

# Complexidade textual em narrativas orais e em narrativas escritas

Juliano Desiderato **ANTONIO\***

\* Doutor em Linguística e Língua Portuguesa (Unesp/ Araraquara). Professor associado - Universidade Estadual de Maringá. Contato: [jdantonio@uem.br](mailto:jdantonio@uem.br).

## **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.

## **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.

# Complexidade textual em narrativas orais e em narrativas escritas

---

Juliano Desiderato ANTONIO

## INTRODUÇÃO

Com as recentes evoluções no campo do Processamento de Linguagem Natural (PLN), tarefas linguisticamente complexas vêm sendo realizadas cada vez mais por ferramentas computacionais. Alguns exemplos são o reconhecimento da fala humana, a tradução automática, a sumarização automática, a análise de sentimentos e, mais recentemente, os diálogos que podem ser estabelecidos com chatbots de inteligência artificial como o ChatGPT. Uma possibilidade interessante que vem se desenvolvendo, segundo Santucci *et al.* (2020), é a análise automática da complexidade de textos do ponto de vista linguístico. Branco *et al.* (2014b), por exemplo, apresentam um sistema que classifica automaticamente textos produzidos em Língua Portuguesa com base no Quadro Europeu Comum de Referência para Línguas. A classificação tem cinco níveis de dificuldade: A1 (mais fácil), A2, B1, B2 e C1 (mais difícil) e utiliza os critérios do Instituto Camões para certificação de proficiência. Os critérios utilizados para classificação, segundo Branco *et al.* (2014a) são leiturabilidade, densidade lexical, quantidade de sílabas por palavra, quantidade de palavras por sentença. Ao reduzirem um grande número de propriedades textuais a um número menor de dimensões de complexidade textual (Goldman; Lee, 2014), as ferramentas podem também auxiliar na seleção de textos adequados para diferentes níveis de aprendizagem (Sheehan; Flor; Napolitano, 2013).

Jensen (2009) alerta para o fato de que dificuldade e complexidade textual são dois conceitos distintos que não devem ser confundidos. Enquanto a dificuldade é subjetiva e pode variar de leitor para leitor, a complexidade é mais objetiva e pode ser calculada a partir de critérios factuais, como os índices de leiturabilidade, a frequência das palavras (quanto mais comum é uma palavra menos esforço cognitivo é dispensado no processamento daquela palavra), a não literalidade (metáforas, metonímias e expressões idiomáticas podem afetar a complexidade textual pelo fato de exigirem maior esforço cognitivo para serem processadas).

McNamara *et al.* (2014) apresentam alguns benefícios da utilização de ferramentas que analisam automaticamente a complexidade de textos: professores podem escolher textos de complexidade adequada ao nível de seus alunos, documentos públicos podem ser redigidos de forma menos complexa tendo em vista maior facilidade de compreensão pelo público em geral, o mesmo podendo ser dito com relação a receitas de remédios e procedimentos médicos.

Para Leal e Aluísio (2024), as principais tarefas de PLN ligadas à complexidade textual são a adaptação textual, composta por simplificação textual, que abrange a simplificação sintática e a simplificação lexical, a sumarização automática e a elaboração textual; e a avaliação da complexidade textual, na qual estão incluídas a avaliação da complexidade textual e a avaliação da complexidade sentencial. A adaptação textual “permite alterar o conteúdo do texto sem mudar seu significado, na maior parte das vezes” (Leal; Aluísio, 2024, p. 539). Embora haja críticas a determinadas tarefas de adaptação textual, como a simplificação, por exemplo, uma vez que ensinar a ler é melhor do que simplificar, há algumas exceções que justificam a simplificação. Em primeiro lugar, ensinar leva tempo, ao passo que simplificar pode ajudar uma pessoa a ter acesso a um determinado texto no momento em que ela precisa, independente do seu nível de competência de leitura. Outra justificativa diz respeito ao nível ideal de complexidade. Atribuir textos demasiadamente complexos a alunos em processo de aprendizagem pode desestimulá-los. O ideal é apresentar textos compatíveis com o nível de aprendizagem.

Nesta pesquisa, investiga-se a complexidade textual de narrativas orais e de narrativas escritas produzidas por alunos de três níveis de escolaridade: fundamental, médio e superior. As narrativas foram coletadas após os alunos assistirem um filme mudo. Primeiramente, as narrativas foram recontadas oralmente e, na sequência, sem poderem ouvir o texto oral que haviam produzido, os alunos produziram a versão escrita. Ao todo, o *corpus* é formado por 80 textos (10 narrativas orais e 10 narrativas escritas de cada um dos grupos): 10 informantes do ensino fundamental, 10 informantes do ensino médio e 20 informantes do ensino superior (10 do curso de Letras e 10 do curso de Comunicação). O objetivo principal do trabalho é verificar se a complexidade textual varia conforme o nível de escolaridade e se há distinções entre as modalidades oral e escrita. A complexidade textual das narrativas foi analisada automaticamente pela ferramenta NILC-Matrix, que leva em conta 200 métricas.

A descrição dessas possíveis diferenças pode fornecer subsídios para que se descrevam os recursos empregados nos textos mais complexos. E esses recursos poderão ser utilizados em sala de aula por professores para auxiliarem seus alunos a produzirem textos narrativos mais complexos, seja na modalidade oral, seja na modalidade escrita. As respostas às perguntas de pesquisa que serão apresentadas na seção de Metodologia poderão fornecer esses subsídios.

Em termos de organização, este trabalho apresenta, para além desta introdução, outras quatro seções. Na seção de Fundamentação, apresentam-se as métricas utilizadas para a análise automática da complexidade textual. Os procedimentos de coleta e transcrição das narrativas são objeto da seção de Metodologia, bem como os procedimentos analíticos e as perguntas de pesquisa. Na sequência, analisam-se os resultados das métricas processadas pela ferramenta utilizada. Por fim, na Conclusão, apresentam-se as respostas às perguntas de pesquisa propostas na seção de Metodologia.

## FUNDAMENTAÇÃO TEÓRICA

Dentre as tarefas de complexidade textual, as que interessam a este projeto são as de avaliação da complexidade. As primeiras fórmulas de avaliação da complexidade textual surgiram no início do século XX e, segundo Leal e Aluísio (2024), por volta de 1980, já havia aproximadamente 200 fórmulas. Com o desenvolvimento do PLN, criaram-se ferramentas que analisam automaticamente a complexidade de textos, e algumas das métricas tradicionais são utilizadas por essas ferramentas.

Uma ferramenta bastante conhecida é utilizada é o Coh-Matrix, desenvolvido por McNamara *et al.* (2014), que investiga 106 métricas do inglês, pertencentes às seguintes categorias: Descriptive, Text Easability Principal Component Scores, Referential Cohesion, Latent Semantic Analysis (LSA), Lexical Diversity, Connectives, Situation Model, Syntactic Complexity, Syntactic Pattern Density, Word Information, Readability.

Para o português destaca-se a ferramenta NILC-Matrix (Leal *et al.*, 2021), que analisa 200 métricas agrupadas em quatorze categorias: Índices descritivos, Métricas de facilidade textual, Coesão referencial, LSA-Coesão semântica, Medidas psicolinguísticas, Diversidade lexical, Conectivos, Léxico temporal, Complexidade sintática, Densidade do padrão sintático, Informação morfossintática das palavras, Informação semântica das palavras, Frequência de palavras, Fórmulas de leiturabilidade.

De acordo com Leal *et al.* (2021, p. 5), o NILC-Matrix ferramenta pode auxiliar os pesquisadores nas seguintes tarefas:

- (i) como as características do texto se correlacionam com a compreensão de leitura;
- (ii) quais são as características mais desafiadoras de um determinado texto, ou seja, quais características tornam um texto ou corpus mais complexo;
- (iii) quais textos possuem as características mais adequadas para desenvolver as habilidades dos alunos-alvo;
- (iv) quais partes de um texto são desproporcionalmente complexas e devem ser simplificadas para

atender a um determinado público. (tradução nossa)<sup>1</sup>

Sheehan, Flor e Napolitano (2013) afirmam que os sistemas automáticos de análise de complexidade textual utilizam métricas de leiturabilidade para avaliar a probabilidade de textos com determinadas características linguísticas serem mais ou menos acessíveis a leitores de diferentes níveis ou com diferentes habilidades de leitura.

Os estudos a respeito da leiturabilidade não são recentes. Flesch (1949) apresenta uma fórmula para calcular a facilidade da leitura dos textos. Em outras palavras, a fórmula quantifica a leiturabilidade dos textos. Leal *et al.* (2021) adaptaram essa fórmula para o português:  $248,835 - [1,015 \times (\text{média de palavras por frase})] - [84,6 \times (\text{média de sílabas por palavra})]$

Quanto maior o resultado da métrica, menor é a complexidade textual, ou seja, métrica alta, complexidade baixa e vice-versa.

## METODOLOGIA

Foram adotados alguns critérios para a coleta dos dados para que se evitasse ao máximo o risco de diferenças nos resultados causados por discrepâncias no corpus. Para que os textos de todos os informantes fossem sobre um mesmo assunto e fossem semelhantes em aspectos como extensão, conteúdo, etc, decidiu-se que a coleta dos dados seria feita a partir da exibição de um vídeo com uma história que seria recontada pelos sujeitos da pesquisa. A opção pela narrativa proveio do fato de que, para a produção desse tipo de texto, o filme serviria como um script a ser seguido pelos informantes, o que permitiria a obtenção de um corpus bastante homogêneo. Para se evitar que houvesse influência das falas do narrador ou de personagens sobre a maneira como os informantes formulariam linguisticamente a história, a solução foi procurar um filme mudo, cuja sequência de cenas fosse suficiente para a compreensão do enredo.

O vídeo escolhido foi “O pavão misterioso”, que se baseia em uma história do folclore nordestino de mesmo nome e que tem como personagens bonecos que representam seres humanos. Com duração de 9 minutos e 20 segundos, o enredo do filme se inicia com a chegada do protagonista à cidade onde acontecerão os fatos. Após conhecer o local e instalar-se em um hotel, o rapaz vai a uma festa popular e conhece uma moça por quem se apaixona. Entretanto, o pai da moça proíbe o amor dos dois. O rapaz vai, então, a uma oficina e constrói uma aeronave em forma de pavão que utiliza para fugir da cidade com sua amada.

O *corpus* da pesquisa é constituído por 80 narrativas, sendo 40 orais e 40 escritas. Essas narrativas foram produzidas por quatro grupos de dez informantes cada, pertencentes a diferentes níveis de ensino, a saber, Ensino Superior, Ensino Médio e Ensino Fundamental. A primeira parte do corpus foi coletada em 1996, com dez alunos do primeiro ano do curso de Letras e 10 alunos do primeiro ano do curso de Comunicação Social de uma universidade pública do Estado do Paraná. As demais narrativas foram coletadas em 2001, em uma escola estadual do município de Sarandi, conurbado com Maringá, Paraná. Os informantes eram dez alunos da 5ª série do Ensino Fundamental e dez informantes da 1ª série do Ensino Médio. A inexistência de diferenças significativas nas normas linguísticas das regiões de Londrina e de Maringá, distantes noventa quilômetros entre si, bem como a diferença de apenas cinco anos entre a coleta da primeira parte e da segunda parte do *corpus* permitem dizer que fatores geográficos ou diacrônicos não influenciaram os resultados da pesquisa.

Cada texto foi processado pela ferramenta NILC-Metrix, e os resultados foram agrupados por nível de ensino (Ensino Fundamental, Ensino Médio, Ensino Superior – Letras, Ensino Superior – Comunicação) e modalidade de língua (língua falada e língua escrita). Por questão de limitação de

---

<sup>1</sup> (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.

espaço, serão analisadas, neste trabalho, quatro métricas: *clauses per sentence*, *words per sentence* e índice Flesch de leituraabilidade.

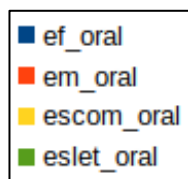
Elaboraram-se cinco perguntas de pesquisa para um melhor direcionamento da análise das métricas. Espera-se que as respostas às perguntas de pesquisa possam fornecer subsídios para que professores auxiliem seus alunos na produção de narrativas mais complexas. As perguntas foram as seguintes:

- 1) Existe correlação entre modalidades orais/escritas e maior/menor complexidade?
- 2) Que propriedades caracterizam as narrativas orais e as narrativas escritas?
- 3) Existe correlação entre nível de escolaridade e maior/menor complexidade?
- 4) Que propriedades caracterizam as narrativas produzidas por alunos de diferentes níveis de escolaridade?
- 5) As narrativas produzidas por alguns alunos são melhores do que as produzidas por outros? Quais são suas características?

## ANÁLISE

Com a finalidade de se comparar os resultados da complexidade dos textos dos estudantes de cada grupo, apresentam-se os dados na forma de gráficos que comparam cada uma das métricas destacadas. A legenda dos gráficos é a seguinte:

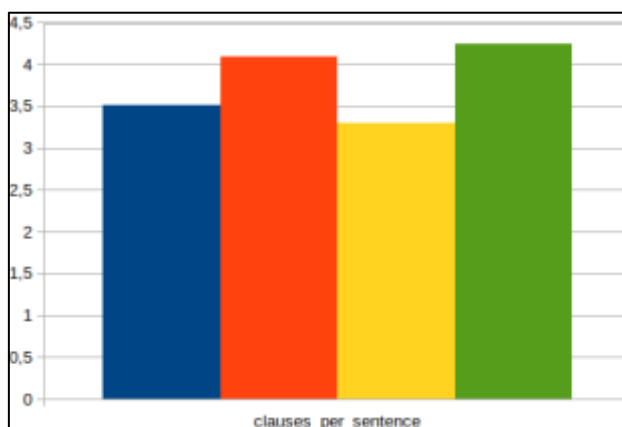
**Figura 1** – Legenda dos gráficos



Fonte: Elaborado pelo autor.

No gráfico 1, encontram-se os dados referentes à métrica *clauses per sentence* nas narrativas orais dos quatro grupos de informantes. Quanto maior a média de orações por sentença, maior a complexidade. O número de orações por sentença vai aumentando conforme aumenta o grau de escolaridade, com exceção das narrativas orais dos informantes do curso de Comunicação Social, que tentam envolver o ouvinte no enredo. Para isso, utilizaram-se de recursos como a repetição de eventos e a descrição detalhada de um subtópico da narrativa, como no exemplo (1). O uso desses recursos dá maior fluidez ao enredo, enfatizando as ações.

**Gráfico 1** - Métrica *clauses per sentence* nas narrativas orais dos quatro grupos de informantes



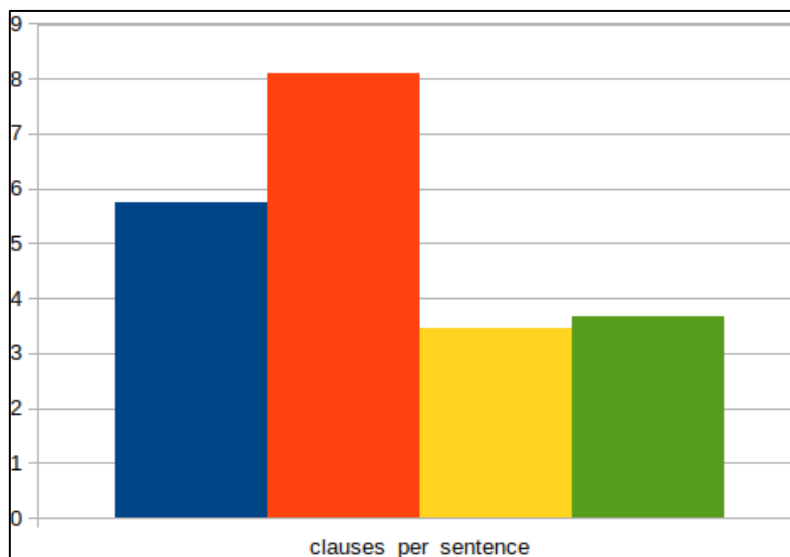
Fonte: Elaborado pelo autor.

(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.

No gráfico 2, encontram-se os dados referentes à métrica *clauses per sentence* nas narrativas escritas dos quatro grupos de informantes. Como pode ser observado, nas narrativas escritas dos alunos do Ensino Superior há menos orações por sentença. Os alunos de Letras e de Comunicação procuraram criar um estilo mais literário, quase poético, com um ritmo mais pausado, mais dramático, como no exemplo (2), em que o autor do texto justapõe três adjetivos em três períodos distintos. No final do período, o autor também utiliza uma oração desgarrada (Decat, 2011) para criar um certo suspense em relação ao motivo de o ritmo da pacata cidade ter sido quebrado.

**Gráfico 2** - Métrica *clauses per sentence* nas narrativas escritas dos quatro grupos de informantes



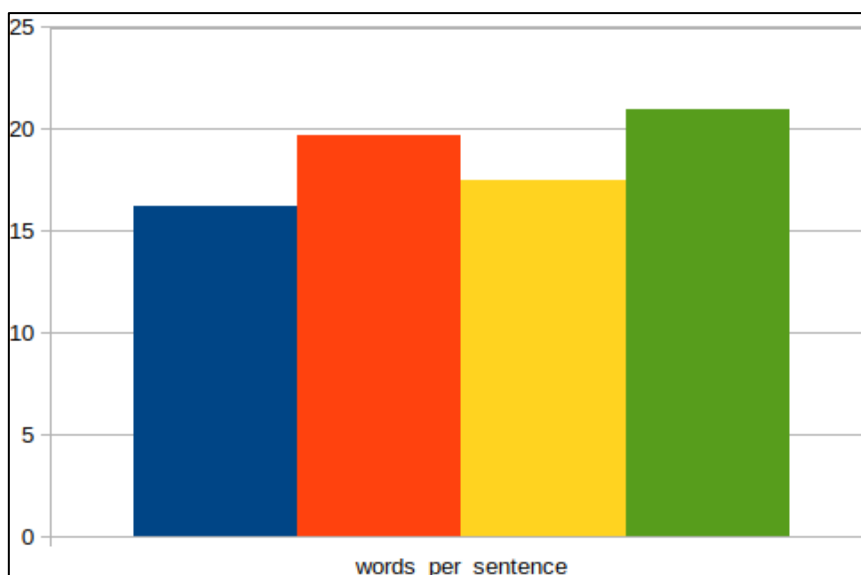
Fonte: Elaborado pelo autor.

(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.

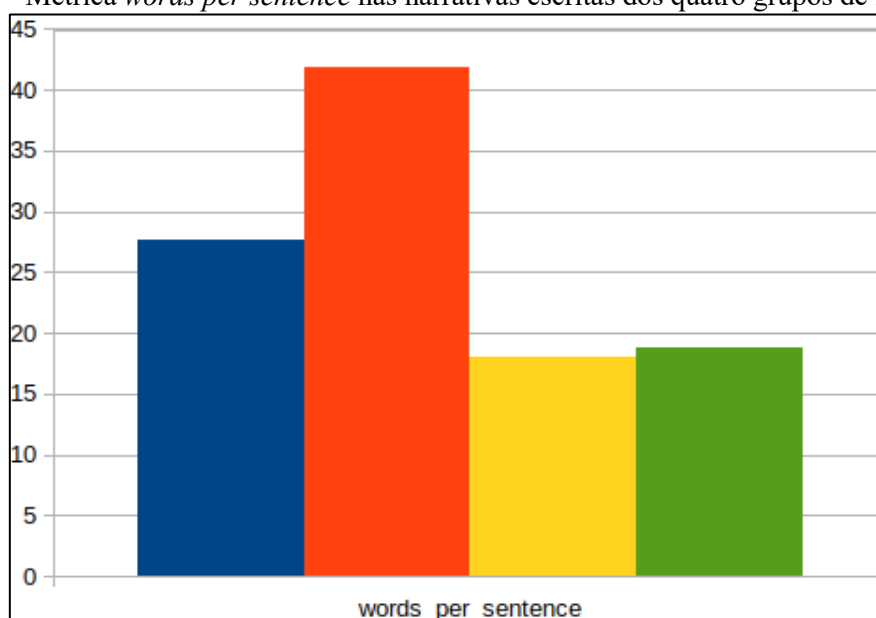
Nos gráficos 3 e 4, apresentam-se os dados referentes à métrica *words per sentence*. Essa métrica calcula a média de palavras por sentença. Quanto mais palavras, maior a complexidade. Observa-se um comportamento diferente nas narrativas dos alunos do ensino superior em relação aos textos dos alunos da educação básica. Enquanto nas narrativas dos alunos do ensino fundamental houve uma diferença de dez pontos a mais nas narrativas escritas e nas narrativas dos alunos do ensino médio, uma diferença 22 pontos a mais também para as narrativas escritas, nos textos dos alunos de Comunicação, a diferença foi de apenas um ponto a mais para as narrativas escritas. Por outro lado, nas narrativas dos alunos do curso de Letras, as narrativas orais é que apresentaram uma média de dois pontos a mais do que as narrativas escritas. Essa maior sofisticação das narrativas escritas está relacionada, como já dito anteriormente, à intenção de os alunos do ensino superior de envolver o ouvinte na recontagem oral da história, aumentando o interesse do interlocutor. Para isso, os informantes do ensino superior produzem textos cheios de detalhes, como em (3).

**Gráfico 3** - Métrica *words per sentence* nas narrativas orais dos quatro grupos de informantes



Fonte: Elaborado pelo autor.

**Gráfico 4** - Métrica *words per sentence* nas narrativas escritas dos quatro grupos de informantes



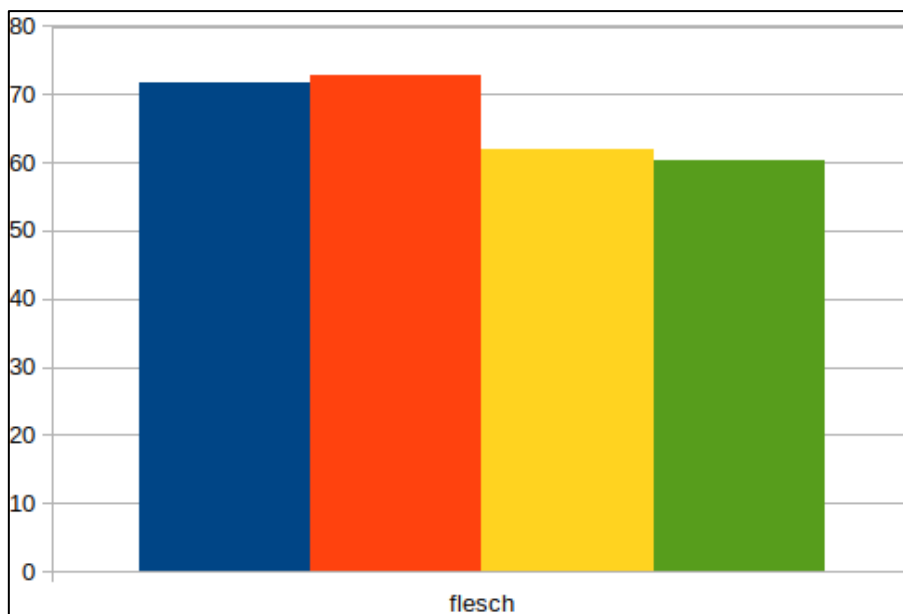
Fonte: Elaborado pelo autor.

(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.

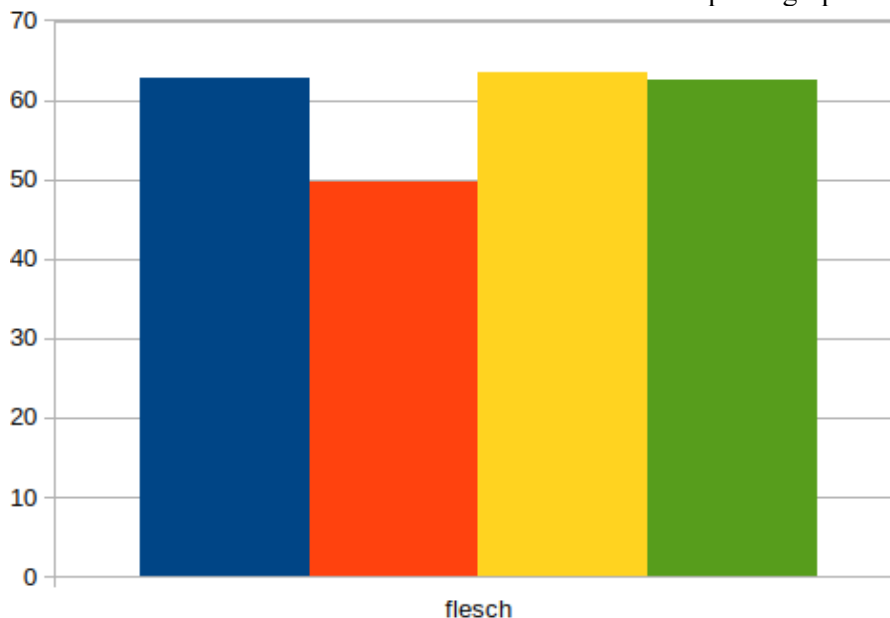
A última métrica analisada no trabalho é o índice Flesch de leitura. Quanto menor o índice, maior a complexidade. Os dados referentes a essa métrica são apresentados nos gráficos 5 e 6.

**Gráfico 5** - Índice Flesch de leitura nas narrativas orais dos quatro grupos de informantes



Fonte: Elaborado pelo autor.

**Gráfico 6** - Índice de leitura Flesch nas narrativas escritas dos quatro grupos de informantes



Fonte: Elaborado pelo autor.

Nas narrativas orais dos alunos do ensino superior (gráfico 5), o índice é menor do que nas narrativas dos alunos da educação básica, indicando maior complexidade nos textos dos alunos dos cursos de Letras e de Comunicação. Já nas narrativas escritas (gráfico 6), os textos dos alunos do ensino médio é que apresentam maior complexidade em relação aos textos dos outros três grupos (ensino fundamental, ensino superior Letras e ensino superior Comunicação). Como essa métrica é fortemente influenciada pelo tamanho dos períodos, a explicação para esses resultados pode ser justificada pelas longas sentenças utilizadas pelos alunos do ensino médio em suas narrativas escritas.

Além dessas métricas, outras diferenças de complexidade podem ser encontradas no que diz respeito aos textos dos alunos da educação básica e da educação superior. As narrativas dos alunos do ensino superior tendem a ser mais elaboradas, a trazer mais detalhes, como no exemplo (4), encontrado em uma narrativa oral de um aluno do curso de Comunicação, em comparação com o exemplo (5), encontrado em uma narrativa oral de um aluno do ensino fundamental.

(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.

As narrativas dos alunos do ensino superior trazem mais detalhes referentes à subjetividade dos personagens, como em (6), encontrado em uma narrativa oral de um informante do curso de Comunicação, ao passo que, nas narrativas dos alunos da educação básica, o foco está em narrar aquilo que é visível no filme, como no exemplo (7), encontrado em uma narrativa oral de um informante do ensino médio.

(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.

Nos textos dos alunos do ensino superior, a linguagem é mais requintada. Há, inclusive, o uso de figuras de linguagem. Observem-se os exemplos (8), encontrado em uma narrativa oral de um informante do curso de Comunicação, e (9), encontrado em uma narrativa oral de um informante do curso de Letras, em comparação com o exemplo (10), encontrado em uma narrativa oral de um informante do ensino médio.

(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.

Nos textos dos alunos de ensino superior, também se encontram interpretações feitas pelos informantes, como em (11), encontrado em uma narrativa oral de um informante do curso de Letras.

(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.

O conflito também é apresentado de forma mais complexa nos textos dos alunos do ensino superior em uma tentativa de criar uma maior tensão, como no exemplo (12). Nas narrativas dos alunos da educação básica, o conflito é apresentado de forma mais direta, como em (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.

Há também diferenças no uso dos adjetivos. Nos textos dos alunos da educação básica, há uma variedade bem menor de adjetivos. Em geral, são adjetivos que designam qualidades que podem ser detectadas facilmente pela visão. Também são adjetivos mais frequentes. Por outro lado, nas narrativas dos alunos do ensino superior, há uma maior variedade de adjetivos, muitos deles mais abstratos. Um exemplo de diferença no uso dos adjetivos pode ser observado em (14), encontrado em uma narrativa oral de um aluno do ensino fundamental, quando contrastado com as caracterizações encontradas em textos produzidos por alunos do ensino superior: “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”. Essa diferença de caracterização não diz respeito apenas ao personagem pai da moça, mas a todos os personagens da narrativa.

(14)

Maria tinha um pai muito bravo.

Nos textos dos alunos do ensino superior, podem ser encontrados exemplos nos quais há um encadeamento de adjetivos para caracterizar uma mesma entidade, como em “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...”, dentre tantos outros.

Também se encontra, nos textos produzidos pelos informantes do ensino superior, uma maior variedade de tempos verbais em relação aos textos dos alunos da educação básica, tanto nas narrativas orais quanto nas escritas. Alguns desses tempos são o pretérito mais que perfeito, como em “ela também havia gostado muito dele”, “sua mulher havia morrido”, “havam percebido”; futuro do pretérito, como em “local onde se hospedaria”, “Sua alma e seu coração só achariam alegria nos braços de seu amor”, “ele iria fugir com sua amada”; pretérito imperfeito do subjuntivo, como em “como quem não acreditasse que estavam se vendo”, “queria que ela pensasse”, “que o ajudasse numa fuga”.

Nos textos dos alunos do ensino superior também se encontra uma maior variedade de orações adverbiais, que estabelecem relações lógico-semânticas diversas. Ao contrário do que acontece com os textos dos alunos da educação básica, as relações de tempo estabelecidas nos textos dos alunos do ensino superior são sinalizadas por conectivos que vão além do “quando”, como em “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”.

As construções concessivas, presentes nas narrativas dos alunos do ensino superior, não foram utilizadas pelos alunos da educação básica. Os exemplos a seguir foram retirados de narrativas produzidas por alunos do ensino superior: “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”.

As orações causais ocorrem com grande frequência, nos textos dos alunos do ensino superior, com a finalidade de justificar ações realizadas pelos personagens, como em “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”.

Embora possa parecer contraditório que nos textos dos alunos do ensino superior as narrativas orais não apresentem maior complexidade do que as escritas, uma vez que o alto grau de escolarização supõe maior domínio desta última modalidade, talvez os resultados apontem para uma conclusão bastante diferente. É nossa hipótese que os alunos da educação superior tenham tido maior preparo em termos de habilidades comunicativas orais. Por exemplo, na universidade, é comum que os alunos precisem apresentar seminários oralmente. Os próprios cursos aos quais pertencem os informantes também devem preparar os alunos para situações de comunicação oral. O curso de Letras forma professores e os alunos fazem estágio. Os profissionais de Comunicação também precisam ter as habilidades comunicativas orais bem desenvolvidas, uma vez que as atividades que realizarão profissionalmente envolvem falar com o público, seja na área jornalística, seja na área comercial. Por outro lado, na educação básica, em geral, a prioridade nas aulas de língua é a modalidade escrita. Poucas escolas e poucos materiais didáticos têm a preocupação de desenvolver as habilidades comunicativas orais dos alunos, embora nos documentos oficiais atuais (como a BNCC), a orientação é que a língua falada seja contemplada assim como a língua escrita.

## CONCLUSÃO

Para concluir este trabalho, apresentam-se as respostas às perguntas de pesquisa propostas na seção de Metodologia.

### 1) Existe correlação entre modalidades orais/escritas e maior/menor complexidade?

As narrativas escritas tendem a apresentar maior complexidade textual em comparação com as narrativas orais, principalmente nos textos dos alunos da educação básica. Na escrita, os períodos tendem a ser um pouco mais longos e, conseqüentemente, apresentar um número médio maior de palavras por sentença, provavelmente devido à possibilidade de se planejar o texto escrito globalmente. Também nas narrativas escritas, a distância média na árvore de dependências é bem maior do que nas narrativas orais e há uma média mais alta de orações subordinadas em relação ao total de orações do texto.

Da mesma forma, o índice Flesch de leitura também aponta maior complexidade nas narrativas escritas do que nas orais. Uma exceção ocorre no ensino superior. As narrativas orais chegam a ser tão ou mais complexas do que as escritas pelo fato de a oralidade ser mais exigida nas atividades que serão desempenhadas pelos futuros profissionais.

Nos textos de alunos do ensino superior, principalmente do curso de Comunicação, a diferença entre as modalidades não é tão marcada. Os alunos utilizam recursos estilísticos para criar um estilo mais literário, quase poético, com a intenção de tentar aumentar o envolvimento do interlocutor.

### 2) Que propriedades caracterizam as narrativas orais e as narrativas escritas?

Nas narrativas orais, devido ao fato de o planejamento ser quase concomitante com a produção, podem ser encontradas no texto marcas do processo de produção. Por outro lado, nas narrativas escritas, a possibilidade de se planejar o texto globalmente permite que o produtor faça suas escolhas de forma antecipada. O caráter editável do texto escrito também contribui para que as marcas de produção não permaneçam no texto.

Nas narrativas orais, os elementos “aí” e “daí” são utilizados com frequência na organização textual. Já nas narrativas escritas, os informantes se utilizam de outros recursos, especialmente da hipotaxe adverbial, para estabelecer as relações temporais sinalizadas primordialmente por “aí” e “daí” nos textos orais. O verbo “pegar” discursivo é utilizado com frequência nas narrativas orais, como em “ele pega e sai correndo”.

Uma outra diferença que provavelmente está relacionada às condições de produção das narrativas orais e das narrativas escritas do corpus diz respeito à apresentação de justificativas para as ações dos personagens, bem mais frequente nas narrativas escritas. Além disso, na escrita, a possibilidade de planejar o texto globalmente acaba por permitir que haja uma melhor caracterização dos personagens, mesmo que sejam secundários.

### 3) Existe correlação entre nível de escolaridade e maior/menor complexidade?

Sim. A complexidade textual aumenta conforme aumenta o grau de escolaridade em diversas métricas.

Na métrica *clauses per sentence*, nas narrativas orais dos quatro grupos de informantes, o número de orações por sentença vai aumentando conforme aumenta o grau de escolaridade, com exceção das narrativas orais dos informantes do curso de Comunicação Social, que, como observado anteriormente, tentam envolver o ouvinte criando um clima de tensão, com pausas dramáticas. Já nas narrativas escritas dos alunos do Ensino Superior há menos orações por sentença do que nos textos escritos dos alunos da educação básica. Como já dito em outro ponto deste trabalho, os alunos de Letras e de Comunicação procuraram criar um estilo mais literário, quase poético, com um ritmo mais pausado, mais dramático.

Relativo à métrica *words per sentence*, a métrica vai aumentando conforme aumenta o grau de escolaridade, com exceção dos textos dos alunos de Comunicação Social, que, como já discutido anteriormente, procuraram criar um estilo mais literário, quase poético, com um ritmo mais pausado, mais dramático. Para isso, utilizaram-se de períodos mais curtos. Já a utilização de períodos mais curtos para criar um estilo mais literário explica a menor média de palavras por sentença nas narrativas escritas dos alunos de ensino superior.

A última métrica é o índice Flesch de leitura. Quanto menor o índice, maior a complexidade. Nas narrativas orais dos alunos do ensino superior, observa-se maior complexidade nos textos dos alunos dos cursos de Letras e de Comunicação. Já nas narrativas escritas, os textos

dos alunos do ensino médio é que apresentam maior complexidade em relação aos textos dos outros três grupos (ensino fundamental, ensino superior Letras e ensino superior Comunicação). Como essa métrica é fortemente influenciada pelo tamanho dos períodos, a explicação para esses resultados pode ser justificada pelas longas sentenças utilizadas pelos alunos do ensino médio em suas narrativas escritas.

- as narrativas dos alunos do ensino superior tendem a ser mais elaboradas;
- as narrativas dos alunos do ensino superior trazem mais detalhes referentes à subjetividade dos personagens, ao passo que, nas narrativas dos alunos da educação básica, o foco está em narrar aquilo que é visível no filme;
- nos textos dos alunos do ensino superior, a linguagem é mais requintada. Há, inclusive, o uso de figuras de linguagem;
- nos textos dos alunos de ensino superior, também se encontram interpretações feitas pelos informantes;
- o conflito também é apresentado de forma mais complexa nos textos dos alunos do ensino superior em uma tentativa de criar uma maior tensão. Nas narrativas dos alunos da educação básica, o conflito é apresentado de forma mais direta;
- nas narrativas dos alunos do ensino superior, há uma maior variedade de adjetivos, muitos deles mais abstratos, ao passo que, nos textos dos alunos da educação básica, há uma variedade bem menor de adjetivos. Em geral, são adjetivos que designam qualidades que podem ser detectadas facilmente pela visão;
- nos textos dos alunos do ensino superior, também se encontra uma maior variedade de tempos verbais em relação aos textos dos alunos da educação básica, tanto nas narrativas orais quanto nas escritas;
- nos textos dos alunos do ensino superior também se encontra uma maior variedade de orações adverbiais, que estabelecem relações lógico-semânticas diversas.

4) Que propriedades caracterizam as narrativas produzidas por alunos de diferentes níveis de escolaridade?

Algumas das propriedades que caracterizam as narrativas produzidas por alunos de diferentes níveis de escolaridade são apresentadas a seguir:

- as narrativas dos alunos do ensino superior tendem a ser mais elaboradas;
- as narrativas dos alunos do ensino superior trazem mais detalhes referentes à subjetividade dos personagens, ao passo que, nas narrativas dos alunos da educação básica, o foco está em narrar aquilo que é visível no filme;
- nos textos dos alunos do ensino superior, a linguagem é mais requintada. Há, inclusive, o uso de figuras de linguagem;
- nos textos dos alunos de ensino superior, também se encontram interpretações feitas pelos informantes;
- o conflito também é apresentado de forma mais complexa nos textos dos alunos do ensino superior em uma tentativa de criar uma maior tensão. Nas narrativas dos alunos da educação básica, o conflito é apresentado de forma mais direta;
- nas narrativas dos alunos do ensino superior, há uma maior variedade de adjetivos, muitos deles mais abstratos, ao passo que, nos textos dos alunos da educação básica, há uma variedade bem menor de adjetivos. Em geral, são adjetivos que designam qualidades que podem ser detectadas facilmente pela visão;
- nos textos dos alunos do ensino superior, também se encontra uma maior variedade de tempos verbais em relação aos textos dos alunos da educação básica, tanto nas narrativas orais quanto nas escritas;
- nos textos dos alunos do ensino superior também se encontra uma maior variedade de orações adverbiais, que estabelecem relações lógico-semânticas diversas.

5) As narrativas produzidas por alguns alunos são melhores do que as produzidas por outros? Quais são suas características?

Sim. Algumas narrativas são melhores do que as outras. Algumas das características são

listadas a seguir:

- tentar aumentar o envolvimento do interlocutor, criando um clima de mais expectativa em relação aos acontecimentos e inserindo com maior frequência hipóteses para os eventos que fazem parte da história;
- relacionar os eventos por meio de explicações e reflexões, e não apenas justapor esses eventos;
- fazer referência ao ambiente utilizando figuras de linguagem;
- explorar o estado emocional dos personagens;
- entrelaçar o cenário com as emoções dos personagens;
- criar tensão, suspense, clima dramático;
- utilizar vocabulário mais requintado;
- trazer mais detalhes;
- justificar os acontecimentos;
- elaborar hipóteses interpretativas;
- apresentar o conflito de forma mais complexa;
- utilizar adjetivos variados, inclusive encadeando mais de um adjetivo para qualificar um referente;
- utilizar uma maior variedade de tempos verbais;
- estabelecer relações lógico-semânticas por meio de orações adverbiais.

Em termos de trabalhos futuros, despontam diversas possibilidades, tanto no que diz respeito ao ensino quanto no que diz respeito aos estudos do texto falado e do texto escrito. Atividades de ensino de escrita de textos narrativos podem ser elaboradas a partir das características descritas nesta pesquisa. Os professores podem sugerir aos alunos, nas atividades de reescrita dos textos, que incluam algumas das características observadas nos textos investigados. Quanto aos estudos de língua falada e língua escrita, diversos fenômenos podem ser investigados, como marcadores discursivos, mecanismos de introdução e manutenção de tópicos, partes estruturais da narrativa, dentre outros.

## REFERÊNCIAS

- 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, Ivandr ; 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](https://doi.org/10.1007/978-3-319-09761-9_29).
- DECAT, Maria Beatriz Nascimento. *Estruturas desgarradas em língua portuguesa*. Campinas: Pontes, 2011.
- FLESCHE, 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. Disponível 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](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*, Dordrecht, 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.