELLI, Filippo; FORTI, Luciana; SPINA, Stefania. Automatic classification of text complexity. *Applied Sciences*, Basileia, v. 10, n. 20, p. 1–19, 2020. Disponível em: https://www.mdpi.com/2076-3417/10/20/7285. Acesso em: 8 out. 2025.

SHEEHAN, Kathleen Marie; FLOR, Michael; NAPOLITANO, Diane. A two-stage approach for generating unbiased estimates of text complexity. *In*: WORKSHOP ON NATURAL LANGUAGE PROCESSING FOR IMPROVING TEXTUAL ACCESSIBILITY, 2., 2013, Georgia. [*Anais...*] Georgia: ACL, 2013. p. 49-58. Disponível em: https://aclanthology.org/W13-1506.pdf. Acesso em: 8 out. 2025.