

PUCRS

ESCOLA DE HUMANIDADES
PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS
DOUTORADO EM LETRAS

ASAFE DAVI CORTINA SILVA

**ACESSIBILIDADE TEXTUAL LINGUODIDÁTICA: ANÁLISE DA COMPLEXIDADE DE TEXTOS
PARA O ENSINO DE INGLÊS COMO LÍNGUA ADICIONAL**

Porto Alegre
2023

PÓS-GRADUAÇÃO - *STRICTO SENSU*



Pontifícia Universidade Católica
do Rio Grande do Sul

**PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO GRANDE DO SUL
PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS
LINGUÍSTICA**

**ACESSIBILIDADE TEXTUAL LINGUODIDÁTICA: ANÁLISE
DA COMPLEXIDADE DE TEXTOS PARA O ENSINO DE
INGLÊS COMO LÍNGUA ADICIONAL**

ASAFE DAVI CORTINA SILVA

ORIENTADORA: PROF^a. DR^a. CRISTINA BECKER LOPES PERNA

Tese apresentada como requisito parcial para a obtenção do título de doutor pelo Programa de Pós-Graduação em Letras da Pontifícia Universidade Católica do Rio Grande do Sul. Área: Linguística.

Orientadora: Prof^a. Dr^a. Cristina Becker Lopes Perna.

**PORTO ALEGRE
2023**

ASAFE DAVI CORTINA SILVA

ACESSIBILIDADE TEXTUAL LINGUODIDÁTICA: ANÁLISE DA COMPLEXIDADE DE TEXTOS PARA O ENSINO DE INGLÊS COMO LÍNGUA ADICIONAL

Tese apresentada como requisito parcial para a obtenção do título de doutor pelo Programa de Pós-Graduação em Letras da Pontifícia Universidade Católica do Rio Grande do Sul. Área: Linguística.

Aprovada com louvor em: 04 de janeiro de 2023.

BANCA EXAMINADORA:

Prof^a. Dr^a. Cristina Becker Lopes Perna (Orientadora)

Prof^a. Dr^a. Elisa Marchioro Stumpf (UFPEL/UFRGS)

Prof^a. Dr^a. Heloísa Orsi Koch Delgado (La Salle/UFRGS)

Prof^a. Dr^a. Lilian Cristine Hubner (PUCRS)

Prof^a. Dr^a. Michele Saraiva Carilo (University of Edinburgh)

Porto Alegre
2023

AGRADECIMENTOS

Em todas as minhas publicações, inicio meus agradecimentos por Aquele que me escolheu desde antes do meu nascimento e que tem me capacitado em cada passo: Deus. Tudo o que tenho alcançado é pela graça dEle, que sempre se manifestou de várias formas e por meio de distintas pessoas na minha vida, sendo a primeira delas, minha mãe.

Volto em 1994, quando eu iniciava minha jornada acadêmica no primeiro ano do ensino fundamental. Sei o quanto a minha mãe trabalhou para que eu tivesse acesso à melhor educação e o quanto ela priorizou o meu bem, muitas vezes até em detrimento de si mesma. Lembro que mesmo trabalhando longas horas, quase todos os dias ela voltava correndo para casa na sua hora do almoço para me levar a pé para a escola. Quando pequeno, talvez eu não enxergasse o quanto isso significava, mas hoje sei que tenho o que tenho e sou o que sou por ela. Toda minha dedicação, todo meu esforço e toda minha vontade de sempre fazer o meu melhor, aprendi com ela. Mãe, espero um dia poder te agradecer por tudo que fizeste por mim e que tenhas orgulho do que **NÓS** alcançamos.

Sou grato ao meu melhor amigo, quem esteve comigo para escrever minha dissertação e agora a minha tese, o Harry. Dizem que só conhecemos o amor mais puro quando somos adotados por um animal e desde que ele entrou na minha vida, tenho certeza disso. Obrigado por me confortar em momentos difíceis e por me fazer sorrir quando eu queria chorar.

Também tenho duas lembranças muito fortes na minha memória e que, certamente, contribuíram para minha formação não só acadêmica, mas também como pessoa: meu avô sentado em uma cadeira de praia na frente de casa lendo um livro e minha avó me ajudando com os temas de Matemática. Como disse Joseph Joubert, "*A palavra empolga, o exemplo ensina.*". Então, tenho certeza de que parte da minha paixão pelas Letras e minha facilidade com os números foi influenciada pelo exemplo dos meus avós maternos. Enquanto uns faziam questão de ostentar títulos, eles me ensinavam pelo exemplo e pela humildade e, para mim, minha mãe e meus avós maternos são os maiores mestres e doutores que já conheci.

Ao meu pai, padrasto, madrasta, primas, primos, tias e tios, que me ajudaram a escrever a minha história.

Ao Dirceu, por estar sempre comigo e me aguentar nos momentos de estresse. Sou grato por me incentivar a crescer e por querer crescer comigo.

Sou grato, também, pelas orientações e paciência da minha querida orientadora, Prof^a. Dr^a. Cristina Perna – uma das pessoas mais inteligentes que já conheci e que, ao mesmo tempo, é um exemplo de humildade. Obrigado por acreditar em mim, obrigado por compartilhar sua sabedoria comigo e obrigado por ser uma inspiração não só acadêmica, mas também profissional. Além de ter tido a incrível oportunidade de ser orientado pela senhora, também tive a honra de dividirmos algumas turmas como professores na graduação e de aprender, também, no campo profissional.

Como disse Simone Weil, “A amizade não se busca, não se sonha, não se deseja; ela se exerce (é uma virtude)”. Então, agradeço aos amigos que exercem essa virtude com tanta paciência, que sempre celebram com minhas vitórias, que me encorajam nos meus momentos difíceis e que foram portos seguros nesses tempos difíceis que estamos vivendo no País. Em especial, Adriana Rocco, Alini Hoenig, Bertha Sonntag, Bruna Bettamello, Bruno Couras, Camila Heck, Eduardo Moll, Larissa Petruzzellis e Paula Giacobbo – com vocês, minha vida fica mais bonita. *“You got troubles, I've got them too, there isn't anything I wouldn't do for you. We stick together and see it through, cause you've got a friend in me.”*

Aos professores e professoras com quem tive a honra de aprender, desde o ensino fundamental (Miriam e Adão Paulo) até a graduação e pós-graduação, há um pouco de cada um de vocês no professor que hoje eu sou. Muito obrigado não só pelos ensinamentos teóricos, mas por terem ensinado com humanidade e com bondade.

Aos meus colegas por todo o compartilhamento e companheirismo, especialmente à Aline Fay (ex-professora, banca de TCC/mestrado/qualificação, chefe, colega e parceira de artigos) por ser uma inspiração de profissional, de pesquisadora e de pessoa.

Aos meus alunos por me fazerem cada dia mais apaixonado pela docência.

E por fim, mas não menos importante, um agradecimento especial às professoras que fazem parte da minha banca de defesa da tese, Prof^a. Dr^a. Elisa Stumpf, Prof^a. Dr^a. Heloísa Delgado, Prof^a. Dr^a. Lilian Hubner e Prof^a. Dr^a. Michele Carilo. Muito obrigado pelo tempo dedicado à leitura deste trabalho e pelas

contribuições que farão. Sou grato por ter a chance de ter um trabalho avaliado por profissionais tão sérias e de excelência acadêmica.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

“Que ninguém se engane, só se consegue a simplicidade através de muito trabalho.”
- Clarice Lispector

“Uma palavra que não representa uma ideia é uma coisa morta, da mesma forma que uma ideia não incorporada em palavras não passa de uma sombra.”
- Lev Vygotsky

RESUMO

Complexidade e Acessibilidade Textual são temas que são estudados no meio acadêmico e linguístico há muitos anos, principalmente tendo *corpora* formados por textos sobre Saúde e/ou Direito. Quase inexistentes são as pesquisas que investigam complexidade e acessibilidade de textos voltados para o ensino de línguas adicionais e mais raras ainda são aquelas que estudam complexidade com base em índices textuais numéricos. Esta tese, portanto, tem como tema o conceito por nós cunhado Acessibilidade Textual Linguodidática (ATLD) que se refere a análises de textos utilizados para o ensino de inglês como língua adicional (ILA) e a adoção de estratégias de simplificação que sejam coerentes com o perfil de estudantes. O principal objetivo deste trabalho é analisar o potencial de complexidade nos níveis lexicais, semânticos e sintáticos dos textos utilizados para atividades de interpretação dos livros *Elementary* e *Pre-intermediate* da quarta edição da coleção *English File* da editora *Oxford*. Para isso, coletamos por meio do *software* Coh-Metrix as seguintes métricas: Polissemia, Idade de Aquisição, Familiaridade com Palavras de Conteúdo, Concretude Lexical, Relação Type-Token, Narratividade, Simplicidade Sintática, Similaridade Sintática, Coesão Referencial e Índice Flesch. Por meio dos indicativos de cada métrica e tomando as descrições das habilidades de leitura definidas pelo *Common European Framework of Reference* (CEFR) como orientações, interpretamos os potenciais de complexidade dos textos do nosso *corpus* de estudo para verificar sua [in]adequação para o público proposto e, a partir desses resultados, apresentar orientações de estratégias de simplificação que possam não só potencializar a acessibilidade textual, mas otimizar o desenvolvimento linguístico. Para corroborar a interpretação das métricas, também realizamos análises comparativas com valores dos índices de textos que são, *a priori*, simples e outros, complexos. Além disso, por meio das métricas, analisamos se existem progressões de complexidade que possam colaborar para uma manutenção linguística coerente, gradual e contínua. Ainda que o foco do trabalho tenha sido os livros *Elementary* e *Pre-intermediate*, também apresentamos uma análise prévia dos textos do livro *Intermediate* com a finalidade de observar a progressão de complexidade e a necessidade de possíveis pesquisas futuras com os demais níveis. Nossos resultados demonstram que existem elementos de alta complexidade nos textos para o perfil de estudante para os quais são propostos especialmente em relação ao léxico e semântica. No livro *Elementary*, 54% das métricas de léxico e semântica indicam baixa complexidade, enquanto 36% e 10% complexidade média e alta, respectivamente. A respeito da sintaxe para esse mesmo nível, 52% dos índices indicam baixa complexidade, 28%, média e 20% alta. Já para o nível *Pre-intermediate*, a indicação das métricas de léxico e semântica configura 43,63% tendo baixa complexidade, 37,27%, média e 19,09%, alta. Os índices de sintaxe do nível pré-intermediário indicam 42,72% de complexidade baixa, 36,36% de média e 20,90% de alta. Também apresentamos nossas interpretações como professores e linguistas para corroborar ou refutar as indicações de algumas métricas (visto que complexidade, para ser estimada com maior precisão, deve levar em consideração o público para o qual um texto é proposto). Por fim, apresentamos indicações de estratégias específicas de simplificação para cada um dos textos destacados pelas métricas como tendo alto potencial de complexidade e, também, desenvolvemos um sucinto manual de utilização de métricas para analisar complexidade e tomar decisões a respeito de estratégias de simplificação.

PALAVRAS-CHAVE: Complexidade Textual. Acessibilidade Textual. Acessibilidade Textual Linguodidática. Inglês como Língua Adicional. Coh-Metrix.

ABSTRACT

Text accessibility and complexity are topics that have been researched in academic and linguistic fields for quite a long time, especially for fields such as Health and/or Law. However, studies that focus on texts used for teaching purposes are almost inexistent (even rarer are the ones that adopt numeric textual metrics to conduct the research). The theme of this Ph.D. thesis is a concept we call 'Linguodidactic Text Accessibility' (LTA), which refers to analyses of texts for teaching English as an Additional Language (EAL) and to adopting simplification strategies that are coherent to the students' profiles. The main purpose of our research is the analysis of the complexity potential of lexical, semantical, and syntactical levels in texts used for interpretation exercises in the fourth edition of the books *Elementary* and *Pre-intermediate* by Oxford Press. To do so, we collected, using Coh-Metrix, the following metrics: Polysemy, Age of Acquisition, Familiarity with Content Words, Lexical Concreteness, Type-Token Ratio, Narrativity, Syntactical Simplicity, Syntactical Similarity, Referential Cohesion, and Flesch Index. By analyzing the orientations of each index and being guided by the description of reading abilities proposed by the Common European Framework of Reference (CEFR), we have interpreted the complexity potential of the texts in order to verify their adequacy (or lack thereof) for the specified audience. Based on the results, we present suggestions of simplification strategies that can be applied to potentialize the accessibility of the texts and, simultaneously, to optimize language learning. To corroborate our interpretation, we have also conducted a comparative analysis between the indexes of our texts and the results for the same metrics of texts that are priorly considered simple, and others that are complex. In addition, we have used the metrics to observe whether or not the complexity is progressively increased to ensure a coherent, gradual and continuous language development. Even though we focused on the books *Elementary* and *Pre-intermediate*, we also present a preliminary analysis of the texts of the *Intermediate* book meant to support the progression analysis and to offer information to justify future research with the other levels of the book collection. Our research shows that there are some elements of high complexity (mainly related to lexicon and semantics) in the texts, especially considering the profile of the students for whom they are proposed. 54% of the lexicon and semantics indexes of the texts from the *Elementary* book indicate low complexity, while 36% and 10%, medium and high complexity, respectively. In terms of syntax for the same level, 52% of the indexes are indicative of low complexity, 28%, of medium, and 20%, of high. The lexicon and semantics indexes of the *Pre-intermediate* book suggest that 43,63% of the metrics are indicators of low complexity against 37,27%, of medium, and 19,09%, of high. The syntax indexes of this level suggest 42,72% of low, 36,36% of medium, and 20,90% of high complexity. Our analyses as professors and linguists are also presented to help corroborate or refute the indication of the indexes once text complexity must take into consideration the reader to whom a text is proposed. In the last two chapters of this thesis, we present some specific suggestions of text simplification for the texts that were classified as having high complexity, and a succinct manual that explains how the indexes can be used to estimate complexity and, based on their results, guide language professionals to choose the most appropriate accessibility strategies.

KEYWORDS: Text Complexity. Text Accessibility. Linguodidactic Text Accessibility. English as an Additional Language. Coh-Metrix.

LISTA DE FIGURAS

Figura 1 - Níveis de referência do CEFR	45
Figura 2 – Capa dos livros <i>English file 4th edition</i>	64
Figura 3: Métricas e níveis textuais	86
Figura 4: Exemplo de classificação de valores de métrica por meio do Excel.....	91
Figura 5 - Árvores sintáticas do texto CP4	98
Figura 6 - Árvores sintáticas do texto CP10	99
Figura 7 - Árvores sintáticas do texto CP13	100
Figura 8 - Exemplo de Saussure (1975) para explicar a relação entre significante e significado.....	112
Figura 9: Exemplo da Escada de Abstração de Hayakawa	113
Figura 10 - Rótulos de cores para interpretação do potencial de complexidade das métricas	121
Figura 11 - Exemplo de texto do livro <i>Elementary</i>	146
Figura 12 - Exemplo de texto com sintaxe complexa no livro <i>Elementary</i>	148
Figura 13 - Exemplo de texto com léxico complexo do livro <i>Elementary</i>	149
Figura 14 - Exemplo de texto curto do livro <i>Pre-intermediate</i>	151
Figura 15 - Exemplo de texto mais extenso do livro <i>Pre-intermediate</i>	152
Figura 16 - Exemplo de texto com léxico complexo do livro <i>Pre-intermediate</i>	153
Figura 17 - Exemplo de texto com léxico balanceado do livro <i>Pre-intermediate</i>	154
Figura 18 - Exemplo de texto que exige conhecimentos de coesão referencial do livro <i>Pre-intermediate</i>	154
Figura 19 - Exemplo de extensão de texto do livro <i>Intermediate</i>	156
Figura 20 - Exemplo de texto com marcadores visuais do livro <i>Intermediate</i>	157
Figura 21 - Exemplo de texto com léxico complexo do livro <i>Intermediate</i>	158
Figura 22 - Parte do texto E1 com auxílio de recursos visuais como disposto no livro.....	171
Figura 23 - Ilustração da seleção de unidades lexicais novas por meio do Excel	179
Figura 24 - Ilustração dos procedimentos lógicos para determinação de aproximação de valores métricos	185
Figura 25 – Exemplo de SS no texto E14 do livro <i>Elementary</i>	195
Figura 26 – Exemplo de SS no texto E10 do livro <i>Elementary</i>	196
Figura 27 - Análise sintática das três primeiras frases do texto mais complexo por SS.....	196
Figura 28 – Texto mais complexo por SIMS do livro <i>Elementary</i>	199
Figura 29 – Texto mais simples por CR do livro <i>Elementary</i>	202
Figura 30 – Texto do livro <i>Pre-intermediate</i> com valor positivo de FPC	222
Figura 31 – Texto do livro <i>Pre-intermediate</i> com valor positivo indicativo de maior complexidade pela FPC	223
Figura 32 – Texto do livro <i>Pre-intermediate</i> com valor de alta acessibilidade por NT	242
Figura 33 – Trecho de texto do livro <i>Pre-intermediate</i> com valor de baixa acessibilidade por NT	244

Figura 34 – Árvores de análise sintática do texto P2	250
Figura 35 – Passos 1 e 2 para o uso do Coh-Metrix	334
Figura 36 – Passos 3 e 4 para o uso do Coh-Metrix	334
Figura 37 – Ilustração de resultados do Coh-Metrix.....	335
Figura 38 – Ilustração de download de métricas do Coh-Metrix	335

LISTA DE QUADROS

Quadro 1 - Três estilos de comunicação para diferentes competências de leitura.....	36
Quadro 2 - Avaliação de competência de leitura de adultos da OECD – resumo.....	42
Quadro 3 - Descrição do CEFR de habilidades globais de leitura	46
Quadro 4 - Descrição do CEFR de habilidades da leitura para informação.....	47
Quadro 5 - Fatores que potencializam a complexidade de textos	52
Quadro 6 – Resumo dos livros <i>English File 4th edition</i>	66
Quadro 7 - Referências de livro e página do <i>corpus</i> de estudo	67
Quadro 8 - Textos do <i>corpus progressivo</i>	69
Quadro 9 - Descrição dos textos simples do <i>corpus</i> comparativo	76
Quadro 10 - Descrição dos textos complexos do <i>corpus</i> comparativo	80
Quadro 11 - Métricas utilizadas para descrição e análise dos textos e de reescritas.....	85
Quadro 12 - Interpretação dos valores de NT para estimar complexidade.....	95
Quadro 13 - Interpretação dos valores de SS para estimar complexidade.....	97
Quadro 14 - Interpretação dos valores de SIMS para estimar complexidade.....	101
Quadro 15 - Interpretação dos valores de CR para estimar complexidade	103
Quadro 16 - Resultados do Índice Flesch.	103
Quadro 17 - Adaptação dos valores de IF para estimar complexidade	105
Quadro 18 - Interpretação dos valores de PS para estimar complexidade.....	107
Quadro 19 - Interpretação dos valores de IA para estimar complexidade	109
Quadro 20 - Interpretação dos valores de FPC para estimar complexidade	111
Quadro 21 - Interpretação dos valores de CLP para estimar complexidade.....	115
Quadro 22 - Interpretação dos valores de TT para estimar complexidade	117
Quadro 23 - Resumo das métricas utilizadas na tese.....	119
Quadro 24 - Valores das métricas do <i>corpus</i> progressivo	123
Quadro 25 - Valores das métricas de sintaxe dos textos simples do <i>corpus</i> comparativo ..	137
Quadro 26 - Valores das métricas dos textos complexos do <i>corpus</i> comparativo.....	141
Quadro 27 - Valores das métricas de léxico e semântica dos textos do livro <i>Elementary</i> ...	159
Quadro 28 - Textos do livro <i>Elementary</i> com exemplos de polissemia.....	163
Quadro 29 - <i>Types</i> e <i>Tokens</i> dos textos do livro <i>Elementary</i>	178
Quadro 30 - Palavras novas introduzidas por cada texto do livro <i>Elementary</i>	180
Quadro 31 - Trechos de textos do livro <i>Elementary</i> para ilustrar resultados de TT	181
Quadro 32 - Comparativo de métricas lexicais-semânticas entre os textos do <i>corpus</i> comparativo e os textos do livro <i>Elementary</i>	186
Quadro 33 - Valores das métricas de sintaxe dos textos do livro <i>Elementary</i>	190
Quadro 34 – Trechos de exemplo de textos com resultados de NT positivos do livro <i>Elementary</i>	192
Quadro 35 – Valores relacionados à extensão de textos do <i>Elementary</i> destacados por IF	207
Quadro 36 – Trechos de textos do <i>Elementary</i> destacados por IF	207

Quadro 37 - Comparativo de métricas de sintaxe entre os textos do <i>corpus</i> comparativo e os textos do livro <i>Elementary</i>	209
Quadro 38 - Valores das métricas de semântica e léxico dos textos do livro <i>Pre-intermediate</i>	212
Quadro 39 - Comparativo dos valores de PS entre os textos E e P	214
Quadro 40 – Palavras dos textos do livro <i>Pre-intermediate</i> que se destacam por PS.....	214
Quadro 41 – Trechos de textos do livro <i>Pre-intermediate</i> que se destacam positivamente por IA	218
Quadro 42 – Trechos de textos do livro <i>Pre-intermediate</i> com valores médios de IA	219
Quadro 43 – Trechos de textos do livro <i>Pre-intermediate</i> com valores médios de FPC.....	224
Quadro 44 – Quantidade de <i>types</i> e de <i>tokens</i> dos textos do livro <i>Pre-Intermediate</i>	230
Quadro 45 – Quantidade de palavras novas/únicas em cada texto do livro <i>Pre-intermediate</i>	231
Quadro 46 – Trechos de textos do livro <i>Pre-intermediate</i> com valores indicativos de alta complexidade por TT	233
Quadro 47 – Trechos de textos do livro <i>Pre-intermediate</i> com valores indicativos de baixa complexidade por TT	235
Quadro 48 - Comparativo de métricas de léxico e semântica entre os textos do <i>corpus</i> comparativo e os textos do livro <i>Pre-intermediate</i>	236
Quadro 49 - Valores das métricas de sintaxe dos textos do livro <i>Pre-intermediate</i>	240
Quadro 50 - Trechos de textos do livro <i>Pre-intermediate</i> com valores altos de NT	242
Quadro 51 – Trechos de textos do livro <i>Pre-intermediate</i> com valores médios de NT	244
Quadro 52 – Trechos de textos do livro <i>Pre-intermediate</i> com valores de alta complexidade por SS.....	247
Quadro 53 – Trechos de textos do livro <i>Pre-intermediate</i> com valores de alta complexidade por SIMS.....	250
Quadro 54 – Trechos de textos do livro <i>Pre-intermediate</i> com valores de alta complexidade por CR	254
Quadro 55 – Trechos de textos do livro <i>Pre-intermediate</i> com valores de alta complexidade por IF	258
Quadro 56 - Comparativo de métricas de sintaxe entre os textos do <i>corpus</i> comparativo e os textos do livro <i>Pre-intermediate</i>	261
Quadro 57 - Valores das métricas de semântica e léxico dos textos do livro <i>Intermediate</i>	265
Quadro 58 - Valores das métricas de sintaxe dos textos do livro <i>Intermediate</i>	271
Quadro 59 - Resumo dos valores e indicativos das linhas na relação métrica-complexidade	278
Quadro 60 - Explicação dos quadros de análise de progressão.....	279
Quadro 61 – Análise de progressão de complexidade da métrica PS	280
Quadro 62 – Análise de progressão de complexidade da métrica IA	282
Quadro 63 – Análise de progressão de complexidade da métrica FPC.....	284
Quadro 64 – Análise de progressão de complexidade da métrica CLP	286
Quadro 65 – Análise de progressão de complexidade da métrica TT	287
Quadro 66 – Análise de progressão de complexidade da métrica NT	290

Quadro 67 – Análise de progressão de complexidade da métrica SS	291
Quadro 68 – Análise de progressão de complexidade da métrica SIMS	293
Quadro 69 – Análise de progressão de complexidade da métrica CR.....	294
Quadro 70 – Análise de progressão de complexidade da métrica IF.....	296
Quadro 71 - Exemplo de apresentação de estratégia de reescrita	300
Quadro 72 – Quadro de estratégias	302
Quadro 73 – Sugestões de estratégias para o texto E1.....	303
Quadro 74 – Sugestões de estratégias para o texto E2.....	304
Quadro 75 – Sugestões de estratégias para o texto E3.....	305
Quadro 76 – Sugestões de estratégias para o texto E5.....	305
Quadro 77 – Sugestões de estratégias para o texto E6.....	307
Quadro 78 – Sugestões de estratégias para o texto E8.....	308
Quadro 79 – Sugestões de estratégias para o texto E14.....	309
Quadro 80 – Sugestões de estratégias para o texto E15.....	309
Quadro 81 – Sugestões de estratégias para o texto E16.....	310
Quadro 82 – Sugestões de estratégias para o texto E17.....	311
Quadro 83 – Sugestões de estratégias para o texto E18.....	312
Quadro 84 – Sugestões de estratégias para o texto E19.....	313
Quadro 85 – Sugestões de estratégias para o texto E20.....	314
Quadro 86 – Sugestões de estratégias para o texto P1.....	316
Quadro 87 – Sugestões de estratégias para o texto P2.....	317
Quadro 88 – Sugestões de estratégias para o texto P3.....	318
Quadro 89 – Sugestões de estratégias para o texto P4.....	318
Quadro 90 – Sugestões de estratégias para o texto P5.....	319
Quadro 91 – Sugestões de estratégias para o texto P6.....	319
Quadro 92 – Sugestões de estratégias para o texto P7.....	320
Quadro 93 – Sugestões de estratégias para o texto P8.....	321
Quadro 94 – Sugestões de estratégias para o texto P10.....	322
Quadro 95 – Sugestões de estratégias para o texto P11.....	322
Quadro 96 – Sugestões de estratégias para o texto P12.....	323
Quadro 97 – Sugestões de estratégias para o texto P13.....	324
Quadro 98 – Sugestões de estratégias para o texto P14.....	326
Quadro 99 – Sugestões de estratégias para o texto P15.....	326
Quadro 100 – Sugestões de estratégias para o texto P17.....	327
Quadro 101 – Sugestões de estratégias para o texto P18.....	329
Quadro 102 – Sugestões de estratégias para o texto P19.....	330
Quadro 103 – Sugestões de estratégias para o texto P22.....	330
Quadro 104 – Referências de complexidade de NT	337
Quadro 105 – Exemplos de NT	337

Quadro 106 – Estratégias para aumento de acessibilidade de NT	338
Quadro 107 – Referências de complexidade de SS	341
Quadro 108 – Estratégias para aumento de acessibilidade de SS	341
Quadro 109 – Referências de complexidade de SIMS	342
Quadro 110 – Estratégias para aumento de acessibilidade de SIMS	343
Quadro 111 – Referências de complexidade de CR.....	345
Quadro 112 – Estratégias para aumento de acessibilidade de CR.....	345
Quadro 113 - Resultados do Índice Flesch	346
Quadro 114 – Estratégias para aumento de acessibilidade de IF.....	347
Quadro 115 – Referências de complexidade de PS	348
Quadro 116 – Estratégias para aumento de acessibilidade de PS	348
Quadro 117 – Referências de complexidade de IA.....	350
Quadro 118 – Estratégias para aumento de acessibilidade de IA	351
Quadro 119 – Referências de complexidade de FPC.....	353
Quadro 120 – Estratégias para aumento de acessibilidade de FPC.....	353
Quadro 121 – Referências de complexidade de CLP	355
Quadro 122 – Estratégias para aumento de acessibilidade de CLP.....	356
Quadro 123 – Referências de complexidade de TT.....	357
Quadro 124 – Estratégias para aumento de acessibilidade de TT	358

LISTA DE GRÁFICOS

Gráfico 1 - Comportamento de NT dos textos do <i>corpus</i> progressivo	124
Gráfico 2 - Comportamento de SS dos textos do <i>corpus</i> progressivo	126
Gráfico 3 - Comportamento de SIMS dos textos do <i>corpus</i> progressivo	127
Gráfico 4 - Comportamento de CR dos textos do <i>corpus</i> progressivo	128
Gráfico 5 - Comportamento de IF dos textos do <i>corpus</i> progressivo	130
Gráfico 6 - Comportamento de PS dos textos do <i>corpus</i> progressivo	131
Gráfico 7 - Comportamento de IA dos textos do <i>corpus</i> progressivo	132
Gráfico 8 - Comportamento de FPC dos textos do <i>corpus</i> progressivo	133
Gráfico 9 - Comportamento de SS dos textos do <i>corpus</i> progressivo	134
Gráfico 10 - Comportamento de SS dos textos do <i>corpus</i> progressivo	135
Gráfico 11 - Classificação percentual de complexidade de PS do livro <i>Elementary</i>	162
Gráfico 12 - Classificação percentual de complexidade de IA do livro <i>Elementary</i>	166
Gráfico 13 - Classificação percentual de complexidade de FPC do livro <i>Elementary</i>	170
Gráfico 14 - Classificação percentual de complexidade de CLP do livro <i>Elementary</i>	173
Gráfico 15 - Classificação percentual de complexidade de TT do livro <i>Elementary</i>	177
Gráfico 16 - <i>Types</i> e <i>Tokens</i> dos textos do livro <i>Elementary</i>	178
Gráfico 17 - Quantidade de palavras novas introduzidas por texto do livro <i>Elementary</i>	181
Gráfico 18 - Classificação percentual de complexidade de NT do livro <i>Elementary</i>	191
Gráfico 19 - Classificação percentual de complexidade de SS do livro <i>Elementary</i>	194
Gráfico 20 - Classificação percentual de complexidade de SIMS do livro <i>Elementary</i>	198
Gráfico 21 - Classificação percentual de complexidade de CR do livro <i>Elementary</i>	201
Gráfico 22 - Classificação percentual de complexidade de IF do livro <i>Elementary</i>	206
Gráfico 23 - Classificação percentual de complexidade de PS do livro <i>Pre-intermediate</i>	213
Gráfico 24 - Classificação percentual de complexidade de IA do livro <i>Pre-intermediate</i>	217
Gráfico 25 - Classificação percentual de complexidade de FPC do livro <i>Pre-intermediate</i>	221
Gráfico 26 - Classificação percentual de complexidade de CLP do livro <i>Pre-intermediate</i>	226
Gráfico 27 - Classificação percentual de complexidade de TT do livro <i>Pre-intermediate</i>	229
Gráfico 28 – <i>Types</i> e <i>tokens</i> dos textos do livro <i>Pre-intermediate</i> em gráfico-linha	231
Gráfico 29 - Classificação percentual de complexidade de NT do livro <i>Pre-intermediate</i>	241
Gráfico 30 - Classificação percentual de complexidade de SS do livro <i>Pre-intermediate</i>	246
Gráfico 31 - Classificação percentual de complexidade de SIMS do livro <i>Pre-intermediate</i>	249
Gráfico 32 - Classificação percentual de complexidade de CR do livro <i>Pre-intermediate</i>	253
Gráfico 33 - Classificação percentual de complexidade de IF do livro <i>Pre-intermediate</i>	258
Gráfico 34 – Progressão de complexidade de PS	280
Gráfico 35 – Progressão de complexidade de IA	282
Gráfico 36 – Progressão de complexidade de FPC	284
Gráfico 37 – Progressão de complexidade de CLP	286

Gráfico 38 – Progressão de complexidade de TT	287
Gráfico 39 – Progressão de complexidade de NT.....	290
Gráfico 40 – Progressão de complexidade de SS.....	291
Gráfico 41 – Progressão de complexidade de SIMS.....	293
Gráfico 42 – Progressão de complexidade de CR	294
Gráfico 43 – Progressão de complexidade de IF	296

LISTA DE SIGLAS

AMD	Análises Multimodais
AT	Acessibilidade Textual
ATLD	Acessibilidade Textual Linguodidática
ATT	Acessibilidade Textual e Terminológica
CEFR	<i>Common European Framework of Reference</i>
CiC	Ciência da Computação
CLP	Concretude lexical
CP	<i>Corpus</i> Progressivo
CLP	Concretude Lexical (Percentual)
CR	Coesão Referencial
CT	Complexidade Textual
FPC	Familiaridade com Palavras de Conteúdo
IA	Idade de Aquisição
IBPAD	Instituto Brasileiro de Pesquisa e Análise de Dados
IF	Índice Flesch
ILA	Inglês como Língua Adicional
LC	Linguística de <i>Corpus</i>
LIC	Língua Inglesa (Complexos)
LIS	Língua Inglesa (Simples)
NT	Narratividade
OECD	<i>Organization for Economic Co-operation and Development</i>
PIAAC	<i>Programme for the International Assessment of Adult Competencies</i>
PLN	Processamento de Linguagem Natural
PS	Polissemia
SIMS	Similaridade Sintática
SS	Simplicidade Sintática
ST	Simplificação Textual
SVO	Sujeito Verbo Objeto
TT	<i>Type-Token</i>
TTR	Relação <i>Type-Token</i>

SUMÁRIO

1 INTRODUÇÃO	23
1.1 ORIGEM DO TRABALHO	23
1.2 QUESTÕES NORTEADORAS	26
1.3 OBJETIVOS	27
1.4 HIPÓTESES	28
1.5 ORGANIZAÇÃO DA TESE.....	29
2 COMPLEXIDADE, ACESSIBILIDADE E SIMPLIFICAÇÃO TEXTUAL	32
2.1 DEFINIÇÕES E ESTUDOS.....	32
2.2 ACESSIBILIDADE TEXTUAL LINGUODIDÁTICA	36
2.3 ORIENTAÇÕES DO CEFR PARA LEITURA EM ILA	44
2.4 ANÁLISES DE ACESSIBILIDADE TEXTUAL	49
2.5 CARACTERÍSTICAS POTENCIALIZADORAS DE COMPLEXIDADE	52
3 EDUCAÇÃO 3.0, 4.0 E 5.0: CUSTOMIZAÇÃO E PERSONALIZAÇÃO COMO DIRETRIZES DA ACESSIBILIDADE TEXTUAL	53
4 PROCESSAMENTO DA LINGUAGEM NATURAL E LINGÜÍSTICA DE <i>CORPUS</i> : UM DIÁLOGO	60
4.1 PROCESSAMENTO DA LINGUAGEM NATURAL	60
4.2 LINGÜÍSTICA DE <i>CORPUS</i>	61
4.3 UTILIZAÇÃO NA PESQUISA	63
5 <i>CORPORA</i> DA PESQUISA	63
5.1 <i>CORPUS</i> DE ESTUDO	64
5.2 <i>CORPUS</i> PROGRESSIVO.....	68
5.3 <i>CORPUS</i> COMPARATIVO.....	74
6 FERRAMENTA E MÉTRICAS	81
6.1 COH-METRIX.....	82
6.2 MÉTRICAS DE LEITURABILIDADE	84
6.2.1 MÉTRICAS COMO SISTEMAS CLASSIFICATÓRIOS DE POTENCIAIS DE COMPLEXIDADE	87
6.2.2 NARRATIVIDADE.....	93
6.2.3 SIMPLICIDADE SINTÁTICA.....	95
6.2.4 SIMILARIDADE SINTÁTICA.....	97
6.2.5 COESÃO REFERENCIAL	102
6.2.6 ÍNDICE FLESCH	103
6.2.7 POLISSEMIA.....	105
6.2.8 IDADE DE AQUISIÇÃO.....	107
6.2.9 FAMILIARIDADE COM PALAVRAS DE CONTEÚDO	110
6.2.10 CONCRETUDE LEXICAL.....	112
6.2.11 RELAÇÃO TYPE-TOKEN.....	116

6.2.12 RESUMO	118
7 PROCEDIMENTOS METODOLÓGICOS.....	120
8 RESULTADOS	121
8.1 ANÁLISE SEMIAUTOMÁTICA DOS <i>CORPORA</i> PROGRESSIVO E COMPARATIVO	122
8.1.2 Análise dos resultados das métricas dos textos do corpus progressivo	123
8.1.2.1 Comportamento da métrica Narratividade (NT).....	124
8.1.2.2 Comportamento da métrica Simplicidade Sintática (SS).....	125
8.1.2.3 Comportamento da métrica Similaridade Sintática (SIMS)	126
8.1.2.4 Comportamento da métrica Coesão Referencial (CR).....	128
8.1.2.5 Comportamento da métrica Índice Flesch (IF)	129
8.1.2.6 Comportamento da métrica Polissemia (PS).....	130
8.1.2.7 Comportamento da métrica Idade de Aquisição.....	132
8.1.2.8 Comportamento da métrica Familiaridade com Palavras de Conteúdo (FPC)	133
8.1.2.9 Comportamento da métrica Concretude Lexical (CLP).....	134
8.1.2.10 Comportamento da métrica Relação <i>Type-Token</i> (TT).....	135
8.2 ANÁLISE DOS RESULTADOS DAS MÉTRICAS DOS TEXTOS SIMPLES DO <i>CORPUS</i> COMPARATIVO.....	137
8.3 ANÁLISE DOS RESULTADOS DAS MÉTRICAS DOS TEXTOS COMPLEXOS DO <i>CORPUS</i> COMPARATIVO.....	141
8.4 DESCRIÇÃO LINGUÍSTICA E PEDAGÓGICA DO <i>CORPUS</i> DE ESTUDO	145
8.4.1 DESCRIÇÃO DOS TEXTOS DO LIVRO ELEMENTARY.....	146
8.4.2 DESCRIÇÃO DOS TEXTOS DO LIVRO PRE-INTERMEDIATE.....	151
8.4.3 DESCRIÇÃO DOS TEXTOS DO LIVRO INTERMEDIATE	155
8.5 ANÁLISE SEMIAUTOMÁTICA DO <i>CORPUS</i> DE ESTUDO	159
8.5.1 RESULTADOS MÉTRICOS DOS TEXTOS DO LIVRO DE NÍVEL ELEMENTARY	159
8.5.1.1 Análise da complexidade lexical-semântica dos textos E.....	161
8.5.1.1.1 Análise da métrica PS dos textos E.....	161
8.5.1.1.2 Análise da métrica IA dos textos E	166
8.5.1.1.3 Análise da métrica FPC dos textos E	169
8.5.1.1.4 Análise da métrica CLP dos textos E	172
8.5.1.1.5 Análise da métrica TT dos textos E	176
8.5.1.1.6 Análise contrastiva das métricas de léxico e semântica de E com os de LIS e LIC	184
8.5.1.2 Análise da complexidade sintática dos textos E.....	189
8.5.1.2.1 Análise da métrica NT dos textos E.....	190
8.5.1.2.2 Análise da métrica SS dos textos E.....	194
8.5.1.2.3 Análise da métrica SIMS dos textos E.....	198
8.5.1.2.4 Análise da métrica CR dos textos E	200
8.5.1.2.5 Análise da métrica IF dos textos E	205

8.5.1.2.7	Análise contrastiva das métricas de sintaxe de E com os de LIS e LIC	208
8.5.2	RESULTADOS MÉTRICOS DOS TEXTOS DO LIVRO DE NÍVEL PRE-INTERMEDIATE	212
8.5.2.1	Análise da complexidade lexical-semântica dos textos P	212
8.5.2.1.1	Análise da métrica PS dos textos P	213
8.5.2.1.2	Análise da métrica IA dos textos P	216
8.5.2.1.3	Análise da métrica FPC dos textos P	220
8.5.2.1.4	Análise da métrica CLP dos textos P	225
8.5.2.1.5	Análise da métrica TT dos textos P	229
8.5.2.1.6	Análise contrastiva das métricas de léxico e semântica de P com os de LIS e LIC	236
8.5.2.2	Análise da complexidade sintática dos textos P	239
8.5.2.2.1	Análise da métrica NT dos textos P	240
8.5.2.2.2	Análise da métrica SS dos textos P	246
8.5.2.2.3	Análise da métrica SIMS dos textos P	249
8.5.2.2.4	Análise da métrica CR dos textos P	252
8.5.2.2.5	Análise da métrica IF dos textos P	257
8.5.2.2.6	Análise contrastiva das métricas de sintaxe de P com os de LIS e LIC	261
8.5.3	RESULTADOS MÉTRICOS DOS TEXTOS DO LIVRO DE NÍVEL INTERMEDIATE	264
8.5.3.1	Análise prévia da complexidade lexical-semântica dos textos I	265
8.5.3.1.1	Análise prévia da PS dos textos I	266
8.5.3.1.2	Análise prévia da IA dos textos I	267
8.5.3.1.3	Análise prévia da FPC dos textos I	268
8.5.3.1.4	Análise prévia da CLP dos textos I	269
8.5.3.1.5	Análise prévia da TT dos textos I	269
8.5.3.2	Análise prévia da complexidade sintática dos textos I	270
8.5.3.2.1	Análise prévia da NT dos textos I	271
8.5.3.2.2	Análise prévia da SS dos textos I	273
8.5.3.2.3	Análise prévia da SIMS dos textos I	273
8.5.3.2.4	Análise prévia da CR dos textos I	275
8.5.3.2.5	Análise prévia da IF dos textos I	276
8.6	ANÁLISE DE PROGRESSÃO DE COMPLEXIDADE	276
9	SUGESTÕES DE ESTRATÉGIAS DE ACESSIBILIDADE PARA OS TEXTOS DO CORPUS DE ESTUDO	299
9.1	SUGESTÕES DE ESTRATÉGIAS DE ACESSIBILIDADE PARA OS TEXTOS DO LIVRO <i>ELEMENTARY</i>	303
9.2	SUGESTÕES DE ESTRATÉGIAS DE ACESSIBILIDADE PARA OS TEXTOS DO LIVRO <i>PRE-INTERMEDIATE</i>	315
10	ORIENTAÇÕES PARA ACESSIBILIDADE TEXTUAL LINGUODIDÁTICA	332
10.1	COMO UTILIZAR O COH-MATRIX E COLETAR MÉTRICAS TEXTUAIS	333

10.2 COMO UTILIZAR A NARRATIVIDADE (NT) PARA ESTIMAR COMPLEXIDADE ..	336
10.3 COMO UTILIZAR A SIMPLICIDADE SINTÁTICA (SS) PARA ESTIMAR COMPLEXIDADE	340
10.4 COMO UTILIZAR A SIMILARIDADE SINTÁTICA (SIMS) PARA ESTIMAR COMPLEXIDADE	342
10.5 COMO UTILIZAR A COESÃO REFERENCIAL (CR) PARA ESTIMAR COMPLEXIDADE	344
10.6 COMO UTILIZAR O ÍNDICE FLESCH (IF) PARA ESTIMAR COMPLEXIDADE	346
10.7 COMO UTILIZAR A POLISSEMIA (PS) PARA ESTIMAR COMPLEXIDADE.....	348
10.8 COMO UTILIZAR A IDADE DE AQUISIÇÃO (IA) PARA ESTIMAR COMPLEXIDADE	350
10.9 COMO UTILIZAR A FAMILIARIDADE COM PALAVRAS DE CONTEÚDO (FPC) PARA ESTIMAR COMPLEXIDADE	352
10.10 COMO UTILIZAR A CONCRETUDE LEXICAL (CLP) PARA ESTIMAR COMPLEXIDADE	354
10.11 COMO UTILIZAR A RELAÇÃO TYPE-TOKEN (TT) PARA ESTIMAR COMPLEXIDADE	356
11 RESUMO DA TESE EM LINGUAGEM SIMPLES.....	360
12 CONSIDERAÇÕES FINAIS	364
REFERÊNCIAS BIBLIOGRÁFICAS	371
APÊNDICES.....	376
QUADROS ORIGINAIS DO CEFR.....	376
LISTAS DE PALAVRAS	378
TEXTOS SIMPLES DO <i>CORPUS</i> COMPARATIVO	431
LIS1	431
LIS2	432
LIS3	433
LIS4	435
LIS5	436
LIS6	437
LIS7	438
LIS 8	439
LIS 9	440
LIS 10	441
LIS11	443
LIS12	444
LIS13	445
LIS14	446
LIS15	447
LIS16	448
LIS17	449

LIS18	450
LIS19	451
LIS20	452
TEXTOS COMPLEXOS DO <i>CORPUS</i> COMPARATIVO	453
LIC1	453
LIC2	454
LIC3	456
LIC5	459
LIC6	460
LIC7	462
LIC8	463
LIC9	464
LIC10	465
LIC11	466
LIC12	468
LIC 13	470
LIC 14	472
LIC 15	474
LIC 16	477
LIC 17	479
LIC 18	481
LIC 19	483
LIC 20	485

1 INTRODUÇÃO

1.1 ORIGEM DO TRABALHO

Acessibilidade e complexidade textual (AT e CT, respectivamente) são assuntos que há muito tempo vêm sendo discutidos no meio acadêmico. Desde os anos 40, por exemplo, pesquisas como as de Rudolph Flesch (1951) se concentram em estudar características que potencializam a complexidade de textos para diferentes tipos de leitores, especialmente adultos com baixa escolaridade. Outros autores e organizações, como DuBay (2004) e o projeto Plain Language¹ do governo dos Estados Unidos, estudam diferentes gêneros e tipologias textuais na busca por estratégias de simplificação que tornem a leitura mais acessível para públicos heterogêneos.

Ainda que AT e CT não sejam assuntos novos, o número de pesquisas e publicações no campo da Linguística – principalmente trabalhos que orientem sobre características e estratégias de acessibilidade textual de forma prática e aplicada – é limitado. As principais temáticas desses trabalhos são, primordialmente, voltadas à acessibilidade de temas da Medicina e do Direito; uma vez que a linguagem dessas áreas tende a ser altamente terminológica e os textos, em geral, complexos (principalmente por serem redigidos por especialistas com hábito de escrever para seus pares). Logo, quando um leitor leigo procura por materiais sobre, por exemplo, algum tópico de saúde, ainda que encontre uma grande quantidade de informativos gratuitos, geralmente tem dificuldade de compreender as informações apresentadas, como demonstraram Silva (2018) com textos sobre a Psiquiatria supostamente publicados para pessoas leigas e Motta (2021) com materiais do Direito.

Textos, como sabemos, são estruturas com múltiplos níveis e características. Cada gênero textual configura e estipula formas de redação e escolhas lexicais específicas. Logo, ao abordarmos complexidade textual, é evidente que a verificação de características dificultadoras está relacionada a diversos fatores, desde traços estritamente ligados ao texto (como léxico, semântica, sintaxe etc.) até características do leitor para o qual um material será destinado. Sendo assim, as características que potencializam a complexidade de um texto e as alternativas para torná-lo mais

¹ <https://www.plainlanguage.gov/>

acessível também são múltiplas – mas delimitadas pelo tipo de texto e pelo perfil de leitor. Pesquisadores da área, como DuBay (2004), Finatto (2018) e Silva (2018) destacam a importância de observar potenciais de complexidade sempre tendo em mente determinados tipos de leitor – uma vez que um texto não é complexo por si só, mas sempre em relação a quem se apresenta. Além disso, os autores supracitados também enfatizam que cada gênero textual e assunto apresenta peculiaridades e, conseqüentemente, diferem nos traços que os tornam mais ou menos complexos.

Inobstante da erudição no que concerne a leitura em língua materna de cada leitor, é indubitável a inexequibilidade de agnição e exegese ante composições abaladamente terminológicas e abstrusas – como neste parágrafo, que embora tente comunicar algo simples, foi escrito com léxico um tanto quanto complexo e, conseqüentemente, dificultou a leitura (ainda que voltado para leitores com titulações em estudos de língua e linguagem). Em palavras mais simples: independentemente do nível de proficiência de um leitor em sua língua materna, é certo que textos complexos e com muitos termos específicos de determinadas áreas podem prejudicar a compreensão das mensagens.

Se existem inúmeras pesquisas que se propõem a estudar elementos que tornam uma leitura complexa em língua materna, podemos concluir que observar tais características em textos escritos em língua adicional para o aprendizado dessas línguas por falantes não-nativos é essencial. Interpretação de leitura é não só uma importante habilidade linguística a ser trabalhada enquanto se busca aprender um determinado idioma, mas também uma forma de ampliar vocabulário, de trabalhar com estruturas gramaticais e de observar aspectos culturais relacionados às línguas. Dessa maneira, textos não são somente para finalidades para interpretação, mas insumos de aprendizagem.

Atualmente, existem [raras] pesquisas que têm sido conduzidas por meio de análises semiautomáticas com o apoio de ferramentas de Processamento de Linguagem Natural (PLN, sobre o qual falaremos adiante) que permitem a coleta de métricas textuais cujos resultados podem ser utilizados como indícios de complexidade. Ao mesmo tempo, nem todos os trabalhos que estimam complexidade comparam os resultados interpretados pelos índices coletados por meio de *software* com análises e interpretações de linguistas especializados no assunto; fato que, conseqüentemente, reforça que tais estudos apontam a **potencialidade** de

complexidade (uma vez que a corroboração sobre um texto ser complexo ou acessível só é dada por meio da leitura de leitores humanos reais).

Este trabalho se originou da experiência com análises de complexidade e acessibilidade textual e com a observação de determinados aspectos de textos didáticos escritos em língua inglesa que dificultam a interpretação (e consequentemente a manutenção linguística²) por alunos brasileiros de inglês como língua adicional (especialmente em níveis iniciais). A complexidade destes materiais é relacionada, primordialmente, ao léxico e à semântica e em segundo lugar, à sintaxe. Os traços de complexidade, além de dificultarem a compreensão dos textos, são fatores de desmotivação para o aprendizado da língua.

Ao realizarmos o Estado da Arte³ (mapeamento de publicações acadêmicas) sobre o tópico desta tese, não conseguimos localizar nenhuma pesquisa com a proposta apresentada aqui: analisar métricas textuais de livros didáticos de inglês como língua adicional para estimar a complexidade desses textos para os níveis propostos e, posteriormente, sugerir estratégias de acessibilidade textual. Conforme detalharemos posteriormente, existem algumas pesquisas que utilizam métricas textuais com o objetivo de estudar complexidade - conforme Finatto (2018) resume - mas todas utilizam como *corpora* textos de especialidades acadêmicas (como Direito e Medicina).

Uma vez que o aprendizado de uma língua adicional se dá por níveis, e tendo em mente que, para cada tipo de leitor, devemos observar e estudar estratégias de acessibilidade específicas, buscamos estudar textos didáticos escritos em língua inglesa para estudantes brasileiros com a finalidade de levantar dados sobre elementos que tenham potencial de impedir ou dificultar o aprendizado da língua para esse público. Como professores de línguas adicionais, temos o conhecimento de que as línguas maternas dos alunos têm grande influência nas facilidades e dificuldades das línguas adicionais estudadas (uma aluna chinesa, por exemplo, tem facilidades e dificuldades diferentes ao aprender inglês como língua adicional do que uma aluna brasileira que, por sua vez, também se difere de uma aluna alemã e assim por diante).

² Nesta tese, utilizaremos o conceito “manutenção linguística” para descrever os resultados e avanços de estudantes no aprendizado de uma língua.

³ Utilizamos plataformas de publicações acadêmicas como *Scielo*, *Google Scholar*, *Connected Papers*, *Research Gate* etc. e pesquisamos por palavras-chave como “complexidade textual + ensino”, “complexidade textual + língua adicional”, entre outras.

Tendo conhecimento e experiência com ferramentas que oferecem métricas para interpretação de complexidade textual, nosso objetivo principal é analisar índices textuais – coletados por meio de um *software* de PLN (Coh-Metrix) – que nos permitam observar quais níveis e características dos textos componentes do nosso *corpus* (o qual descreveremos mais adiante) podem ser considerados potencialmente complexos para o ensino/aprendizado de inglês como língua adicional com nosso perfil de leitores (estudantes de inglês como língua adicional adultos, sobre os quais apresentaremos mais detalhes posteriormente). A partir disso, apresentamos orientações sobre estratégias de acessibilidade que podem ser apropriadas para tornar os textos mais acessíveis - e, simultaneamente, permitir avanço linguístico – e, em outros momentos, destacamos pontos que merecem atenção dos professores que trabalham com customização e personalização de ensino (seguindo as orientações de metodologias educacionais modernas, como a Educação 3.0, 4.0 e 5.0).

Temos a convicção de que a aprendizagem de uma língua adicional não acontece somente por meio da habilidade de leitura. Não excluimos, com esta pesquisa, a importância de outras habilidades linguísticas e de diferentes materiais (como áudio, imagens etc.) que são utilizados para fins didáticos. Ao mesmo tempo, acreditamos que nossa investigação pode colaborar para uma utilização mais eficiente de textos no contexto de ensino-aprendizagem.

Esperamos, também, que nossa pesquisa possa ilustrar o uso da ferramenta adotada (Coh-Metrix) de forma didática, orientar sobre a utilização de dados métricos para estudos de complexidade textual e oferecer informações para autores/editoras de livros didáticos de línguas adicionais e para profissionais de distintas áreas que se preocupem em apresentar textos acessíveis.

Na próxima seção, apresentamos as questões que nortearão o desenvolvimento da nossa pesquisa.

1.2 QUESTÕES NORTEADORAS

As questões norteadoras desta pesquisa são provenientes, conforme já mencionado, da necessidade de determinar quais fatores podem ser considerados potencializadores de complexidade para o ensino de língua inglesa como língua adicional para alunos brasileiros e quais estratégias podem ser eficientes para aprimorar o aprendizado.

Professores de Inglês como Língua Adicional (ILA) aprendem a “intuir”, com o tempo e com a experiência, quais serão as dificuldades relacionadas a cada conteúdo trabalhado e quais pontos dos materiais utilizados tendem a ser difíceis para os alunos. Com base na nossa constatação em sala de aula, de que algumas características dificultam a compreensão e o aprendizado de textos voltados para o ensino de língua inglesa, questionamos se é possível determinar quais fatores pontuais tornam um texto complexo para, principalmente, os níveis iniciais de ensino (a saber: elementar e pré-intermediário) e quais estratégias provenientes de pesquisas sobre acessibilidade textual podem aumentar a acessibilidade aos textos enquanto oferecem insumos suficientes para o aprendizado e para a manutenção linguística.

Sendo assim, as questões específicas que orientam esta pesquisa são:

- a. Os textos didáticos do nosso *corpus*, conforme se apresentam originalmente, têm características que dificultam a compreensão e o aprendizado do inglês como língua adicional para estudantes universitários brasileiros nos níveis linguísticos⁴ elementar e pré-intermediário?
- b. Que características podem contribuir para a possível complexidade desses textos?
- c. Existem textos dentro de um mesmo nível que se diferem dos outros por apresentar características de alto potencial de complexidade textual?
- d. Quais indicações de estratégias de simplificação podem ser adequadas para cada um desses níveis linguísticos de aprendizado?
- e. Existe uma progressão de complexidade gramatical e lexical (naturalmente esperada no desenvolvimento linguístico) nos textos do *corpus* para garantir a manutenção linguística?

1.3 OBJETIVOS

Tendo por base as questões que orientam a pesquisa, nossos objetivos são:

⁴ Decidimos pelo termo “nível linguístico” por acreditarmos que seja mais abrangente que “nível de proficiência” e por englobar todas as habilidades de uma língua (como leitura, escrita, compreensão auditiva, fala, cultura etc.). Além disso, o termo escolhido é equivalente aos documentos do *Common European Framework of Reference (CEFR)*, que será utilizado como uma das referências desta tese.

- I. Analisar a complexidade textual nos níveis lexical, sintático e semântico de textos em língua inglesa presentes nos livros *Elementary* e *Pre-intermediate* da coleção *English File* (quarta edição) da editora Oxford;
- II. Destacar características potencializadoras de complexidade nos textos;
- III. Observar textos que se sobressaem dos demais dentro de um mesmo nível por apresentarem características de potencial complexidade;
- IV. Selecionar estratégias de simplificação que sejam adequadas para cada nível linguístico (elementar e pré-intermediário) com o objetivo de tornar os textos acessíveis e com insumo para o aprendizado;
- V. Verificar se existe progressão de complexidade no desenvolvimento dos níveis dos textos para garantir manutenção linguística;
- VI. Propor diretrizes para um manual de acessibilidade textual para a redação e/ou edição de textos voltados para o ensino de inglês como língua adicional para brasileiros, destacando textos e níveis textuais que necessitam de atenção.

1.4 HIPÓTESES

As hipóteses levantadas são provenientes da experiência do autor desta tese com os textos trabalhados em sala de aula e com a observação da dificuldade de alguns alunos diante de determinadas características desses materiais. Além disso, essas hipóteses estão baseadas em sua pesquisa contínua sobre Complexidade, Acessibilidade e Simplificação textual que nos indica que:

- I. Embora adaptados para o nível linguístico aos quais são propostos, os textos presentes na quarta edição da coleção *English File* da editora Oxford ainda apresentam características que dificultam a compreensão de leitura e, conseqüentemente, a manutenção linguística para estudantes brasileiros universitários de inglês como língua adicional nos níveis elementar e pré-intermediário;
- II. Os principais fatores de complexidade estão relacionados ao léxico e à semântica e os segundos, à sintaxe – especialmente pelo fato de que alguns aspectos desses níveis textuais são apresentados/ensinados em unidades e/ou livros distantes de onde aparecem (por exemplo, um conteúdo trabalhado no livro *Pre-intermediate* que está constantemente presente em textos do livro *Elementary*);

III. Alguns textos apresentam uma ou mais características que os tornam potencialmente complexos, especialmente tendo em vista onde são apresentados no livro;

IV. Há irregularidade e inconsistência na progressão de complexidade lexical, semântica e sintática dentro de cada nível e entre os diferentes níveis;

V. As estratégias de acessibilidade textual que podem ser eficientes são dependentes do nível e da posição de cada um dos textos nos livros e que, quando aplicadas em coerência com as interpretações de indicações de complexidade das métricas podem otimizar a compreensão leitora e, simultaneamente, a manutenção linguística dos estudantes.

1.5 ORGANIZAÇÃO DA TESE

A tese foi organizada da seguinte forma:

Na parte 1, apresentamos nossos pressupostos teóricos iniciando, primeiramente, por uma revisão literária sobre Complexidade, Acessibilidade e Simplificação Textual, por meio da qual apresentamos estudos sobre as áreas e algumas definições relevantes para o trabalho.

Ainda na mesma temática, apresentamos o termo que cunhamos como Acessibilidade Textual Linguodidática, que é a utilização de metodologias e teorias de Acessibilidade Textual, mas voltados especificamente para análises de complexidade de textos utilizados para o ensino de línguas. Nessa seção, discutimos alguns pontos teóricos fundamentais a respeito de aprendizagem de língua adicional e relacionamos esses conceitos com as práticas de estudos de acessibilidade.

Logo após, discutimos e ilustramos como são geralmente conduzidas as pesquisas de complexidade e acessibilidade textual, apresentando algumas definições sobre uso de ferramentas, sobre multidimensionalidade e sobre a necessidade de orientação com base em leitores.

Após nossos capítulos sobre Complexidade, Acessibilidade e Simplificação Textual, discorreremos sobre metodologias e práticas didáticas contemporâneas: Educação 3.0, 4.0 e 5.0 que – atualmente – são vigentes e constantes em escolas que apresentam resultados positivos de aprendizagem (AZEVEDO, DELGADO e SILVA, 2017; PERNA, DELGADO e SILVA, 2020). Uma vez que esse tipo de

metodologia valoriza a customização, a personalização, o uso de tecnologias e o ensino de inglês como língua adicional, acreditamos que nossa pesquisa possa ser útil para orientar professores de inglês sobre formas de customizar e personalizar textos com a finalidade de garantir aprendizados linguísticos eficientes.

Ainda dentro do contexto de revisão literária, nosso último capítulo teórico apresenta conceitos sobre o Processamento de Linguagem Natural (PLN) e sobre a Linguística de *Corpus* (doravante LC) que, embora lidem com a língua e com o texto de formas diferentes (sendo a primeira de forma mais individual e a segunda de maneira mais estatística), servem, em distintos momentos, como fios-condutores na aplicação prática da nossa pesquisa.

Na parte 2 desta tese, apresentamos a organização que adotamos para a aplicação dos procedimentos metodológicos. Primeiramente, detalhamos nossos *corpora* da pesquisa (sendo um, *corpus* de estudo e o outro, *corpus* comparativo) e, posteriormente, as ferramentas computacionais utilizadas para coleta e manipulação de dados e as métricas elegidas para as finalidades deste estudo.

Ainda na segunda parte, explicamos os passos adotados para que nossa coleta e interpretação de dados seguisse uma lógica coerente.

Na parte 3 e final desta tese, apresentamos e discutimos os resultados das métricas textuais e de suas indicações de potencial de complexidade. Antes da análise das métricas, descrevemos as principais características do nosso *corpus* de estudo, destacando pontos importantes sobre os textos com base em nossa experiência como pesquisadores-linguistas e como professores que utilizam os textos do *corpus* em sala de aula.

Posteriormente, realizamos a interpretação das métricas selecionadas a respeito de seus indícios de complexidade enfatizando textos e trechos que se destacam pelos resultados de cada índice. Ao final da análise de cada grupo de métricas (de léxico e semântica e de sintaxe), comparamos os valores dos textos do *corpus* de estudo com as médias dos textos do *corpus* comparativo.

Após a análise das métricas, apresentamos uma discussão e dados sobre possíveis progressões de complexidade entre os textos de diferentes níveis. Uma vez que nosso estudo está voltado para o ensino, e tendo em mente que os textos apresentados deveriam apresentar determinada progressão em dificuldade para garantir a manutenção linguística, propomo-nos a comparar os valores das métricas a fim de relatar se o avanço que é esperado acontece.

Posteriormente, com base nas análises dos dados, apresentamos orientações para adequações de textos do *corpus* de estudo que tenham apresentado traços de alta complexidade a fim de colaborar com professores que busquem personalização e customização desses materiais quando o perfil de alunos exigir.

No último capítulo desta tese, desenvolvemos um manual sucinto de como utilizar as métricas textuais que adotamos na pesquisa para estimar complexidade. Para cada métrica, apresentamos características que impactam nos seus valores e algumas alternativas para potencializar a acessibilidade para as indicações de cada um dos índices.

No capítulo a seguir, apresentamos pressupostos teóricos que sustentam e orientam nossa pesquisa, iniciando por definições e apontamentos sobre complexidade, acessibilidade e simplificação textual.

PARTE 1 – PRESSUPOSTOS TEÓRICOS

2 COMPLEXIDADE, ACESSIBILIDADE E SIMPLIFICAÇÃO TEXTUAL

2.1 DEFINIÇÕES E ESTUDOS

Há muitas décadas, principalmente em países onde o inglês é a língua materna, existe uma preocupação com o *plain language* que, resumidamente, é o tipo de língua utilizada de forma clara para otimizar textos e possibilitar que qualquer pessoa (mesmo com escolaridade limitada ou neuroatípicas) tenha acesso a informações de forma eficiente. Com esse objetivo, existem diversos estudos acadêmicos e projetos (inclusive governamentais) que investigam as características que perfazem a linguagem simples e que oferecem orientações para escrita acessível. Um exemplo é o *Plain Language*⁵ do governo estadunidense, que é um programa formado por trabalhadores federais de diversas agências e especialidades que estudam e instruem a respeito de comunicação clara e efetiva em textos governamentais.

O programa supracitado, além de oferecer guias de orientação para escrita acessível, também disponibiliza cursos e oficinas a respeito do assunto. Seus maiores objetivos são a conscientização e instrução sobre produções textuais que sejam claras para o maior número possível de leitores – tendo consciência de que a população de qualquer país é heterogênea em relação à proficiência linguística. Um dos materiais oferecidos pelo *Plain Language* é um manual de orientações de escrita simplificada que, embora tenha sido composto para orientar a produção de textos governamentais, pode ser utilizado como ponto de partida para a investigação de características de acessibilidade para outros gêneros textuais. Esse manual (que possui mais de cem páginas com instruções claras sobre como escrever para um público leigo) foi adaptado por Silva (2016)⁶ para a língua portuguesa – uma vez que foi originalmente escrito para o inglês.

Outra iniciativa semelhante é a campanha *Plain English*⁷, que desde 1979 se ocupa em estudar e orientar a respeito de formas de comunicação “*crystal clear*” – conforme o *slogan* em seu site. Além de disponibilizar materiais instrutivos a respeito de escrita acessível, o projeto oferece cursos e treinamentos para profissionais de

⁵ <https://www.plainlanguage.gov>

⁶ Disponível em <http://www.ufrgs.br/textecc/acessibilidade/page/simplificar/> (Acesso em 20 de Setembro de 2021)

⁷ <http://www.plainenglish.co.uk>

inúmeras áreas que desejam aprender como escrever de forma que possibilite a compreensão do maior número possível de leitores.

A preocupação a respeito de linguagem acessível não é nova. De fato, um dos primeiros esforços acadêmicos a respeito do assunto foi realizado por Rudolf Flesch nos anos 40 nos Estados Unidos. Flesch – austríaco e doutor em Direito que se refugiou nos Estados Unidos durante a Segunda Guerra Mundial – é visto como uma referência no campo da Acessibilidade Textual. Além de promover a conscientização sobre acessibilidade de textos escritos em língua inglesa para trabalhadores imigrantes refugiados, Flesch se tornou um dos principais pesquisadores acadêmicos da área, tendo desenvolvido e colaborado na criação de fórmulas que estimam a complexidade de um texto com base na extensão de palavras e de frases - fórmulas conhecidas como Índice de Facilidade de Leitura de Flesch e Índice de Legibilidade de Flesch-Kincaid (GOLDIM, 2003).

Além das fórmulas que auxiliam na análise do potencial de complexidade textual, Flesch escreveu uma série de textos que instruem a respeito da escrita simplificada, como “*Say what you mean*”⁸ (1972), “*The Art of Readable Writing*”⁹ (1974), “*The ABC of Style: A guide to Plain English*”¹⁰ (1980), “*A New Guide to Better Writing*”¹¹ (1984), entre outros. Seus livros são, até hoje, utilizados como “pontos de partida” de muitas pesquisas acadêmicas sobre o assunto. Além dos livros mencionados, Flesch também escreveu obras voltadas para linguagem acessível no Direito e na área dos negócios – uma vez que acessibilidade não pode ser considerada no texto por si só, mas depende de uma série de fatores como tipologia/gênero textual, área, público etc.

Outro autor considerado referência para a área é William Henry DuBay, padre americano ativista, que desde o final dos anos 80 tem estudado a linguagem simples e atuado como consultor da área. Em 2004, publicou seu primeiro livro sobre acessibilidade textual, “*The Principles of Readability*”¹², que apresenta fórmulas para calcular complexidade textual e orientações para escrita simplificada. Entre as

⁸ Diga o que você quer dizer

⁹ A Arte da Escrita Legível

¹⁰ O ABC do estilo: um guia para o *Plain English*

¹¹ Um Novo Guia para Melhor Escrita

¹² Os Princípios de Leiturabilidade

principais contribuições da obra, destaca-se a ênfase de DuBay sobre a compreensão e adaptação textual tendo por ponto de partida o leitor-determinado¹³.

Desde os anos 2000, DuBay tem publicado uma série de textos em periódicos e livros sobre a linguagem acessível. Entre eles, estão “*Smart Language: Readers, Readability, and the Grading of Text*”¹⁴ (2007) – que enfatiza que uma leitura acessível é aquela que corresponde ao grau de instrução do leitor-determinado, e “*Unlocking Language: The Classic Readability Studies*”¹⁵ (2007) – que reúne instruções de acadêmicos pioneiros sobre leiturabilidade e relação entre compreensão textual e instrução escolar.

Com o fácil acesso permitido pela internet e com a constante globalização dos últimos anos, a acessibilidade textual tem recebido maior atenção no cenário acadêmico – principalmente alinhada a maiores preocupações com a diversidade sociocultural. No Brasil, contudo, as pesquisas sobre o tema (principalmente os estudos práticos) estão nos passos iniciais se comparados a outros países. Existem alguns projetos, como o Movimento Down¹⁶, que incentivam a comunicação acessível; mas em nível nacional, o tema ainda é muito limitado.

Finatto (2018) apresenta um panorama de pesquisas a respeito da acessibilidade textual no Brasil nos últimos anos. A autora descreve pesquisas de pós-graduandos que buscam compreender a linguagem de áreas como Psiquiatria, Psicologia, Agricultura e Direito e identifica características das linguagens desses campos que podem ser consideradas complexas para leitores leigos e/ou com escolaridade limitada. Dentre essas pesquisas está a de Silva (2018), intitulada “Textos de divulgação para leigos sobre o Transtorno do Estresse Pós-traumático em português: alternativas para a acessibilidade textual e terminológica”, que analisou as características que potencializam a complexidade de textos sobre o Transtorno e, a partir dos resultados coletados, apresentou uma série de instruções para a escrita simplificada em língua portuguesa baseadas na adoção de estratégias de simplificação. Embora a pesquisa tenha sido realizada com textos a respeito do Transtorno do Estresse Pós-Traumático, as orientações a respeito da identificação de

¹³ Nesta tese, utilizaremos o termo leitor-determinado no lugar de “leitor-alvo” por acreditarmos nas orientações da Linguística de Paz que sugerem a substituição de termos com valores bélicos (como “alvo”) por outros, conforme Gomes de Mattos (2006) e Friedrich (2012).

¹⁴ Língua Inteligente: leitores, leiturabilidade e classificação de texto

¹⁵ Desbloqueando a língua: os estudos clássicos de leiturabilidade

¹⁶ <http://www.movimentodown.org.br/jovens-e-adultos/o-que-e-comunicacao-acessivel/>

traços de complexidade e as estratégias de simplificação têm se mostrado eficientes para textos de distintos gêneros e assuntos.

Silva (2018) também adverte que, embora existam estudos a respeito do tema, ainda parece haver falta de clareza na distinção entre complexidade textual (CT), acessibilidade textual (AT) e simplificação textual (ST). O autor explica que CT são todas as características que tornam um texto complexo **para um determinado tipo de leitor**¹⁷, AT são todos os recursos utilizados em um texto (como um resultado final) para que ele possa ser compreendido pelo leitor-determinado e ST são as estratégias utilizadas para composição ou edição de um texto para possibilitar acessibilidade.

Ainda que existam pesquisas sobre acessibilidade textual no Brasil, poucas – ou inexistentes – se concentram na análise de características que potencializam a complexidade de textos voltados para o ensino de língua adicional. A prioridade nas pesquisas sobre AT são, usualmente, textos sobre Saúde e Direito, mas – evidentemente – a área permite (e carece de) estudos a respeito de textos didáticos. Para distinguir nossa pesquisa das que lidam puramente com ATT (Acessibilidade Textual e Terminológica) – já que nosso *corpus* de estudo não é composto por textos de especialidades especializadas ou profissionais - apresentaremos (nas próximas seções) o conceito que denominamos de Acessibilidade Textual Linguodidática – análise de complexidade e acessibilidade de textos voltados para o ensino de línguas.

Flesch (1955), no livro *“Why Johnny can’t read – and what you can do about it”*¹⁸, por exemplo, critica a forma de ensino de leitura nos Estados Unidos da época e “culpa” as metodologias pedagógicas pela dificuldade de compreensão de leitura dos estudantes (e posteriormente, dos adultos). Em 1983, no livro *“Why Johnny Still Can’t Read — A New Look at the Scandal of our Schools”*¹⁹, Flesch retoma o assunto e mantém a crítica ao sistema de ensino de leitura que, em 26 anos, pouco havia evoluído. Com base nessas críticas, o autor advoga por métodos e estratégias mais eficientes que utilizem textos de formas mais assertivas para o ensino de línguas.

Stephens (2021, p. 18) afirma que “mesmo gênios (4% da população) ainda preferem informações claras, simples e bem estruturadas.”²⁰ De forma acessível, a

¹⁷ Grifo nosso

¹⁸ Porque Johnny não consegue ler – e o que você pode fazer sobre isso

¹⁹ Porque Johnny ainda não consegue ler – um novo olhar ao escândalo das nossas escolas

²⁰ Tradução nossa. Original: *“Note that even geniuses (4% of the population) still prefer clear, simple, structured information.”*

autora oferece um resumo de estratégias para diferentes níveis de competência, conforme Quadro 1 abaixo:

Quadro 1 - Três estilos de comunicação para diferentes competências de leitura

1. Baixa literacia:	Gráficos visuais com descrições escritas. Frases curtas e simples.
2. Literacia média:	Palavras comuns. Estrutura sintática padrão. Diagramação clara e simples.
3. Leitores proficientes:	Linguagem clara. Termos específicos. Organização familiar. Boa diagramação e estrutura.

Fonte: Stephens (2021, p.18)

Embora as orientações resumidas por Stephens (2021) sejam concisas, elas englobam os fundamentos de acessibilidade textual apresentadas por outros autores citados nesta seção. Conforme supracitado, uma vez que a estimativa de complexidade só pode ser feita em relação a um determinado tipo de leitor, cada novo perfil exige novas orientações e pesquisas, embora trabalhos prévios sirvam como alicerces estáveis e pontos de partida confiáveis.

Na próxima seção, descreveremos o conceito que denominamos Acessibilidade Textual Linguodidática (ATLD) que inclui orientações sobre AT adaptadas para textos didáticos. Também, apresentaremos alguns conceitos sobre AT que relacionaremos a ATLD.

2.2 ACESSIBILIDADE TEXTUAL LINGUODIDÁTICA

A importância de textos escritos para o ensino de línguas – seja língua materna ou adicional - é inquestionável. A Linguística apresenta inúmeros estudos que relacionam Psicologia e Cognição ao desenvolvimento da leitura (WIGFIELD, GLADSTONE e TURCI, 2016; FERRER et al., 2007), bem como pesquisas que apontam falhas no ensino da leitura em sala de aula (JOHNSON, 2008), vantagens

ou desvantagens de utilizar textos autênticos²¹ para o ensino de línguas adicionais (GILMORE, 2007), relação entre consciências metalinguísticas com habilidade de ler (HODGSON, 1992) etc. Entretanto, estudos sobre a relação do texto e do aprendizado de línguas estão longe de serem esgotados e são essenciais para as áreas de ensino.

Silva, Moll e Perna (2021) enfatizam a importância do papel dos leitores para qualquer observação de acessibilidade textual. Os autores indicam que questões socioculturais são essenciais para a observação de elementos que potencializam a complexidade de um texto e – ainda mais – para a decisão e aplicação de estratégias de simplificação que busquem acessibilidade. Sendo assim, tendo consciência de que simplificação textual é sempre relacionada ao tipo de texto e a relação do texto com o leitor, é lógico afirmar que não existirá uma única fórmula que poderá ser eficiente para todos os tipos de texto: o que funcionará para textos da Medicina, não necessariamente funcionará para textos didáticos e vice-versa.

Logo, ao lidarmos com acessibilidade textual voltado ao ensino (ATLD), os fatores que tornam um texto complexo, bem como as estratégias necessárias para tornar os textos acessíveis dependem não só do tipo de texto com os quais trabalharemos, mas também – e essencialmente – com o nível de instrução e proficiência linguística (mais especificamente proficiência leitora) dos alunos para os quais os textos são voltados.

Para o ensino de línguas adicionais, sabemos que existem determinadas ordens de conteúdos gramaticais, de estruturas sintáticas, de vocabulários e de habilidades que otimizam a assimilação e aprendizado linguístico. Para a língua inglesa, por exemplo, o CEFR – *Common European Framework of Reference*²² apresenta os conteúdos e habilidades que devem ser trabalhados com cada nível linguístico para garantir eficiência de aprendizado. Entretanto, ainda que existam orientações a respeito da ordenação de conteúdos, é comum encontrarmos estruturas e vocabulários complexos em textos de níveis básicos – como a utilização da voz passiva em textos para alunos do nível elementar, por exemplo.

²¹ A definição de “textos autênticos” adotada nesta tese é: materiais multimeios que são produzidos para falantes nativos de uma língua (como jornais, revistas, programas de televisão etc.) e não especificamente para fins didáticos (embora possam ser utilizados para essas finalidades).

²² Quadro de Referência Comum Europeu (www.coe.int/lang-cefr)

Obviamente, a manutenção e evolução linguística ocorrem pela apresentação de novas estruturas e léxico por meio de insumos orais e escritos. Entretanto, qual a extensão e distanciamento entre os conteúdos esperados para um nível e conteúdos divergentes apresentados em textos que não prejudicam o aprendizado e “internalização” da língua?

Krashen (1987) formula uma teoria sobre aquisição de linguagem que é composta por cinco hipóteses:

I. Distinção entre aquisição e aprendizagem: sendo aprendizagem um processo consciente de aprender o sistema linguístico e aquisição um processo subconsciente de assimilação intuitiva e natural provenientes das interações realizadas por meio da língua;

II. Monitor: relaciona aquisição e aprendizagem e defende que os esforços espontâneos e criativos de comunicação são decorrentes da capacidade natural do ser humano de assimilar línguas são regulados pelo conhecimento consciente das regras gramaticais de uma determinada língua e das exceções às regras. Além disso, defende que a personalidade individual tem grande impacto nessa relação, uma vez que pessoas mais introspectivas tendem a bloquear a espontaneidade e, conseqüentemente, interferir na aquisição de novas unidades linguísticas;

III. Ordem natural: afirma que a aquisição de estruturas gramaticais acontece por meio de uma ordem previsível;

IV. Insumo: defende que a quantidade de linguagem oferecida ao aluno para garantir progressão deve ser um pouco além do estágio atual de competência linguística do aprendiz e relevante, interessante, compreensível e oferecido em quantidade suficiente e por meio de meios que garantam o conforto dos alunos. Krashen (1987) ilustra essa quantidade por meio de fórmula $(c+1)$, na qual ‘c’ representa as competências e conteúdos que um aluno já possui e o “+1” o insumo necessário para garantir evolução linguística;

V. Filtro afetivo: número de variáveis afetivas (como motivação, ansiedade e autoconfiança) que facilitam ou dificultam a aquisição de uma segunda língua. Um filtro afetivo baixo (garantido por situações que reduzam quaisquer aspectos negativos relacionados ao aprendizado) é mais efetivo para aquisição linguística, enquanto um ambiente ou recursos que aumentam a ansiedade e a falta de autoconfiança em relação à língua aumentam, também, o filtro afetivo e – conseqüentemente – atrasam a aquisição.

A teoria de Krashen (1987) é relevante para nossa pesquisa por três principais motivos: I. Porque objetivamos investigar se a **ordem natural** (por meio da progressão de conteúdos) está apresentada eficientemente por meio dos textos; II. Porque acreditamos que existem possibilidades de otimizar o **insumo** oferecido pelos textos analisados e; III. Porque nossas experiências como docentes nos demonstram que textos complexos tendem a aumentar o **filtro afetivo** dos nossos alunos e, conseqüentemente, dificultar o aprendizado.

Como sabemos, existe uma forte relação entre fatores psicológicos e aprendizado de língua. Além de Krashen (1987), Gardner, Tremblay e Masgoret (1997) e Yang (1999), por exemplo, apontam que variáveis afetivas influenciam em maior ou menor êxito de aquisição e aprendizado de línguas. Entre essas variáveis, além dos fatores relacionados ao estresse e ansiedade, também está a relação dos alunos diante do desconhecido. No que diz respeito à leitura, textos claramente complexos ou com uma grande quantidade de vocabulário e estruturas difíceis tendem a desestimular a leitura e, conseqüentemente, a aquisição da linguagem.

Stephens (2021) afirma que o sentimento de um leitor sobre a sua fluência em leitura se relaciona com acesso rápido e fácil à informação, e que a avaliação da complexidade de um texto se dá pelo equilíbrio entre o esforço decorrente da leitura e seu ganho informativo. Com base em estudos fundamentados pela Neurociência sobre complexidade textual, a autora apresenta algumas orientações importantes para uma escrita acessível:

1. Uma escrita acessível leva em consideração o interesse do leitor sobre o material. Portanto, devem ser utilizadas técnicas de escrita que direcionem a atenção;
2. A forma do texto deve não só atender as expectativas do gênero textual, mas as expectativas do leitor diante de um determinado texto. Isso significa que esperamos uma determinada forma ao lermos uma notícia e outra forma se lermos um romance; o respeito aos componentes de cada gênero textual é sumário;
3. Uma criação prévia de esquema ou estrutura ajuda a planejar o desenvolvimento da atenção do leitor durante a leitura do texto. Durante a fase denominada “pré-atenção” (que antecede a leitura concentrada de um material), um leitor tende a se organizar e engajar por elementos como:

cores, tamanho de texto, posição de elementos, predicabilidade, número de elementos, impacto emocional e contraste;

4. Ao escrever um texto voltado para acessibilidade, o redator deve “calcular” a facilidade de leitura levando em conta:
 - a. A velocidade de processamento do texto e os possíveis elementos que podem “segurar” o fluxo natural de leitura – causados por língua ou pela diagramação;
 - b. Esforço e energia necessários para processar, entender e decodificar as informações transmitidas pelo texto;
 - c. Facilidade (ou dificuldade) de resgatar informações da memória de longo prazo para relacionar com os elementos textuais. Essas informações podem ser mais facilmente elencadas por meio de estratégias de simplificação que, entre outras, envolvem escolha de palavras e estruturação sintática de frases.

No que diz respeito às estratégias de simplificação, Stephens (2021) destaca que um redator de um texto acessível precisa compreender que a escolha lexical é um fator essencial para a facilitação de leitura e que palavras podem ser compreendidas como tendo uma “função de pensamento”; conseqüentemente, a seleção de palavras para um texto acessível deve ser abalizada tendo o léxico como um dos pilares de simplificação. Assim como outras atividades, a leitura requer que o cérebro busque por padrões similares nas “unidades de armazenamento” (memória de trabalho e memória de longo prazo); uma falta de consistência (como estruturação sintática desconhecida ou palavra complexa) tende a impulsionar estressores psicológicos (como ansiedade e frustração) que, por sua vez, comprometem a compreensão de leitura. Logo, Stephens (2021) demonstra que para o entendimento de um texto, um leitor precisa:

- I. Compreender a relação entre fonema e morfema;
- II. Buscar pela palavra no léxico da sua memória de longo prazo;
- III. Relacionar a palavra com um significado possível e/ou com relação prévia com essa unidade;
- IV. Reter palavras e possíveis significados na memória de trabalho;
- V. Utilizar conhecimentos de estruturas de frases e pontuação para construir o significado das frases;

VI. Identificar a função das palavras por meio de sua posição e relação sintática na frase;

VII. Processar todos os requisitos anteriores para compilar o significado das frases;

VIII. Retornar para algum estágio prévio caso ocorra alguma dissonância ou falta de congruência em algum ponto.

Stephens (2021, p. 17) ainda afirma que:

Quando a compreensão é bem-sucedida, a informação é conectada a uma rede já existente de informações relacionadas ou similares na nossa memória de longo prazo. Somente as diferenças são salvas entre o uso atual da palavra e instâncias prévias. Novas informações precisam ser assimiladas ou alteradas para que ocorram conexões, ou novas conexões e *frameworks* precisam ser criados.²³

Ao observarmos acessibilidade textual de materiais escritos em língua adicional para estudantes, podemos concluir que essa relação necessária entre encontros com instâncias prévias da palavra e/ou criação de novas relações e *frameworks* entre palavras e entre palavras e seus significados é um processo mais complexo do que a assimilação de palavras novas na língua materna. Uma vez que, conforme previamente citado, um dos estágios para a compreensão de palavras é a relação entre morfema e fonema (e que cada língua difere nessas relações), é necessário que os leitores de um texto escrito em língua adicional tenham familiaridade com os morfemas e com os fonemas do idioma estudado para, a partir desses e da possível indução de significado pelos afixos das palavras, conseguir realizar as conexões necessárias para assimilação do léxico. Acreditamos que o primeiro passo para um texto acessível voltado para fins didáticos está estritamente relacionado ao léxico. Antes do estágio V sugerido por Stephens (2021), que afirma que os leitores precisam “utilizar conhecimentos de estruturas de frases e pontuação para construir o significado das frases”, é preciso que os estudantes tenham uma variedade lexical suficiente para assimilação primeiro do léxico e depois da relação entre as palavras. Stephens (2020, p. 17) afirma que “Decifrar o significado das palavras é

²³ Tradução nossa. Original: “*When comprehension succeeds, information is connected to the existing network of related or similar information in our long-term memory storage. Only the differences are saved between the word’s current use and prior encounters. New information must be assimilated or altered so it will connect. Or new connections or frameworks have to be created.*”

inextricavelmente relacionado ao descobrimento da sintaxe da frase [na qual as palavras estão inseridas] para confirmar o significado delas.”²⁴

Além do significado das palavras, um processo essencial para compreensão de um texto são as inferências feitas pelo léxico e pela relação entre as palavras. Essa inferência pode ser compreendida como uma “antecipação” do que está por vir e uma conexão dos elementos sequenciais de um texto. Se um material é redigido de forma que o autor traga muitas informações por meio de implicaturas²⁵, é exigido que um leitor faça as inferências corretas. Evidentemente, não podemos assumir que todo leitor terá a mesma capacidade para fazer as relações necessárias para evocar as inferências corretas das implicaturas presentes em um texto – uma vez que o único contato com o autor se faz por meio do texto (STEPHENS, 2021).

A OECD (*Organization for Economic Co-operation and Development*) tem um programa chamado *Programme for the International Assessment of Adult Competencies (PIAAC)*²⁶ que tem como principal finalidade investigar e estudar diferentes habilidades de adultos, tais como proficiência de leitura, processamento de informação, habilidade de solução de problemas etc. Por meio dos resultados de uma das pesquisas conduzidas pela organização, o grupo organizou um gráfico para avaliar a competência de leitura de adultos. No Quadro 2 abaixo, apresentamos um resumo oferecido por Stephens (2021, p.17) do trabalho realizado pelo PIAAC:

Quadro 2 - Avaliação de competência de leitura de adultos da OECD – resumo

NÍVEL 1	NÍVEL 2	NÍVEL 3	NÍVEL 4	NÍVEL 5
Lê textos relativamente curtos para localizar uma informação específica. Segue orientações escritas simples.	Lê textos mais complexos para localizar uma informação específica ou textos mais simples para localizar múltiplas informações. Faz inferências em baixo nível.	Seleciona e integra informações de diversas fontes ou de diversas partes de um mesmo texto. Faz inferências em baixo nível de múltiplas fontes e identifica informações	Integra e sintetiza informações de múltiplas fontes e de textos complexos e longos. Faz inferências complexas e utiliza conhecimentos prévios generalizados.	Interpreta textos densos e complexos. Faz inferências em alto nível e utiliza conhecimentos prévios especializados.

²⁴ Tradução nossa. Original: “Deciphering the meaning of the words is inextricably linked with discovering the syntax of the sentence to confirm the word meaning. Together, these disclose the semantics – the meaning.”

²⁵ Ver Grice (1981)

²⁶ <https://www.oecd.org/skills/piaac/>

		relevantes e irrelevantes.	Avalia a qualidade de um texto.	
--	--	----------------------------	---------------------------------	--

Fonte: Stephens (2021, p.17)²⁷

Tendo em mente que as informações fornecidas pelas pesquisas da OECD são voltadas para as competências de leitura em língua materna, é preciso compreender que as mesmas diretrizes podem não ser diretamente aplicáveis para leitores diante de textos em línguas adicionais, mesmo leitores com maiores graus de proficiência. Consequentemente, para leitores aprendizes de uma LA, estratégias de simplificação e de acessibilidade são essenciais, conforme a presente pesquisa pretende investigar.

Não é novidade, também, a importância de práticas eficientes de leitura para o aprendizado de uma segunda língua ou língua adicional. Consequentemente, a seleção e adaptação de textos a serem utilizados para o ensino dessas línguas são essenciais. Da mesma maneira que um bom texto pode colaborar para o aprendizado, um texto inadequado pode interferir negativamente. Sendo assim, o objetivo principal da ATLD é analisar conteúdos e características textuais que possam ser complexos para cada nível linguístico e orientar a respeito de estratégias para a composição ou edição de um texto para que ele se torne acessível e, ao mesmo tempo, apresente insumos suficientes para o aprendizado e aquisição.

Outra orientação essencial relacionada à ATLD é proveniente da nossa experiência em sala de aula. Como sabemos, a utilização de textos que estão inseridos dentro de unidades de livros didáticos tem como objetivo, primordialmente, proporcionar contato e impulsionar aprendizado com/dos pontos gramaticais e/ou vocabulários trabalhados na unidade onde o texto está disposto. Ou seja, se – por exemplo – o conteúdo gramatical de determinada unidade é o *Simple Past*, o texto desse capítulo deveria ter exemplos desse tempo verbal com a finalidade de direcionar a atenção do aluno e otimizar a internalização de estrutura e vocabulário. Dessa forma, como docentes, entendemos que, muitas vezes, a presença de estruturas ou vocabulários não relacionados ao conteúdo (quando existem alternativas para substituição) e que não tenham sido apresentados e trabalhados com os alunos acabam se tornando elementos distrativos. Se no texto da unidade do *Simple Past*, por exemplo, há uma frase na voz passiva com um verbo irregular cujo

²⁷ Tradução nossa

Past Participle difere do *Simple Past* (como *gave - given*), a presença dessa frase pode desviar a atenção do aluno do *Simple Past* e/ou interferir na “absorção” do vocabulário relacionado aos verbos irregulares (“Mas o passado de *give* não era *gave*?”). Frequentemente, observamos que uma palavra ou estrutura “diferenciada” se torna disruptiva (em um sentido negativo) para o objetivo do texto e distrai os alunos dos pontos principais – especialmente com alunos que acreditam que precisam compreender todas as palavras de um texto para entender seu sentido e/ou alunos neuroatípicos. Logo, a ATLD deve levar em consideração, também, a presença de unidades lexicais ou sintáticas que possam ser potenciais de distração.

Ainda que não tenham sido desenvolvidas especialmente para análise de complexidade textual, as diretrizes propostas pelo CEFR classificam as habilidades linguísticas dos alunos para cada nível de aprendizado (não só de leitura, mas também das demais habilidades linguísticas). Essas orientações podem auxiliar a observar textos utilizados para o ensino de ILA para verificar se eles estão condizentes com o perfil-padrão dos estudantes de cada nível de aprendizado. Na próxima seção, apresentaremos algumas orientações que são relevantes para nossa pesquisa.

2.3 ORIENTAÇÕES DO CEFR PARA LEITURA EM ILA

O *Common European Framework or Reference for Languages: Learning, teaching, assessment* (CEFR) é um instrumento descritivo-normativo criado em 2001 que delinea as habilidades, os níveis e as estratégias para o aprendizado de línguas. De acordo com a descrição apresentada na mais nova versão do documento:

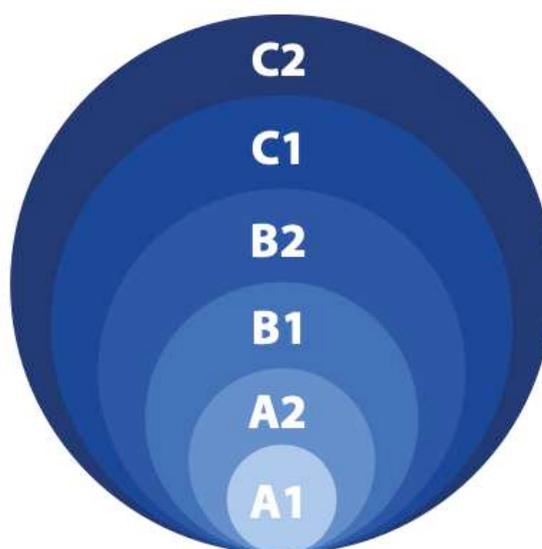
O CEFR é destinado para a promoção de uma educação plurilíngue de qualidade, para facilitar maior mobilidade social e para estimular reflexão e troca entre profissionais das línguas. [...] Além disso, o CEFR fornece uma metalinguagem para discutir a complexidade da proficiência linguística em uma Europa²⁸ multilíngue e intercultural e para auxiliar aqueles que elaboram políticas educacionais a refletir sobre objetivos de aprendizagem e resultados que sejam coerentes e transparentes.²⁹ (EUROPAT, 2020, p.11)

²⁸ Embora o documento cite especificamente a Europa, o CEFR é utilizado como parâmetro em países de outros continentes, incluindo o Brasil.

²⁹ Tradução nossa. Original: “*The CEFR is intended to promote quality plurilingual education, facilitate greater social mobility and stimulate reflection and exchange between language professionals [...]. Furthermore the CEFR provides a metalanguage for discussing the complexity of language proficiency*”

O documento mais atualizado do CEFR foi lançado em 2020 e apresenta com detalhes (em 280 páginas), noções essenciais para promover o ensino e o aprendizado de uma língua adicional. Entre as informações apresentadas, estão quadros que classificam as habilidades linguísticas dos aprendizes e que descrevem as capacidades dos estudantes para cada um dos níveis de aprendizado - classificados entre A1 e C2 - sendo o primeiro considerado um nível básico e o último avançado, como ilustra a imagem a seguir (Figura 1):

Figura 1 - Níveis de referência do CEFR



Fonte: Europarat, 2020, p.36

Para os propósitos da nossa pesquisa, os quadros descritivos do CEFR das habilidades de leitura podem ser utilizados como parâmetro para comparar os conteúdos expostos nos textos e seus possíveis traços de complexidade com as capacidades dos estudantes para cada nível linguístico. No quadro a seguir (Quadro 3), apresentamos uma tradução das habilidades globais de leitura propostas pelo CEFR de 2020³⁰:

for all citizens in a multilingual and intercultural Europe, and for education policy makers to reflect on learning objectives and outcomes that should be coherent and transparent.”

³⁰ O quadro original está nos apêndices desta tese.

Quadro 3 - Descrição do CEFR de habilidades globais de leitura

NÍVEL	DESCRIÇÃO
Pre-A1	Consegue reconhecer palavras/sinais familiares acompanhados por imagens, tais como menus de restaurantes de <i>fast-food</i> ilustrado com fotos, ou livros de imagens que utilizam vocabulários familiares.
A1	Consegue entender uma frase por vez de textos muito curtos e simples, utilizando nomes, palavras e frases básicas e familiares e relendo quando for necessário.
A2	Consegue entender textos curtos e simples em tópicos familiares de um tipo concreto que é redigido com uma alta frequência de palavras do cotidiano ou relacionadas ao seu campo profissional. Consegue entender textos curtos e simples redigidos com um vocabulário que seja altamente frequente ³¹ incluindo palavras utilizadas no léxico internacional ³² .
B1	Consegue compreender em um nível satisfatório textos que vão direto ao ponto ³³ e que sejam sobre tópicos do seu interesse.
B2	Consegue ler com um alto grau de independência, adaptando estilo e velocidade de leitura de acordo com diferentes textos e propósitos e utilizando fontes de referência apropriadas. Tem um amplo vocabulário de leitura, mas pode ter dificuldades com algumas expressões de baixa frequência.
C1	Consegue entender com detalhes textos longos e complexos, independentemente de serem sobre sua área de especialização, desde que possa reler partes difíceis. Consegue entender uma grande variedade de textos incluindo literários, artigos de jornal ou revista e publicações acadêmicas ou profissionais, desde que possam reler os textos e acessar ferramentas de referência.
C2	Consegue entender praticamente todos os tipos de texto, incluindo abstratos e de estruturas complexas, textos literários altamente coloquiais ou não-literários. Consegue entender uma grande variedade de textos longos e complexos, percebendo diferenças sutis de estilo e, também, significados implícitos e explícitos.

Fonte: Tradução de Europarat, 2020, p.54

Além das classificações de habilidades em nível global, o CEFR difere dois tipos de leitura: I. Leitura para orientação (ou seja, para encontrar informações específicas – exemplificado no documento como fazer *skimming/scanning* de um horário de ônibus para encontrar dados sobre o veículo que deseja pegar), e II. Leitura

³¹ O documento do CEFR utiliza 'alta frequência' para se referir a palavras que sejam comumente usadas no cotidiano (como '*morning*', '*food*' etc.) e 'baixa frequência' para o oposto (palavras menos comuns).

³² Ou seja, palavras da língua inglesa que são utilizadas em outras línguas, como '*park*', '*sales*' etc.

³³ *Straightforward*

para informação, também definido como leitura detalhada que, como o nome sugere, é feita para obter informações detalhadas sobre determinados assuntos.

Visto que os objetivos dos textos do nosso *corpus* de estudo estão mais relacionados ao tipo de leitura para informação (uma vez que são utilizados para interpretação e não só para reconhecimento de orientações específicas), acreditamos que seja importante utilizarmos, também, as orientações sobre as habilidades dos estudantes (para cada nível) para esse tipo de texto. No quadro abaixo (Quadro 4)³⁴ apresentamos uma tradução das orientações sobre a leitura para informações:

Quadro 4 - Descrição do CEFR de habilidades da leitura para informação

NÍVEL	DESCRIÇÃO
Pre-A1	Consegue entender os tipos mais simples de materiais informativos, como menus de restaurantes <i>fast-food</i> ilustrados com fotos, ou uma história ilustrada redigida com palavras/sinais muito simples e do cotidiano.
A1	Consegue captar a ideia de materiais informativos simples e curtos e descrições simples, especialmente se houver apoio visual. Consegue entender textos curtos sobre tópicos do seu interesse (por exemplo, notícias sobre esportes, músicas, viagem ou histórias) que sejam escritos em uma linguagem muito simples e apoiados por ilustrações e imagens.
A2	Consegue identificar informações específicas em materiais simples como cartas, brochuras e notícias descrevendo eventos. Consegue acompanhar a ideia geral de notícias sobre tópicos familiares, desde que o conteúdo seja familiar e previsível. Consegue entender partes específicas das informações principais em notícias curtas ou em artigos simples nos quais figuras, nomes, ilustrações e títulos tenham papéis proeminentes no suporte do significado do texto. Consegue entender os pontos principais de textos curtos que abordem tópicos do cotidiano (como estilo de vida, <i>hobbies</i> , esportes e tempo). Consegue entender textos descrevendo pessoas, lugares, acontecimentos diários e cultura, desde que sejam escritos com linguagem simples. Consegue entender informações dadas em brochuras e mapas ilustrados (por exemplo, as principais atrações de uma cidade). Consegue entender os pontos principais em notícias curtas de assuntos de seu interesse (por exemplo, esportes e celebridades). Consegue entender um relatório ou descrição factual curtos sobre seu campo de trabalho/interesse, desde que sejam escritos com linguagem simples e que não contenham detalhes imprevisíveis.

³⁴ O quadro original está nos apêndices desta tese.

	<p>Consegue entender a maior parte do que as pessoas falam sobre si mesmas em anúncios pessoais ou correspondências e do que elas relatam gostar sobre outras pessoas.</p>
B1	<p>Consegue entender textos factuais que vão direto ao ponto em tópicos relacionados aos seus interesses ou campos de estudo.</p> <p>Consegue entender textos curtos sobre tópicos que são familiares ou de interesse atual no qual as pessoas dão suas opiniões (por exemplo, contribuições críticas em uma discussão on-line em um fórum ou uma carta de um leitor para um editor).</p> <p>Consegue identificar as conclusões principais que estejam claramente sinalizadas em textos argumentativos.</p> <p>Consegue reconhecer a linha de argumentação do problema apresentado, embora não de forma detalhada.</p> <p>Consegue reconhecer os pontos significantes em notícias que vão direto ao ponto sobre tópicos familiares.</p> <p>Consegue entender a maior parte das informações factuais que podem encontrar em tópicos familiares de interesse, desde que tenham tempo suficiente para leitura.</p> <p>Consegue entender os pontos principais em notas descritivas, tais como placas de museu e quadros explicativos em exposições.</p>
B2	<p>Consegue obter informações, ideias e opiniões de fontes altamente especializadas dentro do seu campo.</p> <p>Consegue entender artigos especializados fora do seu campo, desde que possa utilizar um dicionário para consultar, ocasionalmente, sua interpretação de terminologia.</p> <p>Consegue entender artigos e relatórios sobre problemas atuais nos quais posturas e opiniões particulares sejam adotadas.</p> <p>Consegue reconhecer quando um texto apresenta informação factual e quando um texto procura convencer leitores sobre algo.</p> <p>Consegue reconhecer diferentes estruturas em um texto discursivo: argumentos contrastivos, apresentações de problema e soluções e relações de causa e efeito.</p>
C1	<p>Consegue entender com detalhes uma grande quantidade de textos longos e complexos que são comumente encontrados em contextos sociais, profissionais e acadêmicos, identificando os pontos mais minuciosos, incluindo atitudes e opiniões mencionadas e implícitas.</p>
C2	<p>Consegue entender os pontos mais minuciosos e implicaturas de um relatório complexo ou artigo, mesmo que fora do seu campo de especialização.</p>

Conforme podemos perceber pelos quadros indicativos do CEFR, a compreensão leitora se torna um pouco mais natural e abrangente (ou seja, que vai além de textos curtos, de tópicos exclusivamente familiares e de materiais apoiados por recursos visuais) a partir do nível B1 (considerado intermediário). Antes desse nível (ou seja, nos níveis Pre-A1, A1 e A2), a leitura é muito condicionada e limitada por diversos fatores contextuais, características linguísticas e extralinguísticas, e gêneros textuais. Logo, textos que são utilizados para o ensino de uma língua adicional para níveis iniciais (elementar e pré-intermediário)³⁵ devem atentar para que suas complexidades não sejam muito além daquilo que é descrito como possível para os estudantes, caso contrário, não só a interpretação, mas também a manutenção linguística pode ser prejudicada.

Em suma, ao conduzirmos análises de complexidade que busquem a acessibilidade linguodidática, devemos cruzar as informações dos textos com as indicações apresentadas pelos quadros de habilidade do CEFR para que, então, possamos estimar a complexidade dos materiais de forma coerente à realidade de um aprendiz de uma língua adicional. Levando essas orientações em consideração, apresentamos, na próxima seção, alguns conceitos fundamentais para pesquisas de acessibilidade textual que serão complementares às observações das indicações do CEFR.

2.4 ANÁLISES DE ACESSIBILIDADE TEXTUAL

É de conhecimento comum que textos são estruturas complexas, com multiníveis e que podem ser estudados por meio de diferentes perspectivas: gramatical, lexical, semântica, sintática, pragmática etc. Por conseguinte, análises de AT também podem ser consideradas complexas e dependem de uma série de fatores.

Para começo, é importante reforçar o que estudiosos da área, como DuBay (2004), enfatizam: análises de AT (bem como práticas de ST voltadas à acessibilidade) só devem ser conduzidas tendo como “fio-condutor” um público leitor determinado – seja ele um público real ou um protótipo de leitor. Ou seja, para definir

³⁵ Textos voltados para aprendizagens de ILA de níveis intermediários (e até avançados) também devem atentar para complexidade textual, especialmente em grupos que tenham alunos com claras limitações linguísticas e/ou neuroatípicos. Entretanto, de forma global, os textos voltados para os níveis iniciais precisam de maiores atenções, especialmente levando em consideração as características das habilidades dos alunos diante da leitura nesses níveis (elementar e pré-intermediário).

se um texto é potencialmente complexo e quais as características que podem torná-lo inacessível (ou menos acessível) dependem – primordialmente – de “para quem” o texto será destinado. Silva, Moll e Perna (2021) apontam práticas de ST como formas de endereçamentos para leitores presumidos que, por sua vez, orientam as escolhas de perspectivas para análise de complexidade e as tomadas de decisão de estratégias de simplificação.

Uma vez que entendemos que análise de AT envolve processos múltiplos e determinados por uma série de fatores, podemos posicioná-la como Análises Multimodais (AMD), que são práticas de estudos e observações que consideram mais de uma dimensão, métrica ou parâmetro para analisar um objeto de estudos. Embora a AMD seja mais tradicionalmente relacionada à LC, ela também pode orientar pesquisas de AT, como a realizada por Silva (2018).

A AMD sugere a observação de dois ou mais aspectos textuais combinados. Biber (1988/1995) defende que nenhum traço textual isolado seria o suficiente para observar características textuais de forma eficiente e que pesquisas baseadas unicamente em características situacionais e/ou fatores linguísticos não dariam conta de descrever fenômenos linguísticos de forma completa. Para o autor, o ideal em análises textuais “seria combinar-se as duas perspectivas, isto é, a descrição firmada em características situacionais com a descrição baseada em traços linguísticos.” (BERBER SARDINHA, 2000, p.100).

Embora o caráter da AMD seja, basicamente, quantitativo e computacional – uma vez que esse tipo de análise é realizada utilizando uma grande quantidade de aspectos linguísticos e por meio de dados estatísticos de traços textuais de *corpora* – a compreensão de que um estudo textual se torna mais completo quando mais de uma característica textual é levada em consideração sustenta as práticas de pesquisa de AT – que também defendem que multimodalidades textuais devem ser levadas em consideração para acurácia de análises de complexidade. O que Berber Sardinha (2000, p. 101) afirma na citação abaixo, embora tenha sido dirigido à AMD em LC, também se aplica a AT:

O analista deve dispor de um arsenal que inclua o maior número possível de características linguísticas, já que o aumento da quantidade de parâmetros implica em número maior de traços linguísticos necessários para cobrir a maior gama de parâmetros.

Feita essa breve caracterização da AMD, é importante destacarmos que abordamos esse tipo de análise como forma de inspiração para o desenvolvimento deste trabalho – ainda que esta tese não seja, frisamos, uma pesquisa de LC. Uma vez que a aplicação da pesquisa levará em conta distintas dimensões textuais (léxico, sintaxe, semântica e pragmática), os preceitos da AMD sustentam nossas decisões de observarmos mais de um parâmetro para observar a complexidade e a [in]eficiência de estratégias dos textos selecionados.

Pesquisas de AT, conforme já mencionado, são direcionadas primeiramente pelo leitor para o qual os textos serão destinados. Portanto, é errôneo pensar que toda análise de complexidade pode ser reduzida a um número específico de métricas e que todo e qualquer texto pode se tornar acessível ao aplicar determinadas estratégias. Ao conduzir pesquisas de ATLD, além de termos a preocupação com a acessibilidade, também precisamos manter em mente que um texto voltado para o ensino precisa ser simultaneamente acessível e oferecer insumos suficientes para a manutenção linguística. Nesta tese, ao utilizarmos o termo “acessibilidade-manutensiva”, estaremos nos referindo às características que permitem acessibilidade textual e, simultaneamente, ofereçam insumos para aprendizado e aquisição de língua adicional.

Por conseguinte, pesquisas de AT (e de ATLD) partem de escolhas metodológicas guiadas pelo leitor-definido e apresentam diversas perspectivas de análise (que podem ser puramente holísticas, unicamente automáticas, ou um equilíbrio de ambas as formas), uma grande quantidade de métricas textuais que podem ser coletadas por meio de programas de PLN, diversas possibilidades de testes e uma série de escolhas que irão diferir entre cada pesquisa.

Ao estipular o leitor-determinado, o analista de acessibilidade textual determina quais níveis textuais observará e quais serão os parâmetros utilizados para a observação. Após, se tiver por objetivo testar eficiência de simplificações e seus resultados de acessibilidade, o analista também precisa eleger quais estratégias são coerentes com seu trabalho e a forma como essas estratégias serão testadas: por meio de comparações métricas, testando com leitores reais, ou das duas formas.

Sendo assim, análises de AT precisam ser coerentemente sustentadas por orientações linguísticas que se adequem ao tipo de leitor estudado e, também, às situações para as quais os textos são destinados. Textos de divulgação científica exigem análises e estratégias diferentes daqueles voltados para o *marketing*, por

exemplo. Não só o gênero textual difere, mas também o público que acessará o texto não é, necessariamente, o mesmo. Logo, as escolhas das características textuais a serem observadas devem ser sabiamente sustentadas. Na segunda parte desta tese, apresentamos a metodologia e as métricas textuais que escolhemos para o desenvolvimento desta pesquisa.

2.5 CARACTERÍSTICAS POTENCIALIZADORAS DE COMPLEXIDADE

Embora análises de complexidade textual devam ser sempre direcionadas pelo perfil do leitor que terá acesso ao texto, pesquisadores da área destacam uma série de fatores que são comumente potencializadores de complexidade. No quadro abaixo (Quadro 5), apresentamos um compilado de algumas indicações de autores sobre acessibilidade textual, como Fry (1977), DuBay (2004), Silva e Finatto (2016), Silva (2018), as diretrizes de simplificação do programa *Plain Language*³⁶ do Governo dos Estados Unidos, entre outros:

Quadro 5 - Fatores que potencializam a complexidade de textos

NÍVEL TEXTUAL	FATORES DE COMPLEXIDADE
Léxico e semântica	<ul style="list-style-type: none"> - Palavras não-cotidianas e “rebuscadas” (especialmente ao tentar variar o vocabulário para evitar repetições); - Palavras extensas; - Termos de áreas técnicas; - Alta frequência de adjetivos e advérbios (principalmente unidades longas); - Substituições frequentes de substantivos por pronomes; - Falsos cognatos; - Palavras homógrafas; - Palavras formadas por prefixos e sufixos complexos;
Sintaxe e organização estrutural	<ul style="list-style-type: none"> - Deslocamento de elementos das suas posições sintáticas canônicas; - Períodos compostos por coordenação e subordinação (orações complexas); - Orações subordinadas posicionadas antes da principal; - Alto número de palavras antes e depois do verbo; - Sintagmas extensos; - Distância grande entre pronome e referente; - Voz passiva (que se enquadra como fator de complexidade tanto por questões de estruturação, quanto por utilização de verbos no particípio que – em língua inglesa – podem ser irregulares); - Remissividade;

³⁶ Programa governamental dos Estados Unidos que advoga e apresenta leis para escrita simplificada, além de oferecer cursos e materiais instrutivos sobre acessibilidade textual. <https://www.plainlanguage.gov/>

	<ul style="list-style-type: none"> - Uso de pontuação não-usual (como utilizar travessões no lugar de vírgulas); - Zeugmas e elipses.
--	---

Fonte: Elaboração própria

As características supracitadas são traços que merecem ser observados com atenção ao analisar complexidade textual. Uma vez que são indicações para textos escritos na língua materna dos leitores, acreditamos que elas devem ser observadas com ainda mais cuidado ao analisarmos materiais escritos em uma língua adicional. Por isso, para esta pesquisa, teremos essas indicações como norteadoras tanto para interpretação das métricas, quanto para a análise holística.

Na próxima seção, apresentamos alguns conceitos essenciais de modelos educacionais denominados 3.0, 4.0 e 5.0 que são usados como diretrizes para a condução desta pesquisa.

3 EDUCAÇÃO 3.0, 4.0 E 5.0: CUSTOMIZAÇÃO E PERSONALIZAÇÃO COMO DIRETRIZES DA ACESSIBILIDADE TEXTUAL

Se olharmos para a história da Educação, podemos perceber que muito mudou em termos de metodologias e práticas de, por exemplo, um século atrás. E se observarmos os últimos anos, podemos notar que o avanço tem sido cada vez mais rápido diante das necessidades do contexto atual.

Graças aos avanços tecnológicos e acadêmicos de áreas como Psicologia, Medicina, Informática, Linguística (entre outras), novas ferramentas e pesquisas permitiram que cientistas de distintas áreas compreendessem melhor o ser humano e suas peculiaridades fisiológicas, neurológicas, psicológicas, linguísticas etc. Por meio de máquinas de ressonância magnética, por exemplo, hoje conseguimos entender como o cérebro humano “se comporta” e “reage” diante de estímulos diferentes – compreensão que abre portas para uma infinidade de possibilidades relacionadas à Educação.

Há muitos séculos (ou talvez, milênios), a humanidade reconhece a importância da erudição. Entretanto, o acesso ao conhecimento, nos tempos bíblicos (e por muito tempo depois disso) era privilégio das classes mais abastadas.

Com o passar dos anos, o acesso à informação foi se tornando mais presente na vida dos cidadãos comuns e hoje, com a internet e a praticidade dos aparelhos que nos conectam a qualquer lugar do mundo, temos oportunidades de obter informações a qualquer momento que – ao mesmo tempo que é uma vantagem imensurável – também altera as expectativas no âmbito educacional e profissional. A escola de modelos antigos, na qual os professores eram “os retentores do conhecimento”, já não atende mais as necessidades do cenário atual que, de acordo com inúmeros pesquisadores, necessita preparar os alunos para a sociedade do conhecimento. (MAGANA, 2017)

É interessante observar o percurso da Educação para perceber como chegamos aonde estamos. Knowles (1984) explica que na idade média, a educação era centralizada em monastério, onde rapazes jovens (estritamente da nobreza) tinham o direito de serem educados por monges por meio de sistemas e metodologias totalmente restritivas e voltadas para o interesse das igrejas. Nesse modelo de educação, os professores eram os únicos responsáveis e – de certa forma, ditadores – sobre as decisões relacionadas ao ensino; escolhendo conteúdos, métodos, formas de avaliação etc. O aluno, por sua vez, deveria ser submisso, obediente e útil para os propósitos do “educador”.

A partir desse momento histórico, o que hoje denominamos de Educação 1.0 começou a predominar as instituições de ensino. Nessa perspectiva educacional – que é existente até hoje em determinados locais – o professor “ensinava” conteúdos enquanto os alunos silenciosamente copiavam e ouviam, sem o direito de participar, de compartilhar ideias e pensamentos, de indagar e de questionar métodos e/ou conteúdo. O aluno era visto como um “cântaro vazio” que só poderia receber o “líquido do conhecimento” de um “cântaro cheio”: o professor.

Entre a Educação 1.0 e 2.0 começou a surgir, também, o “*homeschooling*”, que nada mais é do que uma condução de ensino realizado em casa usualmente por um dos pais ou por um familiar mais velho que, em extrema maioria eram pessoas sem preparo, sem educação formal e sem conhecimentos didáticos para cumprir – eficientemente – o papel de professor.

O modelo de Educação 2.0 pode ser visto como uma forma evolutiva da educação 1.0. Se formos analisar a diferença entre os dois, podemos perceber que o avanço foi considerável (ainda que hoje, muitas metodologias e crenças sejam consideradas defasadas e tenham sido descreditadas). Esse sistema de educação (2.0),

conforme explicam Perna, Delgado e Silva (2021), segue um modelo andragógico que apresenta formas e recursos para ajudar o aluno a “adquirir” informação e desenvolver habilidades. O professor passa a ser o facilitador que prepara metodologias para os aprendizes. Holmes e Abington-Cooper (2000, p.17) explicam que entre esses procedimentos estão o estabelecimento de ambientes apropriados para o ensino, a adoção de mecanismos para planejamento mútuo, o diagnóstico das necessidades de aprendizado, a elaboração de programas e objetivos que atendam essas necessidades, o desenvolvimento de caminhos que permitam a experiência de aprendizagem, a aplicação de técnicas e materiais apropriados, a avaliação de resultados e a reelaboração proveniente das observações de necessidades educacionais interpretados por essas respostas.

Ainda que a Educação 2.0 tenha sido um passo primordial a um acesso mais inclusivo e mais metodológico à informação e ao conhecimento, por muitos anos, escolas mais tradicionais – ainda que contassem com professores preparados para cumprir as exigências mencionadas por Holmes e Abington-Cooper (2000) – continuavam a centralizar o conhecimento no professor e considerava que os mesmos procedimentos metodológicos seriam suficientes para que turmas com grupos heterogêneos de alunos atingissem os mesmos resultados.

De fato, por muitos anos o objetivo da Educação 2.0 acabara por ser esse: que todos os alunos atingissem os mesmos resultados. Acreditava-se, então, que a execução dos mesmos planos de aula para todos os alunos e para turmas distintas conseguiria abranger as necessidades de todos os estudantes, e que os professores poderiam, de uma única forma, garantir que todos tivessem acesso a mesma educação e aproveitamentos semelhantes (metaforicamente similar a um médico receitar o mesmo medicamento para várias pessoas com doenças diferentes). Uma das ferramentas que balizava esse objetivo eram/são as provas padronizadas, cuja criação é atribuída a Horace Mann (por volta de 1843) como forma de assegurar um procedimento metodológico igual para todos os alunos e diferir das formas de avaliação que, até então eram orais; e, principalmente, a Frederick J. Kelly, que em 1914 inventou as provas de múltipla escolha com a crença de que se Henry Ford, por meio do modelo Ford T, conseguia produzir um grande número de automóveis em série e de forma rápida, também seria possível atingir os mesmos resultados de forma eficiente com a educação (DAVIDSON, 2011).

Se os testes eram todos iguais, esperava-se, então, que os alunos recebessem o mesmo tipo de insumo educativo e que, em consequência, atingissem os mesmos objetivos. Por esse motivo, a Educação 2.0 é atualmente ainda presente em muitas escolas e o pensamento de padronização e de resultados iguais – ainda que há muito descreditado – continua sendo o objetivo de muitas redes educacionais. Então, é comum encontrar salas de aulas com uma grande quantidade de alunos sentados um atrás do outro ouvindo um professor “passar o conteúdo”. Ainda que o 2.0 permita que os alunos questionem e participem, o objetivo de alcançar resultados iguais ainda limita as práticas pedagógicas.

Frederick J. Kelly – o responsável pela criação e introdução das provas padronizadas de múltipla escolha nos Estados Unidos – em seu discurso na Universidade de Idaho em 1928 (14 anos após o início das provas padronizadas) advogou com veemência contra sua própria criação. De acordo com Davidson (2011, p.1), uma das partes mais marcantes do discurso de Kelly foi sua frase dizendo que os testes padronizados “são muito brutos para serem utilizados e deveriam ser abandonados.”³⁷. Ainda no mesmo discurso, Kelly argumenta que a aplicação de provas padronizadas produziu um efeito em larga escala de pensamento em baixo nível; ou seja, o tipo de pensamento que não inova e não questiona e que só busca informações na memória para fins de réplica.

É mais do que evidente que a produção de carros e a educação de pessoas são processos extremamente diferentes, e os avanços tecnológicos e de áreas médicas tem dado espaço e fornecido informações que exigem repaginações de metodologias educacionais fato que, em anos recentes, iniciou o movimento da Educação 3.0.

A base da Educação 3.0 é que educadores e estudantes são igualmente partes essenciais do aprendizado. Os professores são profissionais com educação formal que além de cumprir todos os requisitos já orientados por Holmes e Abington-Cooper (2000), também precisam compreender que cada aluno é um ser individual, com necessidades, habilidades, curiosidades e limitações específicas. Os alunos, por sua vez, são integrantes essenciais e centralizadores das práticas pedagógicas que não só tem a função de modelar o ensino junto ao professor, mas também são responsáveis por dar continuidade ao ensino de forma proativa e contínua. O 3.0

³⁷ Tradução nossa. Original “*These tests are too crude to be used and should be abandoned.*”

entende, também, que o aprendizado não é limitado às quatro paredes de uma instituição de ensino, mas é um conjunto de práticas e experiências que estão integradas de forma natural na rotina dos estudantes e dos professores (que também têm o papel de constante atualização).

Perna, Delgado e Silva (2021, p.447) explicam que os pilares da Educação 3.0 são:

- I. Interdisciplinaridade, ou seja, integração entre matérias e conteúdos;
- II. Colaboração entre corpo docente e discente;
- III. Construção de conhecimento com a participação integrada de todos;
- IV. Estudantes em papel ativo;
- V. Desenvolvimento do pensamento crítico;
- VI. Habilidades socioafetivas como trabalho em equipe e resolução de problemas;
- VII. Customização e personalização do aprendizado.

Além disso, Azevedo, Delgado e Silva (2017) incluem como parte da Educação 3.0 a utilização de recursos tecnológicos voltados para o aprendizado (que permitem não só a customização e personalização do aprendizado, mas também contribuem para o acesso constante à informação) e a valorização do inglês como língua adicional – proporcionando maiores contatos com a língua do que os tradicionais um ou dois períodos por semana e contato de forma autêntica e contextualizada.

A customização e personalização, que também estão fortemente presentes nas orientações da Educação 4.0 e da 5.0 (sobre as quais falaremos adiante) são as principais características revolucionárias que alteraram os paradigmas entre a Educação 2.0 e 3.0. Com base nas múltiplas inteligências propostas por Gardner (1983) – psicólogo e pesquisador de Cognição e Educação – esses modelos de educação começaram a compreender que cada estudante é um aprendiz único e que modelos “engessados” e padronizados não poderiam atender a todas as demandas de grupos tão distintos.

Gardner (1983) defende e explica que cada ser humano demonstra e desenvolve suas capacidades cognitivas de formas distintas. Uma vez que cada pessoa, desde o momento que nasce, é exposta a estímulos totalmente diferentes daqueles que outras pessoas recebem, e que cada um se forma e desenvolve de maneiras diferentes, é esperado que suas habilidades e capacidades também sejam

distintas. Portanto, o autor esclarece que inteligência humana é um fator de potencial biopsicológico que se desenvolve por fatores como genética, contexto social, educação formal etc.

Uma vez observada a distinção na forma de desenvolvimento cognitivo, Gardner (1983) propôs um modelo que descreve nove tipos de inteligência (seu primeiro modelo apresentava sete, mas outros dois foram adicionados posteriormente). Esses tipos de inteligência podem ser resumidos da seguinte forma:

I. Inteligência lógico-matemática: capacidade cognitiva de lidar com números, raciocínio dedutivo e conceitos lógicos - exemplo: profissionais da computação;

II. Inteligência linguística: aptidão para comunicação, aprendizagem de línguas, manipulação linguística etc. - exemplo: linguistas;

III. Inteligência espacial: relacionada ao conhecimento espacial e visual, ligado à habilidade de interpretar imagens, manipular espaços, localizar-se etc. - exemplo: arquitetos;

IV. Inteligência físico-cinestésica: habilidades relacionadas ao conhecimento e controle do corpo e ao movimento – exemplo: dançarinos;

V. Inteligência interpessoal: capacidade de compreender outras pessoas – exemplo: psicólogos;

VI. Inteligência intrapessoal: habilidade de autoconhecimento – exemplo: monges;

VII. Inteligência musical: aptidão para reconhecer e produzir sons, melodias etc. – exemplo: músicos;

VIII. Inteligência naturalista: relacionado à natureza e à habilidade de cuidar e compreender animais, plantas etc. – exemplo: veterinários;

IX. Inteligência existencial: propensões a pensamentos filosóficos e metafísicos do comportamento humano – exemplo: filósofos.

Obviamente, as múltiplas inteligências propostas por Gardner (1983) não limitam as formas como um ser humano pode aprender e, ao mesmo passo, não determinam que cada pessoa só tem um tipo de inteligência. Na verdade, o autor sugere que os seres humanos têm mais do que um tipo de inteligência em escala, sendo uma mais forte que a outra.

A sugestão de múltiplas inteligências vai ao encontro das propostas que surgem a partir da Educação 3.0 de personalização e customização do ensino, sendo

personalização a adequação de materiais e metodologias para grupos distintos (como, por exemplo, adaptar o ensino de frações, da Matemática, de forma visual, para um grupo de três alunos que estão com problema de enxergar por meio das demonstrações numéricas), enquanto customização é um tipo de adaptação mais individualizada (como trabalhar com a diferença entre o *Present Perfect* e o *Simple Past* com um aluno já avançado em língua inglesa enquanto o restante da turma está trabalhando unicamente com o *Simple Past*).

Sabendo que salas de aula costumam ter um número grande de alunos, a personalização e customização tem o objetivo de atender diferentes formas de aprendizagem e de contemplar todos os estudantes de forma justa e equalitária. Além da adequação de metodologias e procedimentos, tais práticas também embasam a adaptação de materiais utilizados para o ensino (como reescrita de textos didáticos, por exemplo).

A Educação 4.0 e 5.0 não excluem os princípios da Educação 3.0, mas podem ser vistas como metodologias evolutivas. A 4.0, além das diretrizes já estabelecidas pelo 3.0, entende que o cenário global atual requer o desenvolvimento do pensamento e da inteligência computacional, implementando estratégias que utilizam e ensinam robótica, computação e inteligência artificial. Já a Educação 5.0 incorpora todas as orientações das duas anteriores, mas também a compreensão de que o progresso de inteligências e habilidades socioemocionais (*soft skills*) deve acompanhar os conhecimentos e aptidões técnicas e científicas.

Nesta tese, acreditamos que as indicações de customização e personalização provenientes de sistemas educacionais a partir da 3.0 são práticas que merecem uma atenção especial. Uma vez que sempre haverá alunos distintos, com conhecimentos, inteligências e limitações distintas, podemos concluir que as práticas de customizar e personalizar nunca serão esgotadas. Além disso, uma vez que esses sistemas educacionais, ao enxergar a língua inglesa como língua global, valorizam o ensino contextualizado dessa língua, cremos que a adoção de estratégias de personalização e customização contribuem para esse objetivo. Dentro do nosso contexto educacional – ensino de inglês como língua adicional – compreendemos que uma das formas de respeitar as múltiplas inteligências dos alunos e, conseqüentemente, otimizar o aprendizado e a manutenção linguística é, entre outras, a adequação dos textos utilizados como *input* para garantir o “insumo extra” (c+1) defendido por Krashen (1987) de maneira eficiente e eficaz.

Na próxima seção, abordaremos o PLN e a LC – áreas essenciais para o desenvolvimento prático desta pesquisa.

4 PROCESSAMENTO DA LINGUAGEM NATURAL E LINGUÍSTICA DE CORPUS: UM DIÁLOGO

Ainda que, aparentemente, o PLN e a LC sejam vistos como duas áreas opostas em pensamentos, a união dos dois campos apresenta possibilidades enriquecedoras para uma pesquisa linguística, possibilitando observar manifestações de determinadas línguas de formas mais completas e complexas, conforme defendido por Finatto (2011).

Por essa razão, esta seção apresenta o PLN e a LC com a finalidade de justificar a adoção de ambas as áreas para fins metodológicos desta pesquisa.

4.1 PROCESSAMENTO DA LINGUAGEM NATURAL

Dentro da Ciência da Computação (CiC), o PLN é uma área que tem por finalidade desenvolver e utilizar ferramentas computacionais que facilitem a observação e tratamento de fenômenos e de ocorrências das línguas naturais³⁸. As inúmeras ferramentas de PLN já desenvolvidas pela CiC oferecem recursos para estudos de fonologia, morfologia, semântica, sintaxe etc. Essas ferramentas facilitam práticas de estudos linguísticos e apoiam atividades como tradução automática, busca em acervos textuais, indexação automatizada etc. Por meio da informatização da manipulação de dados, a interface entre a CiC e a Linguística facilita pesquisas que necessitam lidar com um volume grande de textos, o que antes seria inviável sem um apoio computacional.

Apoiados por métodos estatísticos e pelo aprendizado de máquina (capacidade algorítmica de adaptação por meio de dados recebidos), o PLN desenvolve metodologias de interpretação de linguagem tendo como objeto de estudo *corpora* de textos com exemplos típicos da língua utilizada em situações reais, conforme apontam Manning e Schütze (1999). Entre as possibilidades fomentadas pelo PLN,

³⁸ Cada ferramenta de PLN é desenvolvida para uma ou mais línguas específicas (ou seja, utilizar textos escritos em inglês em um programa criado para a língua portuguesa, por exemplo, não apresentará resultados verossímeis.). Visto que cada língua pode ser compreendida como um “sistema” único, programas de PLN utilizam como base de análise as características linguísticas de cada uma delas.

destacamos: contagem de *types* e *tokens* (números de palavras distintas e números totais de palavras de um texto, respectivamente), lematização para fins lexicográficos, análise sintática automatizada, listagem de classes de palavras, detecção automática de línguas, identificação de relações semânticas, análise de morfologia, padrões de tradução de máquina etc.

A empresa SAS (de análise de dados) ao abordar o PLN em seu site³⁹ diz que: "Em termos gerais, as tarefas do PLN segmentam a linguagem em partes menores e essenciais; tentam entender as relações entre elas e exploram como esses pedaços funcionam juntos para criar significados."

Os programas de PLN geralmente "agrupam" textos e automatizam a contagem e a categorização de palavras, além da extração de índices e da apresentação de combinações e probabilidades que podem ser manipulados por linguistas para fins específicos. Embora o PLN possibilite pesquisas com uma grande quantidade de textos, a maior parte das ferramentas também permite estudo e análise de dados de cada texto de forma individualizada – o que difere da LC, sobre o qual falaremos em seguida. Sendo assim, a observação de métricas textuais para cada material analisado permite, também, a coleta de informações sobre apenas um texto que, para determinados fins de pesquisa, faz-se essencial. O *modus operandi* possibilitado pelas ferramentas de PLN facilita pesquisas que necessitam, por exemplo, descrever traços que perfazem uma linguagem complexa de determinado gênero textual.

4.2 LINGUÍSTICA DE *CORPUS*

A Linguística de *Corpus* (LC) é um campo da Linguística que (ainda que não exclusivamente) lida com dados por um viés quantitativo e estatístico com o suporte de ferramentas computacionais.

A LC teve sua origem diante da necessidade de observação e averiguação de usos reais da língua de forma mais extensa possível permitindo, então, teorizações e generalizações sobre língua e linguagem. Filmore (1992, p.35), sobre a LC, afirma que "Todo *corpus* me ensinou coisas sobre a linguagem que eu não teria descoberto de nenhum outro modo".

³⁹ www.sas.com/pt_br/insights/analytics/processamento-linguagem-natural.html

A LC, para o Instituto Brasileiro de Pesquisa e Análise de Dados (IBPAD, 2018)

é:

(...) área da Linguística utilizada para coleta e análise de bases com dados textuais produzidos por falantes reais, a exemplo de discursos, debates em mídias digitais, textos históricos, e outras formas de produção, como as transcrições de entrevistas para análises posteriores.

Em Linguística de *Corpus*, estas bases de dados textuais são objetos de pesquisa chamadas de *Corpus*. *Corpora* é o plural de *corpus* – conjunto de dados linguísticos pertencentes ao uso oral ou escrito da língua e que podem ser processados por computador. Contamos com suporte da tecnologia na Linguística de *Corpus* para potencializar as análises, usando ferramentas como concordanciadores, *corpora on-line*, programas de análise e comparação, dentre outros.

O pesquisador que introduziu a LC no Brasil, Berber Sardinha (2004), defende que muitas das descobertas e dos avanços de estudos linguísticos só foram (e só são) possíveis por meio de uma quantidade significativa de evidências da língua em uso – cujo acesso é facilitado pela LC. Dessa forma, a possibilidade de manipulação e de extração de dados estatísticos de uma grande quantidade de textos simultaneamente permite que linguistas trabalhem e observem manifestações com amostras verossímeis do uso da língua. É válido lembrar que podemos considerar a abordagem da LC como empírica da linguagem, uma vez que lida com a linguagem como um sistema probabilístico. Dessa forma, a teorização a partir das análises promovidas por ferramentas de LC a partir de *corpora* são *a posteriori*.

Estudiosos da LC defendem que a linguagem pode ser compreendida e estudada como um sistema padronizado que se manifesta por meio de repetição, de frequência e de recorrência de instâncias e que para teorizações e observações sejam válidas, cada *corpus* de estudo deve respeitar determinados critérios que, para Berber Sardinha (2004) se resumem em:

- Origem: textos autênticos e compostos por falantes nativos;
- Propósito: os dados devem ser objetos de estudo linguístico;
- Composição: o *corpus* deve ser composto por meio de critérios pré-estipulados;
- Formatação: os textos componentes da amostra devem ser adequados para leitura computadorizada;
- Extensão: para uma representação verossímil de manifestação de uma determinada língua, a quantidade de textos deve ser abrangente.

As ferramentas de LC, dentre outras possibilidades, permitem descrições de funcionamento da linguagem, observações de colocações para o desenvolvimento de dicionários técnicos, composição de listas de concordâncias e combinações, etiquetagem de palavras etc.

4.3 UTILIZAÇÃO NA PESQUISA

Durante o desenvolvimento desta pesquisa, entendemos que embora nossa análise seja primordialmente de forma individualizada em cada texto (apoiada por uma ferramenta de PLN), os princípios estatísticos e probabilísticos da LC darão sustento a observação de padrões da língua que irão colaborar para o levantamento de traços de complexidade levando em consideração cada grupo de textos pertencentes a cada nível de aprendizagem de ILA.

Além disso, a LC nos dará suporte na análise da complexidade lexical ao utilizarmos o *software* AntConc⁴⁰ para gerar listas de palavras de cada um dos textos. Com listas formadas com o auxílio dessa ferramenta, poderemos fazer uma análise individualizada dos vocábulos nos textos indicados como complexos de acordo com métricas que abordem léxico.

Portanto, em determinados momentos, observaremos cada texto de forma isolada e utilizaremos os dados para compreender quais são as características que podem tornar cada um dos materiais analisados potencialmente complexos para os estudantes; em outros momentos, iremos comparar os dados numéricos para observar se determinada frequência métrica indica padronização nas indicações de complexidade ou não.

5 CORPORA DA PESQUISA

Esta pesquisa, que é quantitativa e qualitativa, será realizada com o apoio de três *corpora*:

1. Um *corpus* progressivo, formado por vinte textos redigidos por nós que apresentam progressão controlada de complexidade;

⁴⁰ Ferramenta de Linguística de *Corpus* que, entre outras possibilidades, permite submeter textos e formar lista de palavras provenientes deles. Disponível em <https://www.laurenceanthony.net/software/antconcl/>

2. Um *corpus* comparativo, composto por vinte textos simples e vinte textos complexos em língua inglesa e;
3. O *corpus* de estudo (textos principais da nossa análise que estão nos livros *Elementary* e *Pre-intermediate* da coleção *English File*, sobre o qual falaremos posteriormente).

Os *corpora* comparativo e progressivo são utilizados para dois principais motivos: para observação do comportamento das métricas escolhidas (sobre as quais falaremos adiante) em relação à interpretação de complexidade e para análise contrastiva dos índices desses materiais com os valores dos textos do nosso *corpus* de estudo. Esse contraste nos permitirá observar se as métricas do *corpus* de estudo se aproximam mais de textos simples ou dos materiais considerados complexos.

Neste capítulo, descrevemos e explicamos os textos que são parte dos nossos *corpora*, iniciando pelo *corpus* de estudo.

5.1 CORPUS DE ESTUDO

Esta pesquisa tem como principal *corpus* de estudo os textos utilizados para atividades de interpretação de leitura de todas as unidades dos livros *Elementary* e *Pre-intermediate* da quarta edição da coleção *English File* (Figura 2) da editora *Oxford University Press*. Textos com exercícios de completar lacunas, ou pequenos textos que não tenham finalidades de observar compreensão (ou seja, que não se enquadrem nas características de ‘leitura para informação’ definida pelo CEFR) não serão utilizados.

Figura 2 – Capa dos livros *English file 4th edition*



Fonte: *Oxford University Press* (2020)

Ainda que não seja o foco da nossa pesquisa, também coletaremos as métricas do livro *Intermediate* para apresentar uma análise prévia (mas não tão detalhada) da complexidade desses materiais. Essa análise prévia nos permitirá verificar não só traços que possam ser comuns de complexidade entre os três níveis, mas também a progressão de complexidade.

Optamos por estudar apenas os livros *Elementary* e *Pre-intermediate* da coleção por dois motivos principais:

I. São os níveis nos quais traços de alta complexidade podem impedir a compreensão e desservir a manutenção linguística (visto que, de acordo com as diretrizes do CEFR apresentadas previamente, esses são os níveis cujas habilidades de leitura são mais limitadas e menos “naturais”);

II. São os livros com os quais nossos alunos costumam apresentar dificuldades de interpretação em determinados textos;

Já a análise prévia do livro *Intermediate* será realizada de forma mais globalizada apenas para verificar se ainda existem traços potencialmente complexos para esse nível – visto que, a partir do nível B1, as diretrizes do CEFR indicam maior autonomia e habilidade de interpretação leitora e, conseqüentemente, menores potenciais de traços complexos interferirem na interpretação e manutenção linguística. Logo, não apresentaremos, para o nível intermediário, uma análise detalhada e com exemplos dos traços de complexidade nos textos, mas de forma geral e que nos permita verificar se, futuramente, é válido expandir o estudo para os demais níveis, especialmente para alunos neuroatípicos.

Além disso, também utilizaremos os valores métricos dos textos do livro *Intermediate* para observar se há uma progressão de complexidade coerente e lógica entre diferentes níveis do texto (análise que independe de um estudo detalhado das indicações dos índices textuais).

Optamos por não analisar o livro *Beginner* uma vez que os textos presentes são, em sua maioria, curtos e/ou em forma de diálogo e, também, por não serem utilizados nos nossos contextos como docentes na universidade onde lecionamos.

De forma similar, escolhemos não utilizar os livros *Upper intermediate* e *Advanced* por dois motivos: por ser um material aplicado em níveis de alunos [quase] avançados (para os quais a análise de complexidade textual não seria de grande valia) e por termos turmas enxutas na universidade nas quais utilizamos estes livros.

Os livros da coleção *English File* foram escolhidos pela credibilidade e qualidade dos materiais da editora Oxford, por serem globalmente⁴¹ conhecidos e utilizados por professores de inglês como língua adicional, por serem versões atualizadas (quarta e mais recente edição da coleção) e por nosso uso e experiência com os materiais em sala de aula universitária.

Os conteúdos gramaticais e vocabulários trabalhados em cada um dos livros estão resumidos no quadro abaixo (Quadro 6)⁴²:

Quadro 6 – Resumo dos livros *English File 4th edition*

Livro:	<i>Elementary</i>
Nível do CEFR:	A2
Quantidade de textos:	20
Conteúdos gramaticais:	<i>Verb to be Subject pronouns Possessive adjectives Plurals Imperative and let's Present Simple Word order in questions Genitive case Prepositions of time Position of adverbs Can/can't Present Continuous Object pronouns Past Simple (regular and irregular) There is/are Quantifiers There was/were Countable and uncountable nouns Articles Comparatives Superlatives Be going to Verbs + infinitives Verbs + gerunds Present Perfect</i>
Vocabulário:	<i>Adjectives Adverbs of manner Days of the week Numbers Countries Classroom language Personal objects Feelings Verb phrases Jobs Family Routine verbs Question words Months Adverbs of frequency Seasons and Time Narrative words Dates Ordinal numbers Music</i>
Livro:	<i>Pre-intermediate</i>
Nível do CEFR:	A2+
Quantidade de textos:	22
Conteúdos gramaticais:	<i>Present Simple Present Continuous Past Simple Past Continuous Time sequencers and connectors Be going to Defining relative clauses Present Perfect Comparatives (as...as) Superlatives + Present Perfect Quantifiers Will Infinitives Gerunds Modal verbs First conditional Second conditional Possessive pronouns Present Perfect Phrasal verbs Passive voice Used to So, Neither + Auxiliaries Past Perfect Reported Speech Questions without auxiliaries</i>
Vocabulário:	<i>Common verb phrases describing appearance and personality Clothes Prepositions of place Holidays Airports Paraphrase language Housework Adjectives ending in -ing and -ed Types of numbers Describing cities Health and body Opposite verbs Verb + back Modifiers Verbs + infinitive Verb + Gerund Adjectives + prepositions Get Confusing verbs Adverbs of manner Animals and insects Fears Biography language Sports</i>

⁴¹ https://elt.oup.com/feature/global/english_file/

⁴² Mais informações sobre os livros podem ser encontradas neste link: https://elt.oup.com/feature/global/english_file/?cc=br&selLanguage=pt

	<i>Expressing movement</i> <i>People from different countries</i> <i>School Subjects</i> <i>Noun formation</i> <i>Similarities and differences</i> <i>Time expressions</i> <i>Say and tell</i>
Livro:	<i>Intermediate</i>
Nível do CEFR:	B1
Quantidade de textos:	21
Conteúdos gramaticais:	<i>Present Simple and Continuous</i> <i>Action and non-action verbs</i> <i>Future forms</i> <i>Present Perfect and Past Simple</i> <i>Present Perfect Continuous</i> <i>Choosing between comparatives and superlatives</i> <i>Articles and no article</i> <i>Modal verbs</i> <i>Past Tenses</i> <i>Passive Voice</i> <i>First Conditional and future time clauses</i> <i>Second conditionals</i> <i>Gerunds and Infinitives</i> <i>Reported Speech (sentences and questions)</i> <i>Third conditional</i> <i>Quantifiers</i> <i>Relative Clauses</i> <i>Question tags</i>
Vocabulário:	<i>Food and cooking</i> <i>Family and adjectives or personality</i> <i>Money</i> <i>Strong adjectives</i> <i>Transport</i> <i>Collocations</i> <i>Phone language</i> <i>Adjectives ending in -ing or -ed</i> <i>Sports</i> <i>Relationships</i> <i>Cinema</i> <i>The body</i> <i>Education</i> <i>Houses</i> <i>Work</i> <i>Shopping</i> <i>Nouns from verbs</i> <i>Adjective and adverb formation</i> <i>Electronic devices</i> <i>Compound nouns</i> <i>Crime</i>

Fonte: Elaboração própria

É importante destacar que utilizaremos esses textos da forma original como foram redigidos, sem correções ou alterações que, por vezes, seriam necessárias (principalmente em relação ao uso de vírgulas).

Para fins de futuras consultas – uma vez que, por questões de direitos autorais, não podemos inserir os textos na íntegra nesta tese - o quadro abaixo (Quadro 7) relaciona as etiquetas dos textos utilizadas neste trabalho aos livros e páginas correspondentes:

Quadro 7 - Referências de livro e página do *corpus* de estudo

ELEMENTARY		PRE-INTERMEDIATE		INTERMEDIATE	
ETIQUETA	PÁGINA	ETIQUETA	PÁGINA	ETIQUETA	PÁGINA
E1	21	P1	8	I1	7
E2	23	P2	14	I2	17
E3	32	P3	16	I3	22
E4	37	P4	21	I4	26-27
E5	42-43	P5	26	I5	30-31
E6	46-47	P6	34	I6	36
E7	53	P7	37	I7	38
E8	55	P8	38	I8	42
E9	58-59	P9	43	I9	47
E10	62	P10	47	I10	48
E11	66	P11	53	I11	50
E12	67	P12	58	I12	56
E13	70	P13	66-67	I13	60

E14	83	P14	69	I14	63
E15	85	P15	74	I15	67
E16	86	P16	79	I16	69
E17	88	P17	85	I17	78
E18	91	P18	86-87	I18	80-81
E19	98-99	P19	90-91	I19	86
E20	101	P20	94	I20	96
		P21	97	I21	102-103
		P22	101		

Fonte: Elaboração própria

Embora saibamos que elementos extratextuais podem contribuir para a compreensão textual, para a parte da análise apoiada pela ferramenta Coh-Metrix, precisamos excluir imagens, quadros ou quaisquer elementos não-textuais. Entretanto, a análise desses elementos será realizada na parte holística da pesquisa.

Na próxima seção, apresentamos os *corpora* que utilizaremos para análises comparativas.

5.2 CORPUS PROGRESSIVO

O *corpus* progressivo (rótulo CP) foi inteiramente redigido por nós e contribuirá para que possamos visualizar o comportamento de algumas métricas por meio de textos mais controlados. Para isso, redigimos vinte parágrafos apresentando complexidade progressiva. O primeiro parágrafo é composto por apenas uma frase formada por artigo e substantivo, o segundo apresenta uma oração, e assim por diante. O quadro abaixo (Quadro 8) apresenta os textos e as explicações de progressão de complexidade. Vale salientar que nas descrições das estratégias para progredir a dificuldade do texto, adicionamos somente o que foi realizado naquele texto específico, as estratégias anteriormente aplicadas permanecem de forma cumulativa. Na seção de resultados, apresentaremos e discutiremos os indícios das métricas desses textos, que nos ajudarão a observar o comportamento dos índices com maiores detalhes e nos ajudarão a interpretar esses valores em relação à complexidade, similarmente ao que Silva (2018) propôs para algumas métricas.

Ainda que os textos do *corpus* progressivo sejam curtos, o propósito é podermos verificar as métricas de uma forma mais “microscópica”. Obviamente, textos mais longos terão impactos diferentes nas métricas. Textos curtos, por outro lado, nos permitirão ver o comportamento desses índices de forma mais simplificada.

Quadro 8 - Textos do *corpus progressivo*

ETIQUETA DE REFERÊNCIA NA TESE	TEXTO E DESCRIÇÃO DE PROGRESSÃO DE COMPLEXIDADE
CP1	<i>The girl.</i> Frase simples com um artigo e um substantivo.
CP2	<i>The girl lives.</i> <i>The girl lives.</i> <i>The girl lives.</i> Frase simples seguindo a ordem canônica (sujeito, verbo, objeto) repetida três vezes. Essa repetição é proposital para que métricas como SS pudesse apresentar valores válidos. Uma vez que essas duas métricas fazem uma análise de similaridade entre frases de um texto, precisa-se mais de uma frase. As três frases serem precisamente iguais tem por objetivo forçar essas métricas a demonstrar as valorações que apresentariam para tal fenômeno.
CP3	<i>The girl lives in Paris.</i> <i>The girl lives in Paris.</i> <i>The girl lives in Paris.</i> Frase simples seguindo a ordem canônica (sujeito, verbo, complemento) repetida três vezes.
CP4	<i>The girl lives in Paris.</i> <i>The boy lives in London.</i> Frases simples seguindo a ordem canônica, mas com sujeitos e complementos minimamente diferentes.
CP5	<i>The girl lives in Paris.</i> <i>She likes animals.</i> <i>The boy lives in London.</i> <i>He likes animals.</i> Adição de pronomes para substituir os sujeitos. Pronomes próximos aos referentes.
CP6	<i>The girl lives in Paris.</i> <i>She likes animals. She hates dogs.</i> <i>The boy lives in London.</i> <i>He likes animals. He hates cats.</i> Aumento de quantidade de frases e vocabulário.
CP7	<i>The girl lives in Paris.</i> <i>She likes animals. She hates dogs.</i> <i>The boy lives in London.</i> <i>He likes animals. He hates cats.</i> <i>They are friends.</i> Inserção de uma nova frase com um pronome que se refere a dois sujeitos.
CP8	<i>The girl lives in Paris.</i> <i>She likes animals. She hates dogs.</i> <i>The boy lives in London.</i> <i>He likes animals. He hates cats.</i> <i>They are friends.</i> <i>She ate pizza.</i> <i>He ate pancakes.</i> Inserção de novas frases com novos vocábulos e tempo verbal.
CP9	<i>The girl lives in Paris.</i> <i>She likes animals. She hates dogs.</i> <i>The boy lives in London.</i>

	<p><i>He likes animals. He hates cats. They are friends. She ate pizza yesterday. He ate pancakes two days ago.</i></p>
	Inserção de complementos nas frases com novos vocábulos.
CP10	<p><i>The girl lives in Paris. The boy lives in London. She likes animals. He likes animals. She hates dogs. He hates cats. They are friends. She ate pizza yesterday. He ate pancakes two days ago.</i></p>
	Modificação da ordenação das frases para afastar os pronomes dos sujeitos.
CP11	<p><i>The girl lives in Paris. The boy lives in London. She likes animals but hates dogs. He likes the animals but hates cats. They are friends. She ate pizza yesterday, and he ate pancakes two days ago.</i></p>
	Inserção de conjunção adversativa e aditiva.
CP12	<p><i>The girl lives in Paris, but the boy lives in London. She likes animals but hates dogs. He likes animals but hates cats. They are friends. She ate pizza yesterday, and he ate pancakes two days ago.</i></p>
	Coordenação de frases.
CP13	<p><i>The girl lives in Paris, but the boy lives in London. However, they are friends. Both of them like animals, but she hates dogs, and he hates cats. She ate pizza yesterday, and he had pancakes two days ago.</i></p>
	Coordenação de frases e inserção de advérbio de oposição.
CP14	<p><i>The girl lives in the capital of France, but the boy lives in the capital of England. However, they are friends and in love. Both of them like animals, but she hates dogs, and he hates cats. She ate pizza yesterday, and he had pancakes two days ago.</i></p>
	Adição de informações.
CP15	<p><i>The girl lives in the capital of France, but the boy lives in the capital of England. However, they are friends and in love. Both of them like animals, but she hates dogs, and he hates cats. The female person is going to study Neurobiology, and the male one is going to study Phlebology. They are going to be in college. She ate pizza yesterday, and he had pancakes two days ago.</i></p>
	Substituição de nomes por conjunção substantiva ou sinónimo de maior complexidade.
	Inserção de novo tempo verbal.
	Inserção de termos.
CP116	<p><i>The girl lives in the capital of France, but the boy lives in the capital of England. However, they are friends and in love, and they want to meet up. Both of them like animals, but she hates dogs, and he hates cats – which may be a problem if they get together. The female person is going to study Neurobiology, while the male one, Phlebology. They are going to be in college. She ate pizza yesterday, and he had pancakes two days ago.</i></p>
	Modificação de pontuação.

	Inserção de outras frases coordenadas.
CP17	<p><i>The girl lives in the capital of France, but the boy lives in the capital of England. However, they are friends and in love, and they want to meet up.</i></p> <p><i>Both of them like animals, but she hates dogs, and he despises cats – which may be a problem if they get together.</i></p> <p><i>The female person is going to study Neurobiology, while the male one, Phlebology. They are going to be in college.</i></p> <p><i>She has always been curious about the brain, and she wants to learn about the frontal lobe, the parietal lobe, the temporal lobe, the spinal cord, the occipital lobe, and the cerebellum.</i></p> <p><i>He has always desired to acquire knowledge on diseases associated with veins and venous health and wants to specialize in the differences between veins and arteries. He is willing to study the superior vena cava, the azygos, the left brachiocephalic, the superior epigastric, and the hemiazygos.</i></p> <p><i>She ordered and ate pizza yesterday; and he had pancakes two days ago in a nice restaurant.</i></p> <p>Inserção de frases subordinadas. Adição de mais termos. Os termos inseridos foram especificamente selecionados de áreas nas quais essas unidades teriam poucas possibilidades de assumir outros significados. Optamos por utilizar termos de áreas médicas para evitar que o Coh-Matrix considerasse como palavras simples alguns termos de áreas que o estatuto de termo é dado pela inserção em um texto especializado (como no caso da palavra “planta” que assume o valor de termo ao ser utilizado em um texto de Arquitetura, por exemplo). Alteração de estruturação de elementos.</p>
CP18	<p><i>This is the tale of two people who will end up having their lives merged.</i></p> <p><i>There is a girl, who lives in the beautiful capital of France, Paris; and a boy who resides in the busy capital of England, London. However, they are friends and in love, and they want to meet up to start a relationship.</i></p> <p><i>These two lovebirds have lots of common interests. Both of them like animals, but she hates dogs, and he hates cats – which may be a problem if they get together. Nevertheless, even though they dislike these species of animals, they are not adamant about not having them.</i></p> <p><i>Jaida, the girl, is going to study Neurobiology, while the boy, Thomas, Phlebology. They are going to be in college next year. They are trying to find a university that offers both majors, so they can live in the same city.</i></p> <p><i>She has always been curious about the brain, and she wants to learn about the frontal lobe, the parietal lobe, the temporal lobe, the spinal cord, the occipital lobe, and the cerebellum. She has researched a lot about acoustic neuromas, acromegaly, agnosia, arteriovenous malformation, choroid plexus, and so on.</i></p> <p><i>He has always desired to acquire knowledge on diseases associated with veins and venous health and wants to specialize in the differences between veins and arteries. He is willing to study the superior vena cava, the azygos, the left brachiocephalic, the superior epigastric, and the hemiazygos. He has been studying telangiectasia, ablation, lymphedema, and phlebectomy.</i></p>

	<p><i>Yesterday, she met her friends in a pizzeria and had pizza. He had pancakes two days ago in a pleasant – but crowded – restaurant. They both had an amusing time.</i></p>
CP19	<p><i>Ampliação da estrutura e de informações.</i></p> <p><i>This is a stunning, though succinct, tale of two beings who will end up having their lives intertwined.</i></p> <p><i>There is a gal, who lives in the picturesque capital of France, Paris; and a young man who resides in the busy capital of England, London. Nonetheless, they are friends and nurturing loving feelings for one another, and they crave to meet up to begin a relationship.</i></p> <p><i>These two lovebirds have tons of common interests. Both of them appreciate the fauna, only she hates canines, and he despises felines – which may be a setback in case they get together. Nevertheless, even though they are unappreciative of these animal species, both are not adamant about not adopting either a puppy or a kitten; they may grow accustomed to them.</i></p> <p><i>Jaida, the young lady, is going begin her studies in Neurobiology, while the boy, Thomas, in Phlebology. They shall be in a prestige college in the year to come. They are pursuing to find a well-known university that offers both majors, so they can live in the same conurbation.</i></p> <p><i>She has always been enquiring about the brain, and wishes to learn about the frontal lobe, the parietal lobe, the temporal lobe, the spinal cord, the occipital lobe, and the cerebellum. She has researched a lot about acoustic neuromas (a benign tumor of the hearing nerve), acromegaly (a disorder marked by progressive enlargement of the head, face, hands, feet, and thorax, due to the excessive secretion of growth hormone), agnosia (the absence of the ability to recognize the form and nature of persons and things), arteriovenous malformation (a collection of blood vessels with one or several abnormal connections between arteries and veins, which may cause hemorrhage or seizures), choroid plexus (a vascular structure in the ventricles of the brain that produces cerebrospinal fluid), and so on.</i></p> <p><i>He has always desired to acquire knowledge on diseases associated with veins and venous health and wants to specialize in the differences between veins and arteries. He is willing to study the superior vena cava (a large, valveless vein that conveys venous blood from the upper half of the body and returns it to the right atrium), the azygos (unilateral vessel that ascends in the thorax to the right of the vertebral column, carrying deoxygenated blood from the posterior chest and abdominal walls), the left and right brachiocephalic (major veins in the upper chest, formed by the union of each corresponding internal jugular vein and subclavian vein), the superior epigastric (a blood vessel that carries oxygenated blood to the abdominal wall, and upper rectus abdominis muscle), and the hemiazygos (a vein that receives blood from the lower half of the left thoracic wall and the left abdominal wall, ascends along the left side of the spinal column, and empties into the azygos vein near the middle of the thorax). He has been studying telangiectasia (small, widened blood vessels on the skin), ablation (the surgical removal of body tissue), lymphedema (A condition in which extra lymph fluid builds up in tissues and causes swelling), and phlebectomy (a minimally invasive procedure used to remove varicose veins that lie just beneath the surface of the leg).</i></p> <p><i>Yesterday, she met her friends in a pizzeria and had her favorite dish, a pepperoni and cheese pizza. He had vegan pancakes two days ago</i></p>

	<p><i>in a pleasant – but crowded – restaurant. They both had an amusing time.</i></p> <p>Substituição de palavras por vocabulários mais complexos. Estruturas mais rebuscadas. Alta presença de termos e de apostos explicativos.</p>
CP20	<p><i>This is a stunning, though succinct, tale of two beings who'll end up having their lives intertwined.</i></p> <p><i>There's a gal, who lives in the picturesque capital of France, Paris; and a young man who resides in the busy capital of England, London. Nonetheless, they're friends and nurturing loving feelings for one another, and they crave to meet up to begin a relationship.</i></p> <p><i>These two lovebirds have tons of common interests. Both of them appreciate the fauna, only she hates canines, and he despises felines – which may be a setback in case they get together. Nevertheless, even though they're unappreciative of these animal species, both aren't adamant about not adopting either a puppy or a kitten; they may grow accustomed to them.</i></p> <p><i>Jaida, the young lady, is going begin her studies in Neurobiology, while the boy, Thomas, in Phlebology. They shall be in a prestige college in the year to come. They're pursuing to find a well-known university that offers both majors, so they can live in the same conurbation.</i></p> <p><i>She's always been enquiring about the brain, and wishes to learn about the frontal lobe, the parietal lobe, the temporal lobe, the spinal cord, the occipital lobe, and the cerebellum. She's researched a lot about acoustic neuromas (a benign tumor of the hearing nerve), acromegaly (a disorder marked by progressive enlargement of the head, face, hands, feet, and thorax, due to the excessive secretion of growth hormone), agnosia (the absence of the ability to recognize the form and nature of persons and things), arteriovenous malformation (a collection of blood vessels with one or several abnormal connections between arteries and veins, which may cause hemorrhage or seizures), choroid plexus (a vascular structure in the ventricles of the brain that produces cerebrospinal fluid), and so on.</i></p> <p><i>He's always desired to acquire knowledge on diseases associated with veins and venous health and wants to specialize in the differences between veins and arteries. He's willing to study the superior vena cava (a large, valveless vein that conveys venous blood from the upper half of the body and returns it to the right atrium), the azygos (unilateral vessel that ascends in the thorax to the right of the vertebral column, carrying deoxygenated blood from the posterior chest and abdominal walls), the left and right brachiocephalic (major veins in the upper chest, formed by the union of each corresponding internal jugular vein and subclavian vein), the superior epigastric (a blood vessel that carries oxygenated blood to the abdominal wall, and upper rectus abdominis muscle), and the hemiazygos (a vein that receives blood from the lower half of the left thoracic wall and the left abdominal wall, ascends along the left side of the spinal column, and empties into the azygos vein near the middle of the thorax). He has been studying telangiectasia (small, widened blood vessels on the skin), ablation (the surgical removal of body tissue), lymphedema (A condition in which extra lymph fluid builds up in tissues and causes swelling), and phlebectomy (a minimally invasive procedure used to remove varicose veins that lie just beneath the surface of the leg).</i></p>

	<i>Yesterday, she met her friends in a pizzeria and had her favorite dish, a pepperoni and cheese pizza. He had vegan pancakes two days ago in a pleasant – but crowded – restaurant. They both had an amusing time.</i>
	Além das modificações anteriores, foram aplicadas contrações de palavras.

Fonte: Elaboração própria

Como é possível observar no quadro acima, os textos apresentam traços de complexidade de forma progressiva. O primeiro texto é simples (não só por extensão, mas também nos níveis lexicais, semânticos e sintáticos) e o último, complexo em todos os traços. Conforme já mencionado, a redação dos textos ter sido nossa e as características de complexidade serem controladas nos ajudam a enxergar os comportamentos das métricas de forma mais minuciosa. Sendo assim, além do *corpus* comparativo (sobre os quais falaremos na próxima seção), também teremos referências pontuais que nos permitirão relacionar valores métricos com os elementos potenciais de complexidade estipulados por nós.

5.3 CORPUS COMPARATIVO

Com a finalidade de apresentar referências métricas para fins comparativos, desenvolvemos um *corpus* comparativo subdividido em dois: o primeiro composto por vinte textos em língua inglesa considerados previamente simples (cujo rótulo nos quadros desta tese é LIS – Língua Inglesa Simples) e o segundo, com vinte textos em língua inglesa considerados complexos (rótulo LIC – Língua Inglesa Complexos). Dessa forma, além de interpretações baseadas na organização classificatória proposta por Silva (2018) para algumas das métricas, também podemos utilizar os valores dos índices desses textos como régua. Iremos nos referir a esses conjuntos de textos como *corpus* comparativo e, quando uma diferenciação for necessária, nos referiremos aos textos LIS como textos simples do *corpus* comparativo e aos textos LIC como textos complexos do *corpus* comparativo.

Os textos do *corpus* comparativo e um cruzamento com as informações do *corpus* *progressivo* nos permitirão organizar os valores numéricos de algumas das métricas escolhidas para este trabalho que não tenham sido previamente classificadas por Silva (2018) em relação a possíveis interpretações de complexidade.

Os textos considerados previamente simples do *corpus* comparativo são provenientes de sites que apresentam textos simplificados para crianças: *British Council*⁴³, *Free Kids Books*⁴⁴, *English Club*⁴⁵ e *Mom loves best*⁴⁶. Escolhemos esses textos por serem sites conhecidos e confiáveis e porque adotam padrões e diretrizes apontadas pela literatura a respeito de acessibilidade textual. Além disso, acreditamos que por serem textos voltados para crianças, poderão ser mais próximos da realidade de leitores que estejam em fases iniciais de desenvolvimento linguístico.

Levando em consideração o Princípio da Alcançabilidade – que defende que se um texto é acessível para um leitor com menor proficiência, será também para leitores com maiores proficiências (SILVA, 2020) – e tendo em vista que os textos do *corpus* de estudo são escritos em inglês para estudantes brasileiros (falantes de língua portuguesa como nativa), acreditamos que:

I. Os valores dos índices dos textos em língua inglesa considerados simples (uma vez que são textos autênticos⁴⁷ e acessíveis para crianças nativas do inglês em período de alfabetização) servirão como referência média para observarmos a aproximação ou distância dos valores encontrados nos textos do *corpus* de estudo. Acreditamos que para aprendizes de inglês como língua adicional, as métricas dos textos do livro *Elementary* devem ser similares ou inferiores (na interpretação de complexidade) dos textos simples em inglês. Para os níveis *pre-intermediate* e *intermediate*, cremos que os valores devam se distanciar, progressivamente, das referências;

II. Os valores dos textos complexos em língua inglesa podem ser utilizados como parâmetro para a classificação de interpretação de complexidade das métricas, similarmente ao que foi realizado por Silva (2018) para a TT. Além disso, os índices desses textos podem nos indicar complexidade dos textos do *corpus* de estudo quando seus valores forem próximos.

Os textos em língua inglesa considerados simples (doravante LIS)⁴⁸ estão apresentados e brevemente descritos no quadro a seguir (Quadro 9). Os textos na forma integral estão no apêndice desta tese.

⁴³ <https://learnenglishkids.britishcouncil.org/>

⁴⁴ <https://freekidsbooks.org/>

⁴⁵ <https://www.englishclub.com/kids/stories/>

⁴⁶ <https://momlovesbest.com/short-moral-stories-kids>

⁴⁷ Por “textos autênticos”, referimo-nos a textos escritos para os falantes nativos de uma determinada língua. Esse termo se difere de “textos didáticos”.

⁴⁸ Os textos de forma original e os links de referência estarão disponíveis no apêndice deste trabalho.

Quadro 9 - Descrição dos textos simples do *corpus* comparativo

ETIQUETA DE REFERÊNCIA NA TESE	TÍTULO ORIGINAL	DESCRIÇÃO
LIS1	<i>Stop wasting food!</i>	Fonte: <i>Learn English Kids – British Council</i> Texto simples escrito para crianças. Frases e parágrafos curtos. Léxico simplificado e limitado.
LIS2	<i>Digital citizenship.</i>	Fonte: <i>Learn English Kids – British Council</i> Texto simples escrito para crianças. Frases e parágrafos curtos. Léxico simplificado e limitado. Elencado em etapas com títulos.
LIS3	<i>The Chickens Take a Holiday.</i>	Fonte: <i>English Club</i> Texto simples escrito para crianças pequenas (<i>toddlers</i>). Narrativa curta com desencadeamento lógico e manutenção tema-rema simples. Frases e parágrafos curtos. Léxico simplificado e limitado. Composto basicamente por diálogos e frases descritivas.
LIS4	<i>Linda's picture.</i>	Fonte: <i>Learn English Kids – British Council</i> Texto simples escrito para crianças e por uma criança (personagem). Frases e parágrafos curtos. Léxico simplificado e limitado e aproximado da fala oral. Texto descritivo.
LIS5	<i>Frida Kahlo</i>	Fonte: <i>English Club</i> Texto simples escrito para crianças. Léxico simplificado e limitado. Organizado em tópicos com título. Texto descritivo. Texto biográfico.
LIS6	<i>Too tiny for tea.</i>	Fonte: <i>English Club</i> Texto simples escrito para crianças pequenas (<i>toddlers</i>). Narrativa curta com desencadeamento lógico e manutenção tema-rema simples. Frases e parágrafos curtos. Léxico simplificado e limitado. Composto basicamente por diálogos e frases descritivas.
LIS7	<i>Holi's diary.</i>	Fonte: <i>Learn English Kids – British Council</i>

		<p>Texto simples escrito para crianças. Frases e parágrafos curtos. Léxico simplificado e limitado. Simulação de linguagem infantil utilizada em diário e uso de primeira pessoa. Organizado por títulos/horários.</p>
LIS8	<i>Birthday on the farm.</i>	<p>Fonte: <i>Learn English Kids – British Council</i> Texto simples escrito para crianças pequenas (<i>toddlers</i>). Frases-parágrafo curtas. Léxico simplificado (holisticamente, o mais simples entre todos os textos).</p>
LIS9	<i>Echo and Hera.</i>	<p>Fonte: <i>Learn English Kids – British Council</i> Texto simples escrito para crianças. Frases e parágrafos curtos. Léxico simplificado e limitado. Texto descritivo e narrativo.</p>
LIS10	<i>My father meets the cat.</i>	<p>Fonte: <i>Free kids Books</i> Texto simples escrito para crianças. Frases e parágrafos relativamente curtos. Léxico simplificado e limitado. Texto narrativo.</p>
LIS11	<i>Let's meet these three friends?</i>	<p>Fonte: <i>ISL Collective</i> Texto simples escrito para crianças. Texto altamente simples em estruturação sintática e léxico. Escrito para crianças. Frases e parágrafos curtos. Texto narrativo.</p>
LIS12	<i>I am Steve.</i>	<p>Fonte: <i>ISL Collective</i> Texto simples escrito para crianças. Texto altamente simples em estruturação sintática e léxico. Escrito para crianças. Frases e parágrafos curtos. Texto narrativo.</p>
LIS13	<i>The boy who cried wolf.</i>	<p>Fonte: <i>Mom Loves Best</i> Texto simples escrito para crianças. Frases e parágrafos relativamente curtos. Léxico simplificado e limitado. Texto narrativo. Estruturas sintáticas relativamente simples, mas com presença de coordenação e subordinação e de deslocamento de elementos.</p>
LIS14	<i>The golden touch</i>	<p>Fonte: <i>Mom Loves Best</i> Texto simples escrito para crianças. Frases e parágrafos relativamente curtos.</p>

		Léxico simplificado e limitado. Texto narrativo.
LIS15	<i>The Fox and the Grapes</i>	Fonte: <i>Mom Loves Best</i> Texto simples escrito para crianças. Frases e parágrafos relativamente curtos. Léxico simplificado e limitado. Texto narrativo. Utilização de past e present participle com omissão de sujeito (previamente mencionado).
LIS16	<i>The Proud Rose</i>	Fonte: <i>Mom Loves Best</i> Texto simples escrito para crianças. Frases e parágrafos relativamente curtos. Léxico simplificado e limitado. Texto narrativo. Estruturas sintáticas simples, mas com presença de coordenação e subordinação e de deslocamento de elementos.
LIS17	<i>The Milkmaid and her Pail</i>	Fonte: <i>Mom Loves Best</i> Texto simples escrito para crianças. Frases e parágrafos com extensões medianas. Léxico relativamente simples, mas variado e com a presença de algumas palavras mais complexas. Texto narrativo. Estruturas sintáticas medianas, com presença de coordenação e subordinação, deslocamento de elementos e utilização de pontuações distintas.
LIS18	<i>A Wise Old Owl</i>	Fonte: <i>Mom Loves Best</i> Texto simples escrito para crianças. Frases e parágrafos relativamente curtos. Léxico simplificado e limitado. Texto narrativo. Estruturas sintáticas simples, mas com presença de coordenação e subordinação e de deslocamento de elementos.
LIS19	<i>The Golden Egg</i>	Fonte: <i>Mom Loves Best</i> Texto simples escrito para crianças. Frases e parágrafos relativamente curtos. Léxico simplificado e limitado. Texto narrativo. Estruturas sintáticas simples.
	<i>The Farmer and the Well</i>	Fonte: <i>Mom Loves Best</i> Texto simples escrito para crianças.

LIS20		<p>Frases e parágrafos com extensões medianas.</p> <p>Léxico relativamente simples, mas variado e com a presença de algumas palavras mais complexas.</p> <p>Texto narrativo.</p> <p>Estruturas sintáticas medianas, com presença de coordenação e subordinação, deslocamento de elementos e utilização de pontuações distintas.</p>
-------	--	---

Fonte: Elaboração própria

Os textos em língua inglesa considerados complexos (doravante LIC), por sua vez, foram coletados:

- I. De fonte acadêmica e ou Técnico-científicas: LIC1, LIC3, LIC5, LIC6, LIC7, LIC8, LIC9 e LIC10;
- II. De provas de proficiência de língua inglesa em nível C2: LIC2 e LIC4;
- III. Do livro *Advanced Plus* da mesma coleção dos textos do nosso *corpus* de estudo: de LIC11 a LIC20.

Os textos de fontes acadêmicas e técnico-científicas foram escolhidos por serem tipologias textuais considerados altamente complexos por sua linguagem rebuscada, alta terminologia e estruturas sintáticas complexas. Os materiais de provas de proficiência, por serem classificados como leituras para falantes do nível C2 (mais alto) do CEFR. Por fim, os textos do livro *Advanced Plus* foram selecionados por estarem presentes na mesma coleção que os nossos textos do *corpus* de estudo, mas no livro do nível mais avançado. Sendo assim, podemos comparar as métricas entre os materiais que são foco da nossa análise com os valores dos índices de textos que pertencem ao mesmo “universo”, mas com alta complexidade.

Para buscar a maior fidedignidade possível à extensão dos textos simples do *corpus* comparativo e dos textos originais do nosso *corpus* de estudo, em textos acadêmicos e científicos longos, selecionamos apenas uma seção completa do texto integral; uma vez que os textos científicos na íntegra seriam extensos demais e, conseqüentemente, apresentariam resultados correspondentes ao seu tamanho. Esses textos estão brevemente descritos no quadro abaixo (Quadro 10):

Quadro 10 - Descrição dos textos complexos do *corpus* comparativo

ETIQUETA DE REFERÊNCIA NA TESE	TÍTULO ORIGINAL	DESCRIÇÃO
LIC1	<i>Derivation by phase</i>	Texto acadêmico-científico de Linguística. Altamente terminológico. Sintaticamente complexo. Léxico especializado. Frases e parágrafos longos.
LIC2	<i>CPE Reading C2</i>	Texto narrativo escrito para prova de proficiência de leitura em nível C2 (avançado). Léxico complexo. Frases e parágrafos longos. Sintaticamente complexo.
LIC3	<i>Secondary infection with Streptococcus pneumoniae decreases influenza virus replication and is linked to severe disease</i>	Texto acadêmico-científico de Microbiologia. Altamente terminológico. Sintaticamente complexo. Léxico especializado. Frases e parágrafos longos.
LIC4	<i>Flex your funny bone</i>	Texto de business inglês utilizado para prova de proficiência em nível C2. Altamente terminológico. Com muitas expressões específicas da área. Frases e parágrafos longos.
LIC5	<i>Characterization of synthesized nanoparticles</i>	Texto acadêmico-científico de Microbiologia. Altamente terminológico. Sintaticamente complexo. Léxico especializado. Frases e parágrafos longos.
LIC6	<i>Medication use and risk of amyotrophic lateral sclerosis—a systematic review</i>	Texto acadêmico-científico de Medicina. Altamente terminológico. Sintaticamente complexo. Léxico especializado. Frases e parágrafos longos.
LIC7	<i>Raman microspectroscopy fingerprinting of organoid differentiation state</i>	Texto acadêmico-científico de Microbiologia. Altamente terminológico. Sintaticamente complexo. Léxico especializado. Frases e parágrafos longos.
LIC8	<i>Extracellular membrane vesicles and nanotubes in Archaea</i>	Texto acadêmico-científico de Microbiologia. Altamente terminológico. Sintaticamente complexo. Léxico especializado. Frases e parágrafos longos.
LIC9	<i>A brain mechanism underlying the evolution of anxiety</i>	Texto acadêmico-científico de Neurologia. Altamente terminológico. Sintaticamente complexo. Léxico especializado. Frases e parágrafos longos.
LIC10	<i>Aircraft Maintenance Check Scheduling Using Reinforcement Learning</i>	Texto instrutivo sobre aviação. Altamente terminológico. Sintaticamente complexo. Léxico especializado.

		Frases e parágrafos longos.
LIC11	<i>My most inspiring teacher</i> (Página 15)	<p>Todos os textos apresentam estrutura sintática complexa, com diversos exemplos de coordenação e subordinação, elipses, zeugmas e deslocamento de elementos.</p> <p>No que diz respeito ao léxico, todos os textos também apresentam vocabulário complexo, embora não podendo ser considerado terminológico. Os textos utilizam vocabulários do cotidiano, mas levando em consideração contextos mais amplos que exigem diferentes registros linguísticos.</p> <p>Além disso, alguns textos se assemelham à linguagem jornalística e, dessa forma, são ricos em sinônimos.</p> <p>A variedade lexical também é grande em todos os textos.</p> <p>Embora sejam distintos do gênero textual dos textos entre LIC1 e LIC10, esses do livro Upper Advanced foram escritos para alunos com níveis próximos do C2 no CEFR e são, portanto, os mais complexos em sintaxe, semântica e léxico dentro da coleção <i>English File</i>.</p>
LIC12	<i>Measures to prevent mosquito bites</i> (Página 25)	
LIC13	Sem título (Página 26)	
LIC14	<i>How people in high-pressure jobs cope with stress...</i> (Páginas 48 e 49)	
LIC15	Sem título (Página 55)	
LIC16	<i>A millennial goes back to the eighties</i> (Páginas 56 e 57)	
LIC17	<i>Opinions on Renzo Piano's Shard, London</i> (Página 65)	
LIC18	<i>Why do we care whodunnit?</i> (Páginas 68 e 69)	
LIC19	<i>Food Memories</i> (Página 76)	
LIC20	<i>A Plate of Peas</i> (Página 85)	

Fonte: Elaboração própria

Os valores das métricas serão apresentados e discutidos na seção 9.1: Análise semiautomática do *corpus* comparativo.

No próximo capítulo, descrevemos a ferramenta utilizada para a coleta dos índices e as métricas elegidas para a execução deste trabalho.

6 FERRAMENTA E MÉTRICAS

Neste capítulo, apresentaremos a ferramenta de PLN que utilizamos para coletar os índices textuais (Coh-Metrix) e as métricas adotadas para interpretação de potencial de complexidade dos textos.

6.1 COH-METRIX

O Coh-Metrix⁴⁹ (abreviação de *Cohesion Metrics*) é uma ferramenta de PLN que permite a coleta de métricas referentes a diferentes níveis textuais (lexicais, sintáticos etc.) que podem ser utilizadas para diversos fins de pesquisas linguísticas.

McNamara, et al. (2014) explicam que o Coh-Metrix foi desenvolvido tendo a principal finalidade de facilitar a análise de textos por meio de múltiplas características e níveis da língua e do discurso – uma vez que os índices textuais escolhidos para o *software* se alinham naturalmente com teorias de texto e de discurso.

Ainda que a ferramenta não realize uma análise automática de complexidade textual apontando qual o potencial de complexidade de cada nível de um texto, a interpretação das métricas (seja de forma isolada ou por cruzamento de informações) permite estimativas de complexidade que vêm sendo utilizadas para pré-determinar se um texto poderá ser possivelmente complexo para alguns tipos de leitores. A partir dessas informações provenientes das métricas, linguistas podem corroborar ou refutar potenciais de complexidade com leitores reais e/ou adotar estratégias de acessibilidade que visem adequar um texto para um determinado protótipo de leitor.

O Coh-Metrix é uma ferramenta gratuita e que permite a coleta de cento e oito métricas textuais⁵⁰. As diretrizes do programa apontam que:

Esses valores podem ser usados em muitas diferentes maneiras para investigar a coesão do texto explícito e a coerência da representação mental do texto.

Nossa definição de coesão consiste nas características do texto explícito que cumprem um certo papel em ajudar o leitor a conectar, mentalmente, algumas ideias no texto (GRAESSER, MCNAMARA, e LOUWERSE, 2003).

A definição de coerência é tópico de muito debate. Em teoria, a coerência de um texto é definida pelo número de interações entre representações linguísticas e representações de conhecimentos.

Quando colocamos o foco no texto, contudo, coerência pode ser definida como características do texto (por exemplo, aspectos de coesão) que são prováveis de contribuir para a coerência de representações mentais. O Coh-Metrix fornece indícios de tais características de coesão.⁵¹

⁴⁹ <http://Coh-Metrix.com/>

⁵⁰ Número de métricas disponíveis na versão 3.0 (acessado em 2022)

⁵¹ Tradução nossa. Original: “These values can be used in many different ways to investigate the cohesion of the explicit text and the coherence of the mental representation of the text. Our definition of *cohesion* consists of characteristics of the explicit text that play some role in helping the reader mentally connect ideas in the text (Graesser, McNamara, & Louwerse, 2003). The definition of *coherence* is the subject of much debate. Theoretically, the coherence of a text is defined by the interaction between linguistic representations and knowledge representations. When we put the spotlight on the text, however, coherence can be defined as characteristics of the text (i.e., aspects of

Ainda que existam outros recursos digitais que poderiam ser adotados para a condução desta pesquisa, optamos pelo Coh-Metrix pela quantidade de métricas disponíveis, pela confiabilidade do programa, por ser um *software* que tem sido utilizado para outros estudos de complexidade e acessibilidade textual e pela nossa familiaridade com a ferramenta.

Embora, conforme supracitado, a ferramenta não exiba resultados que digam diretamente se um texto é complexo ou simples, estudos vêm sendo realizados para observar o padrão entre a definição de “complexo” por leitores reais e os intervalos numéricos de cada índice disponíveis pelo Coh-Metrix. Silva (2018) analisou a complexidade de textos sobre o Transtorno do Estresse Pós-Traumático e com o apoio de um *corpus* comparativo e com as definições de cada um dos índices no documento normativo da ferramenta, classificou a valoração numérica de algumas métricas de forma que possam interpretar potencial de complexidade – sobre as quais falaremos mais adiante.

Nesta tese, além de utilizarmos a classificação de Silva (2018), também desenvolveremos os intervalos interpretativos de outras métricas não trabalhadas previamente pelo autor e que colaborarão para análise de potencial de complexidade.

Finatto (2011, p.11) categoriza os índices disponíveis no Coh-Metrix em 6 grupos:

- 1) Identificação Geral e Informação de Referência, Índices de Inteligibilidade, Palavras Gerais e Informação do Texto, Índices Sintáticos, Índices Referenciais e Semânticos e Dimensões do Modelo de Situações. Essa primeira classe corresponde às informações que referenciam o texto, como título, gênero entre outros;
- 2) Índices de inteligibilidade calculados com as fórmulas *Flesch Reading Ease* e *Flesch Kincaid Grade Level*. Essas fórmulas consideram tamanho de sentença, número de palavras por sentença e número de palavras diferentes por sentença;
- 3) Verificação de quatro subclasses: Contagens Básicas, Frequências, Concretude, Hiperônimos;
- 4) Verificação de cinco subclasses: Constituintes, Pronomes, Tipos e Tokens, Conectivos, Operadores Lógicos e Similaridade sintática de sentenças;
- 5) Verificação de três subclasses: Anáfora, Correferência e Análise Semântica Latente;
- 6) Verificação de quatro subclasses: Dimensão Causal, Dimensão Intencional, Dimensão Temporal e Dimensão Espacial.

Na próxima seção, explicaremos as métricas elegidas para esta pesquisa e suas possíveis interpretações de complexidade textual.

cohesion) that are likely to contribute to the coherence of the mental representation. Coh-Metrix provides indices of such cohesion characteristics.”

6.2 MÉTRICAS DE LEITURABILIDADE

Conforme mencionado na seção 2.3 (*Análises de Acessibilidade Textual*), pesquisas que se propõem a analisar CT devem fazer escolhas orientadas pelo leitor-determinado, tipologia textual e formas de análise que – vale lembrar – são multimodais. Dentre essas escolhas, o analista de CT precisa escolher as métricas de textualidade que melhor se adequem aos propósitos da sua pesquisa.

O propósito de pesquisas de CT multimodais é garantir que a interpretação das métricas não seja limitada a apenas um nível textual (caso o objetivo seja uma análise global dos textos) e que, dentro de cada nível escolhido para ser observado, a investigação não seja direcionada e reduzida a apenas uma métrica.

Vamos tomar como exemplo um texto de Medicina onde todas as frases seguem exatamente a mesma estrutura sintática, mas que o vocabulário seja altamente terminológico. Em uma análise de complexidade realizada com o apoio de métricas, um índice que seja utilizado para indicar similaridade sintática irá sugerir que o texto é simples, enquanto um índice de variação lexical ou outro de análise semântica irão nos destacar que, na verdade, este texto é complexo. Ao mesmo tempo, uma métrica de polissemia poderia indicar que o texto é simples por não haver grande presença de palavras polissêmicas (uma vez que termos da área médica têm, em grande maioria, valoração semântica única). Sendo assim, a utilização das métricas nos direciona para pontos que precisam de atenção e a verificação de um linguista dará uma corroboração mais precisa sobre a complexidade.

No caso do nosso exemplo, ao perceber que uma métrica de sintaxe e que uma de semântica e léxico (a da polissemia) indicaram simplicidade, enquanto duas métricas de semântica e léxico indicaram complexidade, um linguista precisa averiguar os níveis textuais indicados para chegar à conclusão de que o texto é, em realidade, potencialmente complexo por ser altamente terminológico e especializado, mas que a estrutura sintática é similar entre todas as ocorrências de frases (e, em consequência, a métrica de sintaxe compreendeu essa manifestação como indicativo de um texto simples) e que a grande presença de termos “enganou” a leitura da máquina em relação à quantidade de palavras polissêmicas.

Para o desenvolvimento do estudo desta tese, escolhemos métricas que estimem a complexidade lexical, semântica e sintática dos textos do nosso *corpus*.

Esses índices serão coletados por meio do Coh-Metrix (detalhado na seção 7.1) e seguirão as diretrizes do programa.

As métricas escolhidas para subsidiar a presente pesquisa estão relacionadas no quadro abaixo (Quadro 11) bem como o nível textual para qual utilizaremos cada uma delas para a observação dos potenciais de complexidade:

Quadro 11 - Métricas utilizadas para descrição e análise dos textos e de reescritas

MÉTRICA	SIGLA	NÍVEL TEXTUAL
Polissemia	PS	Semântica e Léxico
Idade de aquisição	IA	
Familiaridade com palavras de conteúdo	FPC	
Concretude lexical	CLP	
Relação type-token	TT	
Narratividade	NT	Sintaxe
Simplicidade sintática	SS	
Similaridade Sintática	SIMS	
Coesão Referencial	CR	
Índice Flesch	IF	

Fonte: Elaboração própria

Com a finalidade de explorar a maior quantidade possível de níveis textuais dos materiais voltados para o ensino, selecionamos as métricas descritas acima porque acreditamos que elas sejam suficientes para nos apresentar um panorama geral dos possíveis traços de complexidade nos níveis sintáticos, semânticos e lexicais.

Destacamos que durante o processo de aplicação da nossa pesquisa, estudamos e coletamos outras métricas que optamos por excluir da versão final. Entre essas métricas, podemos citar a Análise Semântica Latente (excluída por não demonstrar um padrão de comportamento claro relativo à complexidade textual) e a Leiturabilidade em L2 (que embora pareça condizente com os propósitos da pesquisa, não apresenta descrições sobre o funcionamento e tipo de análise no manual do Coh-Metrix – diferentemente de todas as outras métricas – e por não apresentar padrão de comportamento que permitisse inferir/compreender qual característica linguística utiliza como base).

A figura abaixo (Figura 3) ilustra quais métricas nos auxiliarão a analisar a sintaxe, a semântica e o léxico dos textos:

Figura 3 - Métricas e níveis textuais

Fonte: Elaboração própria

Como é possível perceber pela Figura 3, as métricas Narratividade (NT), Simplicidade Sintática (SS), Similaridade Sintática (SimS), Coesão Referencial (CR) e Índice Flesch (IF) nos permitirão observar a estruturação dos textos em nível sintático. Já as métricas Polissemia (PS), Idade de Aquisição (IA), Familiaridade com Palavras de Conteúdo (FPC), Concretude Lexical (CLP) e Relação *Type-Token* (TT) nos auxiliarão a estudar as indicações semânticas e lexicais dos nossos materiais de estudo.

É importante destacar que a maior parte dos índices providenciados por meio do Coh-Metrix não são voltados especificamente para análise de complexidade textual. Para isso, é preciso estudar o comportamento numérico de cada uma das métricas e realizar testes comparativos e estatísticos para identificar faixas de valoração que possam ser utilizadas para interpretação de potencial de complexidade, conforme explica Silva (2018).

Antes de explicarmos as métricas com detalhes e de apresentarmos nossos quadros classificativos de complexidade, explicaremos, a seguir, os procedimentos e critérios que nos auxiliaram a desenvolver os quadros classificatórios de potenciais de complexidade.

6.2.1 MÉTRICAS COMO SISTEMAS CLASSIFICATÓRIOS DE POTENCIAIS DE COMPLEXIDADE

Conforme mencionado anteriormente, a maior parte das métricas disponíveis no Coh-Matrix não é direcionada especificamente para estimativas de complexidade textual.

Com a finalidade de classificar as valorações de cada índice de forma eficiente para interpretações de complexidade - conforme Silva (2018) realizou para alguns índices - os seguintes procedimentos devem ser seguidos para cada métrica:

1. Compreensão da finalidade específica, por meio de estudos sobre as métricas pelas indicações do Coh-Matrix e publicações sobre cada índice;

2. Entendimento da faixa de valores possíveis, determinado pelas diretrizes do Coh-Matrix e corroborado pela observação com um *corpus* de complexidade progressiva;

3. Observação do comportamento por meio de um *corpus* controlado de complexidade progressiva – por meio do qual podemos verificar quais valores são resultantes para cada inserção de característica de maior complexidade inserimos para cada texto;

4. Análise do comportamento dos valores em situações extremas utilizando *corpora* comparativos – sendo um *corpus* com textos que são simples e outro com textos complexos;

5. Delimitação da variação da faixa de valores em textos reais, pela observação de faixas de valores que ocorrem em textos que não sejam formados apenas por poucas frases soltas, mas que tenham frases, parágrafos, exemplos de coordenação e subordinação etc. Ou seja, textos que tenham dimensões “regulares” e que sejam similares em tamanho aos que encontramos no dia a dia, desde que proporcionais em extensão aos textos do nosso *corpus* de estudo;

6. Demarcação dos valores em categorias interpretativas de acordo com as finalidades propostas pela análise de complexidade tendo por base princípios organizacionais de Análise de Banco de Dados da Tecnologia da Informação (DUNKEL e SOPARKAR, 1999). Para os propósitos da nossa pesquisa, os seguintes procedimentos foram adotados na organização dos dados:

- 6.1. Disposição dos índices no Excel;

6.2. Classificação por ordem numérica decrescente;

6.3. Comparação dos valores com a ordenação dos números do *corpus* progressivo;

6.4. Verificação de ocorrências irregulares (que “fogem do padrão” como, por exemplo, um texto simples que apresentou valor de determinada métrica próximo aos valores desse índice para os textos complexos);

6.5. Análise holística do nível textual das ocorrências que apresentam valores diferenciados para confirmar se a alteração é coerente (por exemplo, se um texto simples apresenta valor de Simplicidade Sintática mais próximo aos valores dos textos complexos, analisamos as formações sintáticas presentes nesse texto para que possamos confirmar se – para essa métrica – a indicação é condizente);

6.6. Divisão dos valores na quantidade de faixas interpretativas que sejam coerentes ao trabalho proposto (no caso da nossa pesquisa, três níveis de complexidade: alta, média e baixa).

Para exemplificar os procedimentos metodológicos de classificação de valores das métricas para interpretar potencial de complexidade, vamos tomar a métrica Simplicidade Sintática (SS) como exemplo e ilustrar os procedimentos descritos acima.

1. Compreensão da finalidade específica:

No procedimento 1, é preciso compreender que a SS é uma métrica que reflete uma média de complexidade de formação sintática de todas as frases do texto levando em consideração quantidade de palavras e relação entre elas (como quantidade de unidades lexicais antes do verbo, deslocamento de elementos da ordem canônica, etc.).

2. Entendimento da faixa de valores:

Para o passo 2, precisamos observar que os valores possíveis para a SS. O Coh-Metrix e os nossos textos do *corpus* progressivo nos afirmam que os valores de SS variam entre zero e cem (por ser uma métrica percentual).

3. Observação do comportamento por meio de um *corpus* controlado:

No procedimento 3, utilizamos um *corpus* de textos que apresenta complexidade progressiva e observamos o impacto da complexidade relacionada ao nível textual da métrica nos números dispostos por ela.

Na nossa pesquisa, nosso *corpus* progressivo foi redigido por nós apresentando, a cada novo texto, um aumento na dificuldade em diferentes níveis textuais (sintáticos, semânticos, lexicais etc.).

No que diz respeito à sintaxe, por exemplo, nosso primeiro texto era composto por um fragmento de frase, enquanto nosso último texto por diversas frases e parágrafos com estruturas sintáticas em diferentes configurações.

Sendo assim, a progressão de complexidade (como, por exemplo entre as frases com “sujeito e verbo” e outras com “adjunto, sujeito, verbo, objeto direto, objeto indireto”) nos permite observar o comportamento dos valores de métricas como a SS.

Para a SS, observamos que textos que seguem precisamente um paralelismo na formação sintática das frases (como, por exemplo, a forma “sujeito, verbo, objeto” em todas as frases do texto), com a mesma quantidade de palavras para cada função sintática e sem nenhuma ocorrência de frases com diferentes estruturas, têm resultados com valor 100. Entretanto, como linguistas, sabemos que são raros os textos “naturais” que apresentam todas as frases com exatidão em formação sintática.

Já textos com a maior diversidade de formações sintáticas e com diferentes quantidades de palavras para cada função apresentam os valores que tendem a ser entre 5 e 10.

4. Análise do comportamento dos valores em situações extremas:

Para que possamos observar os valores que as métricas tendem a apresentar para textos de diferentes complexidades, utilizamos um *corpus* composto por textos que são considerados simples (como textos para crianças pequenas e textos previamente simplificados seguindo orientações de acessibilidade textual) e outro *corpus* de textos considerados complexos (como materiais técnicos ou acadêmicos). Dessa forma, podemos perceber padrões e faixas de valoração para cada um dos índices e de acordo com a complexidade de cada texto.

Os valores de SS no nosso *corpus* de textos simples apresentaram uma média de 75,5, sendo que a maior parte dos textos tem valores entre 70 e 95. Poucas são as instâncias abaixo de 65. Já entre os textos complexos, o valor médio de SS foi de

46,1, sendo que a maior parte deles variou entre 30 e 65, sendo poucas instâncias acima de 65 e poucas abaixo de 30.

5. Delimitação da variação da faixa de valores:

Tendo conhecimento dos valores possíveis de cada métrica – como no caso de SS que varia entre zero e cem – e observando o comportamento desses valores por meio dos *corpora* comparativos, podemos delimitar a faixa de valores que ocorre em textos reais.

Para a SS, por exemplo, não foi possível observar valores iguais a zero, somente um caso de valor aproximado de 1 e somente outra instância abaixo de 16. Ao realizarmos as análises holísticas desses textos, percebemos que esses dois textos são exemplos raros de formações sintáticas que se diferem entre a maior parte das frases. Portanto, poderíamos excluir das interpretações de classificação de complexidade os valores abaixo de 16 ou incluir os valores de 0 a 16 em um intervalo classificatório extremo.

Na maior parte de textos reais – principalmente os textos mais extensos – sabemos que existem diferentes formações sintáticas entre as frases, mas que, ao mesmo tempo, temos ocorrências de frases com formações sintáticas iguais ou muito similares. Considere os exemplos “*A menina joga futebol.*” e “*João come pizza.*”. Embora tenhamos duas palavras na posição de sujeito na primeira frase e uma na segunda, a formação sintática é sujeito – verbo – objeto para as duas. Logo, é evidente que valores muito próximos a zero ou muito próximos a cem são incomuns.

Para algumas métricas, é possível observar que existem ocorrências de exemplos entre todas as faixas de valor possíveis, sem que seja possível eliminar qualquer intervalo.

6. Demarcação dos valores em categorias interpretativas:

Por meio dos valores resultantes no *corpus* de textos simples e complexos e contrastando com os valores dos textos do *corpus* progressivo, podemos selecionar faixas de valores que são comuns para cada tipo de complexidade indicativa daquela métrica.

Se, por exemplo, consideramos classificar complexidade em três níveis, como no caso dessa pesquisa, podemos definir três faixas de valores que sejam coerentes com os comportamentos das métricas.

A imagem abaixo (Figura 4), cujo propósito é apenas ilustrar a disposição dos valores/cores nesta etapa metodológica, permite-nos observar os passos descritos no procedimento 6 para organizar e classificar as métricas. As células do Excel (nas colunas A, D, G etc.) que estão destacadas em azul representam os textos fáceis do corpus comparativo, enquanto em laranja, os complexos. Se observarmos a disposição das cores, podemos perceber determinados padrões e/ou equilíbrio na organização dos textos (textos mais simples frequentemente posicionados na parte de cima e textos mais complexos, na de baixo). Por meio dessa organização, podemos seguir os processos 6.4 (verificação de ocorrências fogem do padrão) e 6.5 (análise holística do nível textual das ocorrências que apresentam valores diferenciados) para, então, determinar as faixas de valores para fins interpretativos de complexidade. Além da organização de cores para as etiquetas (colunas à esquerda) dos textos simples (azul) e complexos (laranja), os valores numéricos (colunas à direita) foram organizados de acordo com os indícios de complexidade. As cores seguem as seguintes indicações:

- a. Verde: valores indicativos de complexidade baixa. Quanto mais forte a gradação de verde, maior acessibilidade indicada por aquela métrica.
- b. Amarelo: valores indicativos de complexidade média. Quanto mais forte a gradação de amarelo, maior acessibilidade indicada por aquela métrica.
- c. Vermelho: valores indicativos de complexidade alta. Quanto mais forte a gradação de vermelho, menor acessibilidade indicada por aquela métrica.

Figura 4 - Exemplo de classificação de valores de métrica por meio do Excel

Fonte: Elaboração própria

Ao dispormos os valores de SS no Excel e os classificarmos em ordem decrescente, podemos observar que os textos LIS (que são os textos considerados simples do nosso *corpus* comparativo, destacados em azul) de fato estão classificados nos valores mais altos da métrica, enquanto a maior parte dos textos LIC (textos complexos, destacados em laranja), nos valores mais baixos.

Por meio dessa organização, podemos contrastar as faixas de valores com os provenientes do nosso *corpus* progressivo para verificar se existe coerência entre os textos demarcados como complexos para aquele nível textual no *corpus* progressivo com os componentes dos *corpora* comparativo.

Uma vez que percebemos a presença de alguns textos LIS entre os LIC e vice-versa, realizamos uma análise holística desses textos para compreender os motivos para os valores de SS.

É importante lembrar que, algumas vezes, é comum termos textos simples que se classificam com determinados valores métricos próximos aos números dos textos complexos (e vice-versa) porque, conforme já mencionado, complexidade e acessibilidade devem ser observados por meio de mais de uma métrica. No caso do exemplo acima, os textos simples (como o LIS19) que estavam entre os textos complexos pelos valores de SS são simples em léxico e em semântica, mas complexos em estruturação sintática (observado por distintas formações de frases e de quantidade de palavras para cada função sintática) – fato que justifica os valores de SS próximos aos textos complexos.

No caso das métricas de sintaxe, para corroborarmos a análise holística, podemos gerar árvores de análise sintática (como apresentaremos, posteriormente, na explicação da métrica SIMS) de textos com valores extremos para verificar – manualmente – a quantidade de formações sintáticas diferentes nesses textos. Para isso, é possível utilizar *software* de geração de árvores sintáticas, como o VISL⁵², por exemplo.

Após a verificação dos textos que “fugiram do padrão”, podemos classificar os valores de acordo com o número de faixas determinados para a pesquisa (no caso desta tese, três níveis de complexidade: alta, média e baixa).

Para a SS, por exemplo, foi possível constatar que:

- Textos complexos apresentam valores entre 0 e 49;

⁵² Disponível em <https://visl.sdu.dk/visl/en/parsing/automatic/trees.php>

- Textos médios, entre 50 e 74;
- Textos fáceis, entre 75 e 100.

É preciso destacar que as faixas de valores não são necessariamente sempre proporcionais. É preciso estimar estatística e proporcionalmente as quantidades de ocorrências relativas a cada tipo de texto nos *corpora* utilizados para a classificação.

Nas próximas subseções, as métricas escolhidas serão explicadas com mais detalhes e os quadros classificatórios de potencial de complexidade serão apresentados.

6.2.2 NARRATIVIDADE

A métrica Narratividade (NT) é a mais simples entre as escolhidas e o nome é intuitivo. De acordo com os documentos do Coh-Metrix⁵³, uma narratividade é um texto que conta uma história com elementos como personagens, eventos, lugares e outras coisas que são familiares para os leitores. Textos narrativos são próximos de conversas orais e, conseqüentemente, cheios de palavras simples e estruturais similares à fala do cotidiano. Os textos não-narrativos, por outro lado, geralmente são aqueles com tópicos não necessariamente familiares e que, no quesito de linguagem e estrutura, estão no oposto do *continuum* textual das narrativas. As narrativas são, portanto, textos cuja acessibilidade se constrói no desencadear da história por meio de estruturas lógicas e palavras comuns. (MCNAMARA, et al., 2014)

Além disso, é de senso comum que narrativas são – geralmente – os primeiros textos com os quais temos contato. Desde que nascemos, estamos cercados de narrativas: nas conversas que escutamos, nos desenhos animados que assistimos, nas histórias de dormir que nossos cuidadores leem para nós, nas novelas etc. Sendo assim, desde pequenos temos nossas mentes preparadas para esse tipo de texto (seja ele escrito ou oral) que, por sua vez, segue estrutura e desencadeamento lógicos e com manutenções tema-remas acessíveis ao leitor.

A métrica disponibilizada pelo Coh-Metrix, então, analisa o texto submetido e retorna um percentual de similaridade às estruturações de narrativas. Quanto mais próximo esse valor for de 100 (considerando que é um percentual), mais similar é a uma narrativa e – conseqüentemente – mais acessível para um leitor.

⁵³ Disponível em http://cohmetrix.memphis.edu/cohmetrixhome/documentation_indices.html

Para ilustrar o funcionamento da métrica, submetemos os três parágrafos abaixo ao Coh-Metrix para verificar se:

- a. O parágrafo narrativo (parágrafo 1) apresenta resultados mais próximos ao 100.
- b. O terceiro é o que mais se afasta de características de narrativas (mais próximas ao zero).
- c. O segundo parágrafo – por ser uma instrução aproximada de narrativa – apresenta um NT com valor entre os parágrafos 1 e 3.

Parágrafo 1: *Once upon a time there was a beautiful girl who lived on a farm. This girl used to wake up early every morning. After waking up, she had breakfast, brushed her teeth, and fed the animals. One day, she woke up and heavy rain was falling. She put on her raincoat and boots and went to the stables to check on the animals.* (NT = 78.230)

Parágrafo 2: *Follow these steps to enhance your photo using Photoshop. First, open your picture, and under the menu 'Image', click 'Image Size' and observe whether the resolution is at least 250 pixels/inches. Under the menu 'Filter', you can apply the 'Unsharp Mask' to make the edges of the subject more defined. It is also possible to redefine HDR toning and exposure to balance the channels of your image.* (NT = 35.914)

Parágrafo 3: *Peperomia scandens is also commonly known as the Cupid Peperomia and is the variegated form of the green Peperomia scandens. This plant is a vining, epiphytic plant that can reach 24-36 inches in height. It is a very popular indoor plant because of its hardiness and simplicity of care. The plant has thick, glossy, heart-shaped light green leaves and irregular, cream-colored yellow borders.* (NT = 1.970)

Como foi possível prever antes da submissão dos textos ao Coh-Metrix, o parágrafo 1 é, de fato, o que mais se aproxima de uma narrativa (com 78% de equivalência a um texto narrativo), enquanto o 3 é o que menos se aproxima (1,97%). O parágrafo 2, por sua vez, apresenta resultado entre os dois outros (com 35% de

equivalência), mas menos próximos de um valor considerado similar a um texto narrativo.

Embora essa métrica seja simplória em sua proposta, acreditamos que seja válida para a aplicação desta pesquisa principalmente quando adotada com os demais índices escolhidos. Os textos presentes no nosso *corpus* de estudo nem sempre são narrativas – uma vez que a coleção de livros tenta apresentar gêneros textuais distintos para que os estudantes possam se familiarizar com elas em língua inglesa – a métrica nos possibilitará observar se, por exemplo, os materiais do livro *elementary* se afastam muito das estruturas das narrativas e quais as possíveis consequências disso para a ATLD.

Uma vez que a NT é um índice de percentual, podemos classificar seus valores em relação à complexidade da seguinte forma (Quadro 12):

Quadro 12 - Interpretação dos valores de NT para estimar complexidade

COMPLEXIDADE	FAIXA DE VALORES DE NT
Fácil:	80 a 100
Médio:	39 a 79
Difícil:	0 a 40

Fonte: Elaboração própria

É importante observar que temos a consciência, antes da análise dos valores de NT, que de forma isolada, alguns textos podem apresentar valores de NT que seriam considerados potencialmente complexos. Entretanto, uma vez que nem todos os textos do nosso *corpus* seguem uma estruturação de narrativa, essa métrica sozinha não é suficiente para estimar complexidade. Utilizaremos a NT como uma das características que podem indicar complexidade; mas priorizaremos uma análise combinada desse índice com outros para que os resultados sejam coerentes.

6.2.3 SIMPLICIDADE SINTÁTICA

A métrica Simplicidade Sintática (SS) apresenta um percentual da simplicidade ou complexidade da formação das frases do texto. McNamara, et al. (2014) explicam que a interpretação de um texto tende a ser mais fácil quando as frases são escritas com estruturas sintáticas familiares: com poucas palavras, com vocabulário de uso cotidiano e seguindo estruturação canônica. Já frases com mais palavras, com frases

subordinadas e coordenadas e que fujam da ordenação sujeito-verbo-objeto tendem a ser mais complexas.

De acordo com McNamara, et al. (2014, p. 85), a métrica SS “reflete o grau em que as frases em um texto contêm menos palavras e utilizam estruturas sintáticas mais simples e familiares e que sejam menos desafiadoras para processar.”⁵⁴

Silva (2018) aponta como um dos indícios de sua pesquisa e de revisões bibliográficas sobre CT que a ordenação sintática é um dos fatores que mais facilita ou dificulta a compreensão de textos. Um dos conselhos do autor sobre escrita acessível é dar preferência para a estruturação canônica.

A fim de ilustrar os valores de SS, vamos considerar os seguintes exemplos:

Frase 1: *Jake studies English.*

Frase 2: *English is one of the things that Jake, who lives next door, studies.*

Ao submetermos a primeira frase ao Coh-Metrix, podemos observar que seu valor de SS é de 100%, uma vez que utiliza palavras simples e segue a forma canônica direta.

A segunda frase, por sua vez, apresenta um valor de SS de 29.1% - que reflete a estruturação não-canônica e a presença de um aposto.

Com base nas valorações de SS dos nossos *corpora* progressivo e comparativo (cujos índices apresentaremos na seção 9.1), foi possível observar que os textos simples apresentam uma média de 80% para a SS, sendo que a maior parte, valores dessa métrica superiores a 70. Os textos que apresentam baixos valores de SS são as narrativas com muitos apostos e/ou de advérbios e expressões de frequência deslocados para o início da frase.

Os textos complexos do nosso *corpus* comparativo apresentam média de SS de 36.5%, tendo valores que variam de 5% até 74%. Por meio de uma análise holística, podemos observar que o texto que apresenta o índice mais alto de SS no *corpus* de textos complexos (texto LIC4) é composto, de fato, por estruturas mais simples. Os fatores de alta complexidade deste texto estão relacionados ao léxico e à terminologia.

⁵⁴ Tradução nossa. Original: “This component reflects the degree to which the sentences in the text contain fewer words and use simpler, familiar syntactic structures that are less challenging to process.”

Portanto, com base nas observações de comportamento das métricas e no estudo manual da composição sintática dos textos dos nossos *corpora* progressivo e comparativo, podemos classificar os valores de SS a respeito de complexidade do seguinte modo (Quadro 13):

Quadro 13 - Interpretação dos valores de SS para estimar complexidade

COMPLEXIDADE	FAIXA DE VALORES DE SS
Fácil:	75 a 100
Médio:	50 a 74
Difícil:	0 a 49

Fonte: Elaboração própria

De forma resumida, os valores de SS variam entre zero e cem, de forma que valores mais próximos a cem indicam menores potenciais de complexidade, enquanto índices próximos a zero, maiores.

Na próxima subseção, apresentamos a métrica Similaridade Sintática.

6.2.4 SIMILARIDADE SINTÁTICA

A Similaridade Sintática (SIMS), conforme o nome sugere, procura por paralelismos de formação sintática entre as frases do texto. McNamara et al. (2014, p. 70) explicam que:

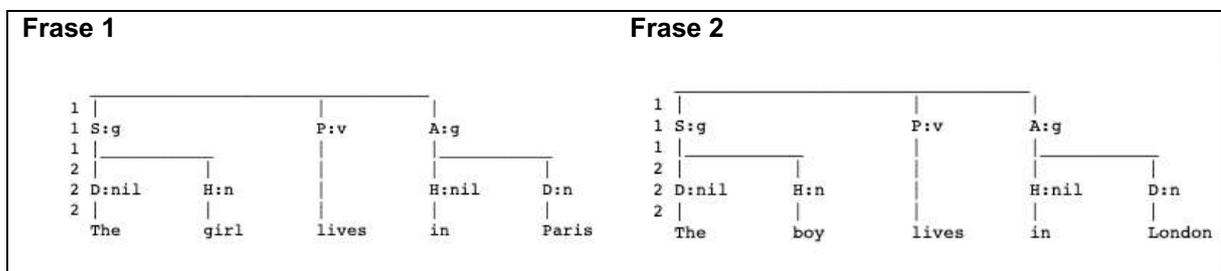
[...] algumas frases são curtas e têm uma formação sintática simples que segue um padrão sujeito-verbo-objeto. Elas têm poucas (se tiver) frases justapostas e que seguem a voz ativa em vez da voz passiva. Algumas frases têm formações sintáticas mais complexas e entrucadas que demandam mais da memória de trabalho. A sintaxe em um texto tende a ser mais fácil de processar quando tem frases mais curtas, poucas palavras antes do verbo principal da oração principal e poucas palavras nas frases substantivas. [...] Frases com construções sintáticas mais difíceis incluem o uso de constituintes justapostos e são, frequentemente, estruturalmente densas, sintaticamente ambíguas ou agramaticais (GRAESSER et al., 2004); como conseqüências, elas são mais difíceis de processar e de entender (PERFETTI, LANDI, e OAKHILL, 2005).⁵⁵

⁵⁵ Tradução nossa. Original: “[...] some sentences are short and have a simple syntax that follow an actor-action-object syntactic pattern, have few if any embedded clauses, and have an active rather than passive voice. Some sentences have complex, embedded syntax that potentially places heavier

Com a finalidade de demonstrar o comportamento da métrica SIMS, realizamos a análise sintática (*parsing*) de três textos do nosso *corpus* progressivo: um com valor alto (CP4, SIMS: 100), outro com valor mediano (CP10, SIMS: 58,3) e, por fim, um texto com valor baixo (CP13, SIMS: 17,1).

Ao analisarmos a árvore de análise sintática do texto CP4 (Figura 5), podemos observar que as duas orações seguem precisamente o mesmo tipo de formação: Sujeito – verbo - adjunto.

Figura 5 - Árvores sintáticas do texto CP4



Fonte: Elaboração própria por meio da plataforma visl.sdu.dk

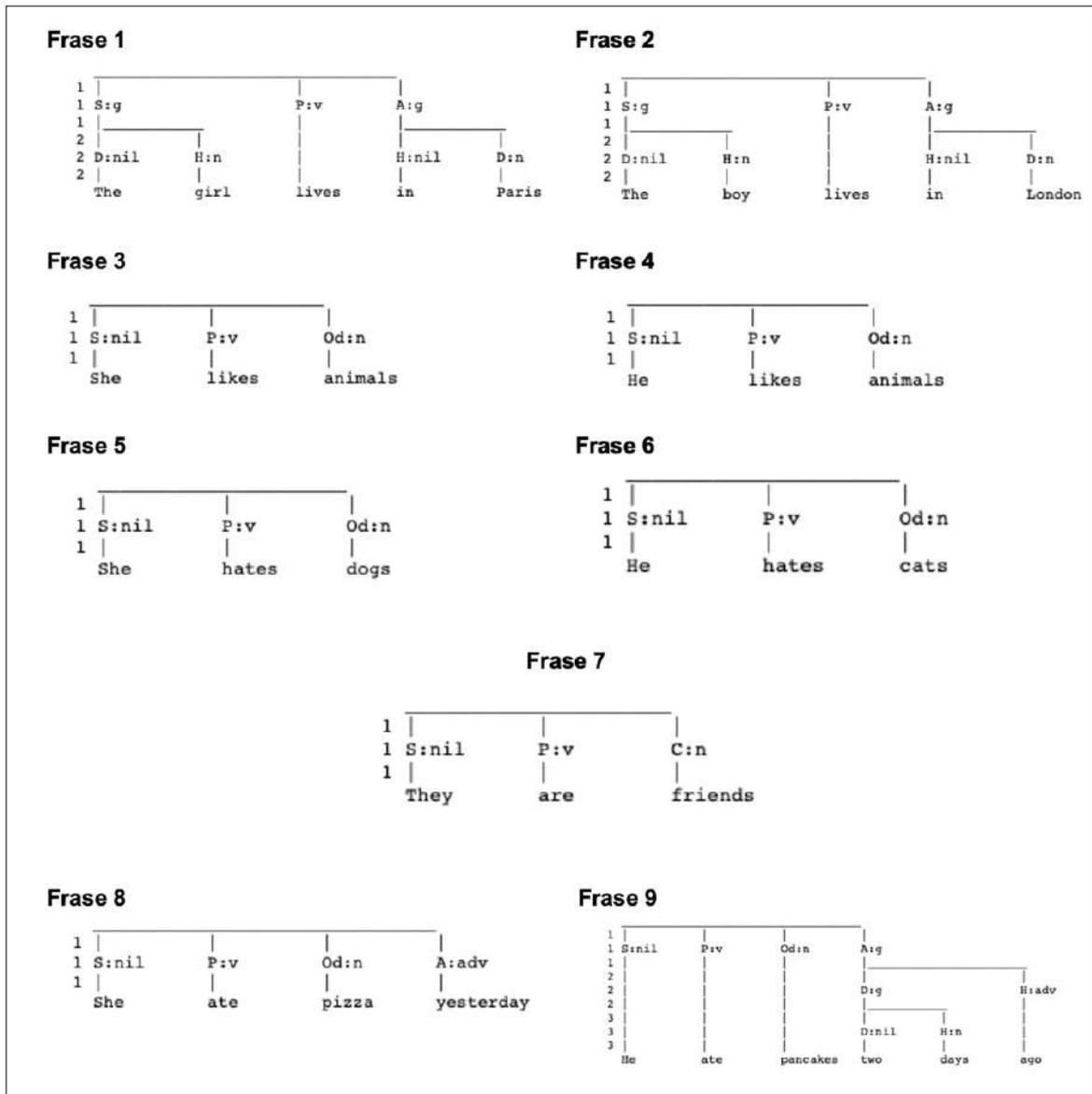
Visto que as duas orações do texto CP4 têm não só a mesma formação sintática, mas também o mesmo número de palavras em cada função, a similaridade é exata e, como consequência, o valor de SIMS é 100.

Já ao analisarmos a árvore sintática do texto CP10 (Figura 6) - cujo valor de SIMS foi mediano (58,3), constatamos quatro tipos de formação:

- Frases 1 e 2: sujeito – verbo - adjunto.
- Frases 3, 4, 5 e 6: sujeito – verbo - objeto direto.
- Frase 7: sujeito – verbo - complemento.
- Frases 8 e 9: sujeito – verbo - objeto direto - adjunto.

demands on working memory. The syntax in text tends to be easier to process when there are shorter sentences, few words before the main verb of the main clause, and few words per noun-phase. [...] Sentences with difficult syntactic constructions include the use of embedded constituents and are often structurally dense, syntactically ambiguous, or ungrammatical (Graesser et al., 2004). As a consequence, they are more difficult to process and comprehend (Perfetti, Landi, e Oakhill, 2005)."

Figura 6 - Árvores sintáticas do texto CP10



Fonte: Elaboração própria por meio da plataforma visl.sdu.dk

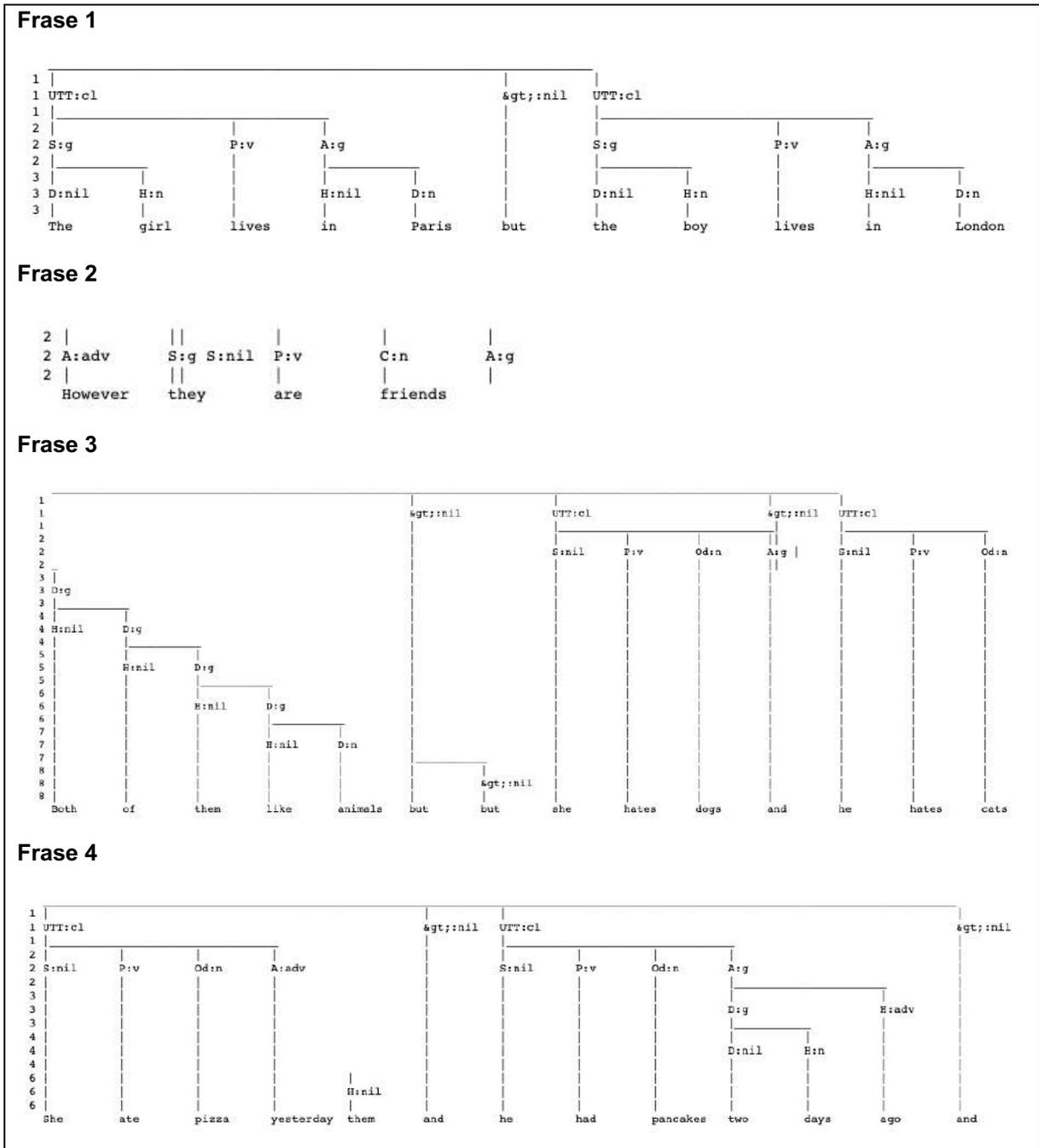
Uma vez que não há paralelismo completo entre as estruturas sintáticas de todas as frases do texto – visto que a frase 7 não tem um par com a mesma formação - o valor da métrica SIMS foi considerado mediano.

Por fim, a Figura 7 abaixo exhibe as árvores de *parsing* do texto CP13, cujo valor de SIMS é baixo como consequência da maior diversificação e disparidade sintática. Podemos verificar as seguintes formações:

- Frase 1: [sujeito – verbo – adjunto] + conjunção + [sujeito – verbo – adjunto]
- Frase 2: adjunto – sujeito – verbo - complemento

- Frase 3: [sujeito – verbo - objeto direto] + conjunção + [sujeito – verbo - objeto direto]
- Frase 4: [sujeito – verbo - objeto direto – adjunto] + conjunção + [sujeito - verbo - objeto direto – adjunto]

Figura 7 - Árvores sintáticas do texto CP13



Fonte: Elaboração própria por meio da plataforma visl.sdu.dk

As árvores de análise sintática nos mostram que embora o texto seja curto, não existem paralelismos frequentes entre todas as frases do texto; todas apresentam

formações sintáticas diferentes (ainda que as coordenadas tenham a mesma estrutura), fato que resultou em um valor baixo de SIMS (17,1).

No que diz respeito à utilização da métrica SIMS como forma de interpretar complexidade, é preciso destacar que os valores desse índice não podem ser distribuídos de forma uniforme nas classificações de complexidade. Obviamente, existe um número limitado de formações e estruturas sintática possíveis dentro da língua inglesa e – conseqüentemente – um texto com mais de dois parágrafos e com um tamanho “médio” (como os textos didáticos do nosso *corpus* de estudo) geralmente não apresenta exatidão em formações sintáticas entre suas frases.

Se tomarmos uma narrativa como exemplo, sabemos que a estruturação sintática dessa tipologia textual tende a ser diferenciada entre cada frase, já que o deslocamento de elementos sintáticos exerce funções específicas relacionadas à atenção, ao destaque etc. Na realidade, como linguistas, podemos corroborar que são poucos gêneros textuais que seguem estruturas sintáticas de formações paralelas entre todas as frases do texto: geralmente receitas, manuais ou textos redigidos para tal finalidade. Logo, por meio dessa constatação e por intermédio dos comportamentos dos nossos *corpora* progressivo e comparativo, compreendemos que raras são as ocorrências onde teremos valores de SIMS próximos a cem. Textos que são simples (mas não similares ao tipo de texto que encontramos diariamente, tanto por uma análise holística, quanto por outros indícios métricos) tendem a apresentar valores de SIMS entre 80 e 100. Textos mais complexos (com frases longas, elementos deslocados, coordenações e subordinações irregulares etc.), por outro lado, resultam valores de SIMS abaixo de 80.

Portanto, por meio das observações do comportamento da métrica SIMS e pelos dados coletados pelos nossos *corpora* progressivo e comparativo (sobre os quais falaremos posteriormente), podemos classificar os valores de SIMS de acordo com potencial de complexidade conforme descrito pelo Quadro 14 abaixo:

Quadro 14 - Interpretação dos valores de SIMS para estimar complexidade

COMPLEXIDADE	FAIXA DE VALORES DE SIMS
Fácil:	Acima de 101
Médio:	80 a 100
Difícil:	Abaixo de 80

Fonte: Elaboração própria

Ainda é válido apontar que McNamara et al. (2014) destacam que textos com frases com maiores semelhanças em formação sintática são, geralmente, mais simples. Entretanto, não podemos utilizar apenas essa métrica como indicativa de complexidade sintática. Ainda que exista uma grande diferença entre a estruturação das frases, é possível que as composições não sejam complexas. Além disso, o vocabulário utilizado nas frases pode ser acessível. Então, nossa análise holística e as observações contrastivas da SIMS com as outras métricas nos auxiliarão a estimar potenciais de complexidade mais eficientemente - estimativa que será apoiada também pela métrica CR, sobre a qual falaremos a seguir.

6.2.5 COESÃO REFERENCIAL

Coesão referencial (CR) é uma característica linguística que diz respeito à relação ou ligação entre termos de um texto permitindo (ou impedindo) que o leitor identifique “o que se refere a quê”, como, por exemplo, na frase “*The boy cried. He was sad.*”, o pronome ‘*he*’ é referência à combinação de artigo e substantivo ‘*the boy*’.

A coesão referencial é uma das responsáveis pela sequenciação e desenvolvimento lógico e conectivo de um texto. Quando há falha na coesão, o texto (ou um fragmento dele) pode ser considerado complexo. (SILVA, 2018)

No que diz respeito à CR, Silva (2018) demonstrou que a relação entre pronomes e substantivos pode ser um fator de alta complexidade, principalmente quando o pronome está distante do sujeito ao qual se refere. Essa distância pode não permitir uma referenciação lógica, principalmente para leitores menos proficientes. Um exemplo é “*The boy cried, and his father comforted him. He was sad.*”. A primeira interpretação de um leitor proficiente é relacionar o pronome ‘*he*’ ao ‘*the boy*’, uma vez que “se o garoto chorou, provavelmente era ele quem estava triste.”. Contudo, é importante lembrar que nunca podemos fazer suposições sobre outros leitores com base nas nossas experiências. No caso do exemplo, um leitor em níveis iniciais de aprendizado de inglês como língua adicional pode conectar o pronome a ‘*his father*’, uma vez que ‘*father*’ também é uma palavra masculina e está mais próxima do pronome.

Pelos motivos supracitados, uma das métricas escolhidas para essa pesquisa é a CR, mais especificamente a coesão referencial percentual que, de acordo com as diretrizes normativas do Coh-Matrix, verifica justaposição de substantivos e pronomes

entre as frases de um texto e exibe um valor percentual de coesão. Essa justaposição ocorre quando há uma ocorrência de um substantivo/pronome em uma frase e o mesmo substantivo (seja no singular ou plural) e/ou pronome em outra frase. (MCNAMARA, et al., 2014). Quando há baixos valores de coesão em um texto, seu processamento se torna mais complexo porque isso reflete que existem poucas conexões lógicas que podem ajudar a conectar as ideias dentro de um texto.

Seguindo nossa classificação de complexidade previamente explicada, desenvolvemos a seguinte classificação (Quadro 15) dos valores de CR em relação aos seus indícios de potencial de complexidade:

Quadro 15 - Interpretação dos valores de CR para estimar complexidade

COMPLEXIDADE	FAIXA DE VALORES DE CR
Fácil:	50 a 100
Médio:	31 a 49
Difícil:	0 a 30

Fonte: Elaboração própria

A seguir, apresentaremos a única métrica cujo desenvolvimento foi especialmente voltado para estimativas de complexidade textual: o índice Flesch.

6.2.6 ÍNDICE FLESCH

O Índice Flesch – originalmente dividido em duas fórmulas – foi desenvolvido pelo austríaco Rudolf Flesch (1981) como métrica estimativa de complexidade textual baseada na extensão de palavras e de frases. Esse índice, que atualmente é uma unificação de duas fórmulas originais, apresenta um número que classifica a dificuldade de leitura de um texto e a quantidade de anos escolares formais necessárias para a compreensão de um determinado material (GOLDIM, 2003).

Algumas ferramentas unificam as duas fórmulas originais do IF e apresentam os resultados conforme o quadro abaixo (Quadro 16):

Quadro 16 - Resultados do Índice Flesch.

ÍNDICE	CLASSIFICAÇÃO
100-90	Muito fácil de ler. Facilmente compreendido por um aluno médio de 11 anos de idade.
90-80	Fácil de ler. Inglês conversacional.

80-70	Razoavelmente fácil de ler.
70-60	<i>Plain English</i> . Facilmente compreendido por alunos de 13 a 15 anos de idade.
60-50	Razoavelmente difícil de ler.
50-30	Difícil de ler
30-10	Muito difícil de ler. Melhor compreendido por estudantes de graduação.
10-0	Extremamente difícil de ler. Melhor compreendido por graduados e pós-graduados.

Fonte: Goldim (2003)

Para fins ilustrativos, consideremos os seguintes exemplos:

Exemplo 1: *A great number of technological tools – such as Quizlet, Kahoot, and Quizizz – have been widely used for gamification purposes in classrooms all around the world and, simultaneously, support teaching practices that, following the premises of modern educational methodologies, have been transformed into a more student-centered practice; allowing students to be the center and “tailors” of learning processes. (IF = 29.1)*

Exemplo 2: *There are tools that teachers have been using for gamification reasons. Some examples are Quizlet, Kahoot and Quizizz. These tools help teaching practices. These practices put the students in the center. By doing so, students are allowed to be the “tailors” of their learning. (IF = 71)*

Por meio de uma simples leitura dos dois exemplos, podemos perceber que o exemplo 1 foi escrito em um único e longo parágrafo e com palavras mais extensas do que o texto no exemplo 2. O segundo texto, por sua vez, dividiu as informações em frases mais curtas e preferiu utilizar algumas palavras e termos de extensões menores. Uma vez que o IF é um índice cujo resultado é calculado por meio da extensão de palavras, frases e parágrafos é evidente que, então, o resultado de IF do exemplo 1 indicaria, de fato, um texto mais complexo do que o segundo. Ao submetermos os exemplos ao Coh-Metrix, coletamos a informação de que o primeiro exemplo tem 29.1 de IF, sendo classificado como muito difícil de ler e melhor compreendido por estudantes de graduação, enquanto o exemplo 2 apresenta IF de

71 (que o classifica como *Plain English* - Facilmente compreendido por alunos de 13 a 15 anos de idade.).

Uma vez que nosso leitor não é falante de inglês como primeira língua, adaptamos as indicações de IF para que se enquadrem nas nossas etiquetas interpretativas. Com base na análise dos nossos *corpora* progressivo e comparativo (sobre os quais falaremos mais adiante), organizamos o quadro abaixo (Quadro 17) para que possamos interpretar os potenciais de complexidade com as mesmas nomenclaturas utilizadas para as demais métricas:

Quadro 17 - Adaptação dos valores de IF para estimar complexidade

COMPLEXIDADE	FAIXA DE VALORES DE IF
Fácil:	70 a 100
Médio:	61 a 69
Difícil:	0 a 60

Fonte: Elaboração própria, adaptado de Flesch (1981)

Para nossa pesquisa, utilizaremos o IF apenas como um “ponto de partida” indicativo de complexidade. Entretanto, como critica Silva (2021), utilizar apenas o IF para estimar CT é uma prática limitadora, uma vez que a fórmula se reduz à extensão e complexidade não deve ser reduzida a apenas um atributo.

A seguir, explicamos a primeira métrica relacionada ao léxico e à semântica.

6.2.7 POLISSEMIA

Polissemia, como conhecemos linguisticamente, é o fenômeno de múltiplos sentidos possíveis de uma palavra. Exemplos são as palavras ‘tênis’ e ‘casa’. A primeira pode significar tanto um jogo que utiliza raquetes quanto um tipo de sapato; a segunda, por sua vez, pode invocar tanto o significado de ‘moradia’, quanto o verbo ‘casar’ conjugado na terceira pessoa do singular.

Ao termos consciência de que existem inúmeras palavras que possuem mais de um possível significado, podemos concluir que o fator polissemia pode ser um elemento que aumenta o potencial de complexidade dos textos (seja em primeira ou em língua adicional) não só porque exige maior vocabulário do leitor, mas também porque diante de significados não conhecidos, é preciso utilizar ou conhecimentos morfológicos para inferir significado, ou recursos contextuais ou, na impossibilidade

de obtenção de significação por meio desses passos, pesquisar em dicionários ou tradutores, por exemplo.

Sendo assim, uma vez que o Coh-Metrix disponibiliza uma métrica que calcula uma média de valores de polissemia nos textos submetidos, acreditamos que esse índice seja de grande valor aliado aos demais previamente mencionados para estimativa de complexidade textual, principalmente no que diz respeito ao léxico.

Para explicar a métrica, McNamara, et al. (2014, p.75) apresentam o seguinte parágrafo:

Polissemia se refere ao número de sentidos (principais sentidos) de uma palavra. Por exemplo, a palavra 'bank' tem pelo menos dois sentidos [em língua inglesa], um sendo um prédio ou instituição para depositar dinheiro e outro sendo a costa de um rio. O Coh-Metrix apresenta uma média de polissemia das palavras de conteúdo em um texto. As relações polissêmicas na métrica são baseadas em *synsets* (grupo de itens lexicais relacionados), que são usados para representar conceitos similares, mas faz distinção entre sinônimos e sentidos de palavras (Miller et al., 1990). Esses *synsets* permitem uma diferenciação de sentidos e proveem a base para examinar o número de sentidos associados a uma palavra. O Coh-Metrix reporta a média dos valores de polissemia para todas as palavras de conteúdo do texto. Polissemia é considerada um indicativo de ambiguidade uma vez que quanto mais sentidos uma palavra tiver, maiores interpretações lexicais ela tem. Contudo, palavras mais frequentes também tendem a ter mais significados, então, valores altos de polissemia em um texto também podem indicar alta presença de palavras comuns.⁵⁶

Com base nas observações apresentadas pelo documento do Coh-Metrix, compreendemos que a utilização da métrica de polissemia irá permitir interpretar os possíveis múltiplos significados que palavras nos textos podem evocar, e – ao cruzarmos com outras métricas relacionadas ao léxico, verificar se altos índices de polissemia são consequências de presenças de palavras comuns.

Diante da verificação dos níveis de polissemia, podemos observar os textos de forma individual e averiguar se as palavras polissêmicas são – supostamente – parte do vocabulário prévio que os alunos necessitam (e que aprenderam em unidades ou

⁵⁶ Tradução adaptada nossa. Original: *Polysemy refers to the number of senses (core meanings) of a word. For example, the word bank has at least two senses, one referring to a building or institution for depositing money and the other referring to the side of a river. Coh-Metrix provides average polysemy for content words in a text. Polysemy relations in WordNet are based on synsets (i.e., groups of related lexical items), which are used to represent similar concepts but distinguish between synonyms and word senses (Miller et al., 1990). These synsets allow for the differentiation of senses and provide a basis for examining the number of senses associated with a word. Coh-Metrix reports the mean WordNet polysemy values for all content words in a text. Word polysemy is considered to be indicative of text ambiguity because the more senses a word contains relates to the potential for a greater number of lexical interpretations. However, more frequent words also tend to have more meanings, and so higher values of polysemy in a text may be reflective of the presence of higher frequency words.*

níveis anteriores) ou se são novas. Sendo palavras novas, pretendemos analisar a possibilidade de utilização de sinônimos mais simples ou estratégias para que a evocação de significado de palavras polissêmicas não seja ambígua.

Ao utilizarmos nosso *corpus* comparativo para observar o comportamento da métrica PS, observamos que textos mais simples apresentam valores de PS mais altos, enquanto textos mais complexos, mais baixos. A média de valores dessa métrica foi de 438.387 para os textos simples, e de 396.614 entre os textos mais complexos. O valor de PS mais alto (e, conseqüentemente, com indicação de maior acessibilidade) foi do texto LIS 3, cujo valor é de 472.75 – um texto narrativo escrito para *toddlers* (crianças entre 1 e 3 anos de idade). Já o valor mais baixo de PS (que indica maior complexidade) foi de 356.714, do texto LIC1: texto acadêmico de linguística altamente terminológico, complexo em terminologia e com variação lexical grande.

Os textos do *corpus* progressivo (sobre o qual falaremos adiante) comprovaram que textos mais simples apresentam valores mais altos da métrica polissemia (mas, menor presença de palavras polissêmicas). Uma vez que esses textos foram redigidos com complexidade progressiva, um gráfico-linha dos valores de PS nos permitiu observar que o declínio dos dados numéricos desse índice simboliza maior complexidade.

O quadro abaixo (Quadro 18) organiza os valores de PS como interpretativas de complexidade:

Quadro 18 - Interpretação dos valores de PS para estimar complexidade

COMPLEXIDADE	FAIXA DE VALORES DE PS
Fácil:	Acima de 4,05
Médio:	3,49 a 4,04
Difícil:	Abaixo de 3,50

Fonte: Elaboração própria

Além da PS, também utilizamos a métrica Idade de Aquisição, que apresentamos na próxima seção, para auxiliar a interpretar os potenciais de complexidade lexical e semântica dos nossos textos.

6.2.8 IDADE DE AQUISIÇÃO

A primeira observação importante sobre a métrica Idade de Aquisição (IA) é que ela não tem por objetivo estimar a idade média com a qual leitores que tenham o

inglês como primeira língua devam “adquirir” as palavras de um texto. Na verdade, essa métrica é baseada nas normas de aquisição de léxico propostas por Gilhooly e Logie (1980) e apresenta como resultado um valor de IA com média das palavras do conteúdo de um texto. Esse índice tem sido utilizado com grande frequência desde os anos 80 como estimativa de aquisição de vocabulário, como observações de interpretações de texto etc.

O IA reflete a noção de que algumas palavras (como ‘mãe’, ‘comida’ etc.) são apresentadas e se tornam parte do léxico de uma criança antes do que outras palavras (por exemplo, ‘matemática’, ‘imposto’ etc.); ou seja, existe uma ordenação natural na apresentação de vocabulários desde a primeira infância – que se alinha à ideia de ordenação proposta por Krashen (1987), anteriormente mencionado. Sendo assim, o IA atribui uma pontuação calculada pela média da pontuação de cada palavra de conteúdo do texto - *score* que é proveniente dos valores estabelecidos no estudo de Gilhooly e Logie (1980) com 1903 palavras únicas - de forma que valores mais altos indicam a presença de palavras mais complexas (ou seja, que aparecem no vocabulário de pessoas com idades mais avançadas); e um número mais baixo indica que o texto apresenta léxico com vocabulário mais próximo ao que uma criança tem.

McNamara et al. (2014) exemplificam esse índice ilustrando que palavras como ‘*cortex*’, ‘*dogma*’ e ‘*matrix*’ têm IA de valor 700 – uma vez que essas palavras não são comuns no vocabulário de uma criança e começam a aparecer com mais naturalidade no vocabulário de adolescentes e adultos. Palavras como ‘*milk*’, ‘*smile*’ e ‘*pony*’, por outro lado, têm IA de valor 202 uma vez que são palavras mais comuns para crianças e presentes nos seus *inputs* linguísticos desde o nascimento. Além disso, o documento explica que textos com uma média de IA com valores altos demonstram a presença de léxico que só são adquiridos com idades mais avançadas.⁵⁷

Tendo em mente que embora nosso leitor seja um adulto, os textos são escritos em uma língua adicional; consequentemente, acreditamos que as médias de IA para os textos do livro *elementary* devem ser relativamente próximas aos valores de referência ao vocabulário infantil, enquanto a progressão dos níveis permite valores de IA mais elevados.

Yamazaki et al. (1997) explicam que essa métrica desenvolvida por Gilhooly e Logie (1980) classificam as palavras em uma escala de 1 a 7, sendo o número 1

⁵⁷ No estudo original, a escala de pontuações dessa métrica varia de 1 a 7. Entretanto, o Coh-Metrix multiplica esses números por 100 para fins de arredondamento.

representativo de palavras aprendidas antes dos dois anos de idade e 7 palavras apresentadas mais distante da infância (após os 13 anos ou em idade adulta).

Zevin e Seidenberg (2002) explicam que a idade na qual uma palavra é aprendida por intermédio da língua falada influencia na proficiência de leitura e que embora o aprendizado lexical não seja estritamente relacionado à idade, existem estudos (Markson e Bloom, 1997; McCandliss, Posner, e Givon, 1997) e evidências que apontam que palavras aprendidas mais cedo na infância facilitam e impactam no desenvolvimento da compreensão leitora. Logo, a métrica IA vem sendo pesquisada em inúmeros contextos linguísticos relacionados à capacidade de leitura e interpretação.

Ainda que essa métrica não seja estritamente utilizada para estimar complexidade, com o auxílio das diretrizes de Gilhooly e Logie (1980) e por meio da observação dos valores de IA para os textos dos nossos corpora comparativo e progressivo, organizamos o quadro abaixo com faixas classificativas dos valores de IA em relação a interpretação de potenciais de complexidade, (Quadro 19):

Quadro 19 - Interpretação dos valores de IA para estimar complexidade⁵⁸

COMPLEXIDADE	FAIXA DE VALORES DE IA
Fácil:	Abaixo de 275
Médio:	274 a 349
Difícil:	Acima de 350

Fonte: Elaboração própria

É interessante observar que o intervalo numérico dessa métrica é muito pequeno, com variações geralmente nos decimais. Entretanto, alguns exemplos de textos longos e com muitas palavras distintas e especializadas demonstram a possibilidade de valores mais próximos a 2, enquanto textos mais curtos, com pouquíssimas palavras e léxico simples chegam a apresentar valores próximos a 3.

A terceira métrica utilizada para estimar potenciais de complexidade em nível lexical e semântico é a Familiaridade com Palavras de Conteúdo, sobre a qual falaremos na seção que segue.

⁵⁸ Tendo em mente que o Coh-Metrix multiplica os valores propostos por Gilhooly e Logie (1980) por cem, também organizamos nosso quadro classificativo nessa escala numérica.

6.2.9 FAMILIARIDADE COM PALAVRAS DE CONTEÚDO

A métrica Familiaridade com Palavras de Conteúdo (FPC), conforme o nome indica, apresenta uma estimativa com base em um banco de palavras contendo cerca de 3488 palavras classificadas por nível de familiaridade. Dessa forma, o Coh-Metrix estuda os textos submetidos e retorna um valor que calcula a média da familiaridade das palavras de conteúdo de um texto.

McNamara et al. (2014, p.74) explicam a FPC da seguinte forma:

Essa é uma classificação de quão familiar uma palavra parece para um adulto. Frases com mais palavras familiares são processadas mais rapidamente. Essa métrica apresenta classificações para 3488 palavras. O Coh-Metrix provê uma média das pontuações para as palavras de conteúdo de um texto. As classificações por familiaridade são apresentadas por meio de escalas de 7 pontos, sendo 1 ponto destinado para palavras que adultos provavelmente nunca tenham visto e 7 para palavras que adultos veem com grande frequência (quase todo dia). As pontuações são multiplicadas por 100 e arredondadas para números não-quebrados.

Por exemplo: as palavras 'milk' (FPC: 588), 'smile' (594) e 'pony' (524) têm uma média de familiaridade de 569. As palavras 'cornet' (364), 'dogma' (328) e 'manus' têm uma média de familiaridade de 268. Exemplos de palavras com alto grau de familiaridade são 'mother' (632), 'water' (641) e exemplos de palavras com baixo grau de familiaridade são 'calix' (124) e 'witan' (110).⁵⁹

Utilizaremos o FPC, então, para auxiliar a estimar a complexidade lexical dos textos do *corpus* de estudo. Acreditamos que junto às demais métricas (como a Idade de Aquisição, Polissemia e Relação *Type-Token*), poderemos ter resultados fidedignos sobre a complexidade do vocabulário dos materiais analisados.

A coleta de dados dos nossos *corpora* progressivo e comparativo (que apresentaremos na seção 9.1) e as diretrizes dadas por McNamara et al. (2014), nos permitiram observar que a FPC se comporta da seguinte forma: textos mais simples apresentam valores mais altos, enquanto textos mais complexos têm valores mais baixos.

⁵⁹ Tradução adaptada nossa. Original: *This is a rating of how familiar a word seems to an adult. Sentences with more familiar words are words that are processed more quickly. MRC provides ratings for 3488 unique words. Coh-Metrix provides the average ratings for content words in a text. Raters for familiarity provided ratings using a 7-point scale, with 1 being assigned to words that they never had seen and 7 to words that they had seen very often (nearly every day). The ratings were multiplied by 100 and rounded to integers.*

For example, the words milk (588), smile (594), and pony (524) have an average Familiarity of 569 compared to the words cornet (364), dogma (328), and manus (113), which have an average Familiarity of 268. Words with very high Familiarity include mother (632) and water (641), compared to calix (124) and witan (110).

No *corpus* comparativo, os textos LIS8 (FPC: 592.839) e LIS6 (FPC: 590.416) são os que apresentaram valores mais altos de FPC (logo, maior acessibilidade), enquanto os textos LIC7 (FPC: 545.283) e LIC3 (FPC: 546.943), valores mais baixos (maior complexidade). Uma vez que os textos LIS8 e LIS6 são narrativas infantis com vocabulário adequado para crianças de 1 a 3 anos de idade, e os textos LIC7 e LIC3 são textos acadêmicos de Microbiologia, esperávamos que os primeiros apresentassem, de fato, valores mais altos de FPC; enquanto os segundos, índices mais baixos.

Da mesma forma, os textos do nosso *corpus* progressivo demonstram uma curva de declínio (sobre a qual falaremos posteriormente), corroborando a relação entre valores de FPC e complexidade lexical.

Com base nos valores de FPC observados nos nossos *corpora* comparativo, progressivo e de estudos, compreendemos que a variação dessa métrica em textos reais está na faixa de 540 a 600. Valores mais próximos a 600 são os textos que apresentam vocabulário simples do cotidiano, enquanto números mais próximos a 540 apresentam vocabulário mais especializado.

Seria de estranhar que os valores dos textos simples do *corpus* comparativo apresentassem valores tão “próximos” de FPC quanto os dos textos complexos (ambos grupos na faixa entre 540 e 600). Entretanto, nossos estudos nos permitiram perceber que a proporção de palavras do cotidiano para termos (palavra: termo) é geralmente de 95%: 5% – ou seja, para que os textos tenham coerência, termos especializados compõem apenas 5% do vocabulário. Caso estivéssemos analisando uma lista de termos não contextualizados em um texto, os valores de FPC seriam, então, mais baixos e fora da escala 540-600.

Com base nas observações acima, desenvolvemos o quadro abaixo que organiza os valores de complexidade indicados pela FPC (Quadro 20):

Quadro 20 - Interpretação dos valores de FPC para estimar complexidade

COMPLEXIDADE	FAIXA DE VALORES
Fácil:	Acima de 580
Médio:	564 a 579
Difícil:	Abaixo de 565

Fonte: Elaboração própria

Além da FPC e das demais métricas já apresentadas, outro índice que utilizaremos é a Concretude Lexical, sobre o qual falaremos na próxima seção.

6.2.10 CONCRETUDE LEXICAL

A métrica concretude lexical, como o nome sugere, exibe um valor que expressa o percentual de palavras do texto estudado cujo significante evoca um significado concreto no mundo real. Ou seja, palavras como “carro” e “gato”, por exemplo, trazem à mente ideias concretas. Embora cada pessoa imagine veículos e felinos distintos ao ler as palavras, elas ainda apontam para algo concreto no mundo real e a relação entre significante e significado é mais concreta, como o exemplo (Figura 8) apresentado no Curso de Linguística Geral (SAUSSURE, 2012 [1975], p.107) que ilustra essa relação por meio da palavra árvore que, também, elicitava uma ideia concreta na mente do leitor.

Figura 8 - Exemplo de Saussure (1975) para explicar a relação entre significante e significado



Fonte: Saussure (2012 [1975], p.107)

De acordo com McNamara et al. (2014, p. 75), “palavras que são mais concretas são aquelas coisas que você pode ouvir, provar ou tocar.”. Com a finalidade de ilustrar concretude lexical, os autores afirmam que exemplos de palavras com valores altos dessa métrica são “box” e “ball”, enquanto “protocol” e “difference” recebem valores mais baixos.

Outras palavras não possibilitam a evocação de um significado concreto. As palavras “amor” e “difícil”, por exemplo, não têm um único (ou concreto) referente no mundo real. Sendo assim, a significação de algumas palavras é abstrata e o esforço cognitivo necessário para relacionar o significado ao significante não é tão direto quanto ao das palavras concretas.

Uma vez que, como explica Saussure (2012 [1975]), línguas estão restritamente relacionadas às culturas nas quais estão inseridas, alguns conceitos não-concretos podem ser dependentes da língua e relação social dos seus falantes, como o clássico exemplo das diversas unidades lexicais que comunidades esquimós

têm para significar “gelo” e, também, para “branco”. Logo, é evidente que palavras abstratas podem ser potencialmente complexas para falantes não-nativos (especialmente aprendizes) de uma língua.

Embora as definições de “concreto” e “abstrato” pareçam simples, elas têm alguns desdobramentos que podem influenciar no potencial de complexidade de um texto. O professor e linguista americano Hayakawa (1978) apresenta um conceito denominado “escada de abstração” que descreve como as pessoas pensam e se comunicam seguindo determinados graus de concretude e abstração. Para explicar esse conceito, ele apresenta o exemplo de uma vaca chamada Bessie e a forma como ela é “categorizada” em distintos contextos, conforme ilustra a imagem abaixo (Figura 9):

Figura 9 - Exemplo da Escada de Abstração de Hayakawa



Fonte: (HARVEY, 2016)

Hayakawa (1978) explica que na base da escada, está a definição mais concreta de Bessie e no topo da escada, a forma mais abstrata. O linguista explica os níveis da seguinte forma:

I. No primeiro degrau (*process level*) vê-se Bessie como a soma de todos seus átomos para formar uma vaca;

II. No segundo degrau (*name*), nossa experiência começa a entrar na equação, uma vez que entendemos que vaca é um animal mamífero quadrúpede com chifres e

manchas. Ao passarmos de carro pela fazenda em que Bessie está e a virmos de longe, nossa experiência (seja de forma direta ou indireta) nos faz reconhecer que ela é uma vaca;

III. No terceiro degrau (*Bessie*), a atribuição do nome “Bessie” à vaca específica nos permite separá-la das demais. Bessie não é apenas uma vaca, ela é aquela vaca particular. Os filhos do dono da Bessie, por exemplo, ao ouvir seu nome, fazem uma relação direta com a “vaquinha que gosta de dormir e que tem uma mancha em forma de coração”, por exemplo, entre todos os outros animais da mesma espécie;

IV. No quarto degrau (*cow*), a abstração se torna maior e Bessie perde sua identidade única e passa a ser vista como mais uma vaca entre todas as outras;

V. No quinto degrau (*livestock*), a abstração é ainda maior e Bessie começa a ser vista como parte de um grupo que é definido por traços em comum. Nesse grupo estão incluídos os porcos, as ovelhas, as cabras e todos os outros animais pertencentes à pecuária;

VI. No sexto degrau (*farm assets*), Bessie passa a ser categorizada não como parte dos animais, mas como parte dos ativos da fazenda (que inclui maquinários, produtos e tudo que tenha valor monetário);

VII. No sétimo degrau (*asset*), Bessie “perde” quase todos seus traços particulares e é categorizada como parte de qualquer coisa na fazenda que tenha algum valor, não só monetário;

VIII. No oitavo e último degrau (*wealth*), Bessie alcança o nível mais alto de abstração e passa a ser vista como “riqueza”. A palavra “riqueza” é muito abstrata e subjetiva e Bessie não tem mais nenhum traço que a caracterize em um nível individual.

O exemplo apresentado por Hayakawa (1978) sobre a vaca Bessie nos ajuda a entender que a maior parte das palavras não pode ser vista de uma forma binária (concreta ou abstrata) e que a abstração pode ser entendida por uma forma escalar de acordo com o contexto e com a perspectiva adotada.

Tendo em mente a ideia da escada de abstração proposta por Hayakawa (1939), entendemos que essa característica lexical quando relacionada à potencial de complexidade deve ser cuidadosamente analisada, especialmente quando estivermos trabalhando com textos voltados para o ensino de uma língua.

A complexidade expressa pela concretude deve ser estudada de duas formas: 1. pela complexidade da palavra em si (extensão, morfologia etc.) e; 2. Pela facilidade ou dificuldade de associação entre significante e significado (a palavra “amor”, mesmo que abstrata é fácil de ser compreendida. Já a palavra “altruísmo”, por exemplo, pode ser um pouco mais difícil de ser interpretada).

O Coh-Metrix nos auxilia a verificar a proporção de palavras que evocam conceitos concretos pelas unidades que têm um significado abstrato. No quadro abaixo (Quadro 21), classificamos os valores da métrica Concretude Lexical (CLP⁶⁰) em relação à complexidade:

Quadro 21 - Interpretação dos valores de CLP para estimar complexidade

COMPLEXIDADE	FAIXA DE VALORES
Fácil:	80 a 100
Médio:	54 a 79
Difícil:	0 a 55

Fonte: Elaboração própria

Como é possível perceber, nossos estudos e cálculos a respeito da CLP demonstram que a divisão da faixa de valores para cada nível de complexidade não é uniforme. Os textos do *corpus* comparativo que têm grande frequência de palavras abstratas obtiveram valores abaixo de 55.

Diante da ideia da escada de abstração supracitada e das nossas observações em relação ao uso de concretude lexical para estimar complexidade, é preciso destacar que essa métrica tem seu uso otimizado quando adotada com outros índices indicativos de léxico. Visto que a abstração pode ser compreendida por níveis e que existem palavras que mesmo abstratas permitem uma fácil relação de significado, os indicativos de CLP devem ser utilizados como forma de sinalizar textos com muitas unidades abstratas para que professores de ILA observem esses materiais e analisem o potencial de complexidade do vocabulário abstrato.

Na próxima seção, explicamos a última métrica relacionada a léxico que adotamos na nossa pesquisa, a Relação *Type-Token*.

⁶⁰ A letra “P” é a sigla de ‘percentual’, visto que essa métrica resulta em um número percentual.

6.2.11 RELAÇÃO TYPE-TOKEN

O índice *Type-Token* (TT), ou *Type-Token Ratio* (TTR), é uma métrica de relação entre os tipos de palavra em um texto (ou seja, instâncias únicas de cada palavra diferente) e os tokens (quantidade total de palavras em um determinado texto). Por esse motivo, o TT pode ser utilizado como indicativo de variedade lexical.

Ao calcular o TT, o Coh-Metrix utiliza apenas palavras de conteúdo (substantivos, verbos, adjetivos, advérbios e numerais) e não leva em consideração as palavras de função (artigos, preposições, conjunções, pronomes e interjeições).

O TT, de acordo com McNamara, et al. (2014), ao ser utilizado como estimativa de diversidade lexical, indica que quanto mais diverso for o vocabulário apresentado no texto, maior a probabilidade de problemas de coesão e, quanto menor a heterogeneidade, maior coesão. Sendo assim, índices altos de TT sugerem que o texto apresenta uma gama muito variada de palavras, menor coesão e – em consequência, maior complexidade. Para ilustrar o funcionamento da métrica, vamos observar os dois exemplos abaixo:

Frase 1: O gato é bonito. O gato é grande. O gato é interessante. O gato é legal. (TT: 0.438)

Frase 2: O gato é bonito. O felino é grande. O bichano é interessante. O gatinho é legal. (TT: 0.625)

Para fins ilustrativos, as frases acima são curtas. Podemos perceber, sem a necessidade de auxílio de *software*, que a variação lexical da frase 2 é muito maior do que da frase 1; portanto, o índice de TT (comprovado por meio do *software*) é maior.

É evidente que, pela pequena extensão dos exemplos, não percebemos falhas de coesão textual em consequência de variedade lexical maior da segunda frase. Textos mais longos, por outro lado, como na pesquisa de Silva (2018), demonstraram que índices mais altos de TT estavam quase sempre relacionados a maiores problemas de coesão.

A relação entre extensão de vocabulário e instrução formal não é novidade. Desde a época dos escribas, somente as altas classes tinham acesso à leitura e, conseqüentemente, expansão lexical. Nos tempos atuais, pessoas com mais anos de educação formal recebem maiores *inputs* textuais e, por conseguinte, tendem a apresentar um leque de palavras mais extenso do que pessoas com menos tempo de

educação formal. Ao levarmos em conta que os textos do nosso *corpus* de estudo são voltados para aprendizes de inglês como língua adicional, podemos afirmar que o léxico desses estudantes é – na grande maioria das vezes – mais reduzido se comparados com seus domínios de palavras na língua portuguesa. Dessa forma, podemos também considerar que textos com alta presença de vocabulário variado pode ser considerados potencialmente complexos.

Outro fator de suma importância que pode relacionar TT à complexidade textual linguodidática é que variação lexical pode influenciar na coesão referencial. Quando palavras são substituídas por sinônimos (ou por pronomes), o leitor precisa “enxergar” a ligação entre elas para compreender que uma faz referência à outra. Quanto maior for o grau de coesão referencial de um texto, maior o grau de legibilidade e maior a acessibilidade que ele oferece, como apontam Silva et al. (2018).

Utilizaremos o índice TT nesta pesquisa para observar se os textos apresentam variação de vocabulário muito grande e, por meio de uma comparação com vocabulários estudados nas unidades e níveis anteriores aos que os textos se apresentam, verificar se a variabilidade é sustentada pela prévia apresentação das palavras (ou seja, se as palavras utilizadas no texto já foram apresentadas/estudadas anteriormente e, caso não tenham sido, se a inferência de significado é possível). Além disso, por meio da pesquisa com os alunos reais, iremos observar se a variabilidade lexical corresponde também ao desconhecimento de determinadas palavras que sejam do léxico do inglês britânico (fato que observamos com frequência em sala de aula).

Conforme mencionado anteriormente, a maior parte das métricas disponíveis pelo Coh-Metrix não tem como principal finalidade a análise de complexidade textual. Por esse motivo, Silva (2018, p.144) desenvolveu um quadro que organiza os intervalos numéricos obtidos pela métrica de TT em relação a possíveis interpretações de complexidade, conforme Quadro 22 abaixo:

Quadro 22 - Interpretação dos valores de TT para estimar complexidade

COMPLEXIDADE	FAIXA DE VALORES DE TT
Fácil:	Abaixo de 680
Médio:	679 a 749
Difícil:	Acima de 750

Fonte: Elaboração própria

Conforme já realizamos com SIMS, também multiplicaremos os valores de TT por mil para facilitar a interpretação, uma vez que o Coh-Metrix resulta valores de TT entre zero e um e com três números após a vírgula.

Em suma: maior valor de TT representa maior variedade lexical que, por sua vez, pode representar menor coesão e, conseqüentemente, maior potencial de complexidade.

6.2.12 RESUMO

A próxima página, propositalmente isolada das demais, apresenta um resumo dos índices que serão trabalhados nessa tese e suas referências de valor para a complexidade textual. Com o objetivo de facilitar a leitura e interpretação, a página pode ser destacada da tese impressa para servir de apoio na leitura da seção de resultados.

Caso a leitura esteja sendo feita em tela, é possível usar o comando de busca e digitar “resumo das métricas” para vir diretamente para esta página.

O quadro presente na próxima página (Quadro 23) apresenta um resumo das métricas trabalhadas nesta tese.

Após o quadro, apresentamos uma seção com a descrição dos procedimentos metodológicos desta pesquisa.

Quadro 23 - Resumo das métricas utilizadas na tese

MÉTRICA	UTILIZAÇÃO	VALORES DE REFERÊNCIA	
Narratividade (NT)	Proximidade a estruturas de narrativas.	Fácil:	80 a 100
		Médio:	40 a 79
		Difícil:	0 a 39
Simplicidade Sintática (SS)	Simplicidade ou complexidade da estrutura e relação entre as palavras das frases.	Fácil:	75 a 100
		Médio:	50 a 74
		Difícil:	0 a 49
Similaridade Sintática (SIMS)	Representa a similaridade ou disparidade na formação sintática de todas as frases do texto.	Fácil:	Acima de 101
		Médio:	80 a 100
		Difícil:	Abaixo de 80
Coesão Referencial (CR)	Analisa coesão entre pronomes e substantivos.	Fácil:	50 a 100
		Médio:	31 a 49
		Difícil:	0 a 30
Índice Flesch (IF)	Complexidade baseada em extensão de palavras e frases.	Fácil:	70 a 100
		Médio:	61 a 69
		Difícil:	0 a 60
Polissemia (PS)	Média de quantidade de palavras polissêmicas.	Fácil:	Acima de 4,05
		Médio:	3,49 a 4,04
		Difícil:	Abaixo de 3,50
Idade de Aquisição (IA)	Valor que representa o quão próximo ou distante do léxico de uma criança (mais acessível) é a média das palavras do texto.	Fácil:	Abaixo de 275
		Médio:	274 a 349
		Difícil:	Acima de 350
Familiaridade com palavras de conteúdo (FPC)	Média da familiaridade das palavras de conteúdo do texto.	Fácil:	Acima de 580
		Médio:	564 a 579
		Difícil:	Abaixo de 565
Concretude Lexical (CLP)	Percentual de palavras cujo significante evoca um significado concreto no mundo real.	Fácil:	80 a 100
		Médio:	55 a 79
		Difícil:	0 a 54
Relação Type-Token (TT)	Variedade lexical.	Fácil:	Abaixo de 680
		Médio:	679 a 749
		Difícil:	Acima de 750

Fonte: Elaboração própria

- Fim da página destacada para referência -

7 PROCEDIMENTOS METODOLÓGICOS

Conforme citado na seção 2.3 ('Análises de Acessibilidade Textual'), análise de AT e de ATLD são multimodais e orientadas de acordo com os objetivos de cada estudo e com o perfil de cada leitor-determinado.

Para esta pesquisa, nossas escolhas metodológicas foram orientadas tendo em mente estudantes de inglês como língua adicional, adultos, brasileiros e cursando disciplinas de inglês na universidade. Além disso, como o principal objetivo desta pesquisa é uma observação abrangente das características que perfazem a acessibilidade-manutensiva, consideraremos alunos dos níveis elementares e pré-intermediários (classificados entre A1 e B1 de acordo com o CEFR⁶¹). Portanto, nossos procedimentos metodológicos seguem a seguinte ordem:

I. Coleta e preparo do *corpus* de estudo: nesta etapa, coletamos todos os textos presentes nos livros dos níveis *elementary* e *pre-intermediate* da coleção *English File* (conforme descrito na seção 6.1 *Corpus*) e os preparamos para submissão ao Coh-Metrix – removendo imagens, formatações e gráficos;

II. Coleta e preparo do *corpus* de comparação: nesta etapa, coletamos 10 textos simples e 10 textos complexos em língua inglesa (conforme descrito na seção 6.2) cujas métricas servirão como *tertio comparationis* para a análise semiautomática. O preparo se deu da mesma maneira conduzida no primeiro procedimento;

III. Redação de textos para *corpus* progressivo: redigimos 20 textos apresentando elementos e níveis com progressão de complexidade;

IV. Submissão dos *corpora* ao Coh-Metrix: os *corpora* – de estudo, comparativo e progressivo – foram submetidos ao Coh-Metrix para que as métricas fossem coletadas. Uma vez que a ferramenta é de PLN (e não de LC), os textos foram submetidos individualmente;

V. Coleta das métricas (definidas no capítulo 7): as métricas estipuladas foram coletadas por meio do Coh-Metrix e organizadas em tabelas comparativas;

⁶¹ É importante destacar que compreendemos que o CEFR oferece orientações que auxiliam na “padronização” de conteúdos, características e habilidades de uma língua adicional. Contudo, precisamente por oferecerem modelos de padronização, os documentos normativos tendem a ser excludentes. Para estudantes de línguas adicionais que falem o Português como língua materna, por exemplo, o CEFR proporciona informações/diretrizes coerentes. Contudo, acreditamos que os documentos devem ser utilizados como referências e não como normas absolutas, visto que cada estudante tem peculiaridades específicas (conforme discutiremos posteriormente) e que línguas não-ocidentais (como o Mandarim, por exemplo) não são tão eficientemente consideradas pelas diretrizes do CEFR.

VI. Análise semiautomática dos potenciais de complexidade: este procedimento foi conduzido em quatro etapas: 1. Análise e interpretação dos valores métricos com base em seus sistemas próprios de classificação de potencial de complexidade; 2. Contraste entre os valores métricos dos textos do *corpus* de estudo e dos *corpora* comparativo e progressivo; 3. Estudo holístico e comparativo das estruturas, conteúdos gramaticais e lexicais dos textos e; 4. Verificação da progressão das métricas com finalidade de observar avanço nos insumos linguísticos para garantir a manutenção linguística.

VII. Proposições de diretrizes de acessibilidade linguodidática: com base nos resultados de complexidade provenientes das métricas e das opiniões dos alunos, elaboraremos instruções para redação e/ou edição de textos em língua inglesa voltados para o ensino de forma que os textos possam potencializar o aprendizado do idioma.

Na próxima parte desta tese, apresentaremos e discutiremos os resultados obtidos da coleta das métricas dos nossos *corpora*, bem como as demais análises propostas por este trabalho.

8 RESULTADOS

Neste capítulo, apresentamos e discutimos os resultados da nossa pesquisa, iniciando pelos dados coletados por meio do Coh-Metrix e interpretações relativas aos potenciais de complexidade das métricas escolhidas.

Com o objetivo de facilitar a interpretação das métricas a respeito do potencial de complexidade apresentaremos, nesta seção, os dados numéricos coletados por meio do Coh-Metrix dispostos em tabelas com os seguintes rótulos de cores (Figura 10):

Figura 10 - Rótulos de cores para interpretação do potencial de complexidade das métricas



Fonte: Elaboração própria

Além da classificação das cores verde, amarelo e vermelho, as gradações de cor também indicam maior ou menor complexidade dentro de cada classificação. Sendo assim, entende-se que:

- a. O verde mais escuro indica menor complexidade, enquanto o verde mais claro, maior complexidade dentro da escala de complexidade baixa;
- b. O amarelo mais escuro indica menor complexidade, enquanto o amarelo mais claro, maior complexidade dentro da escala de complexidade média;
- c. O vermelho mais claro indica menor complexidade, enquanto o vermelho mais escuro, maior complexidade dentro da escala de complexidade média;

Iniciaremos a análise na próxima subseção por meio dos dados dos textos do *corpus* comparativo e progressivo.

8.1 ANÁLISE SEMIAUTOMÁTICA DOS *CORPORA* PROGRESSIVO E COMPARATIVO

Com a finalidade de desenvolvermos *tertio comparationis*, formamos dois *corpora* de textos: um composto por dez textos simples (LIS) e por dez textos complexos (LIC) em língua inglesa (*corpus* comparativo) e outro formado por vinte textos escritos por nós com progressão de complexidade (*corpus* progressivo).

O propósito do *corpus progressivo* é auxiliar a observar as métricas elegidas de maneira mais detalhada com a finalidade de permitir comparações mais ricas e, também, de estipular diretrizes para interpretação de algumas métricas em relação a potenciais de complexidade. Para isso, utilizaremos os valores das métricas em gráficos-linha que nos permitirão observar o comportamento das métricas em relação aos recursos de complexidade adicionados por nós no *corpus* progressivo.

Já os objetivos do *corpus* comparativo são:

I. Contrastar os valores de textos simples e de textos complexos com os textos do nosso *corpus* de estudo, uma vez que acreditamos que os materiais do livro *Elementary* devem apresentar valores mais próximos aos valores dos textos simples, enquanto o dos textos do livro *Intermediate* devem apresentar métricas mais próximas (mas nem tanto) dos valores dos textos complexos do *corpus* comparativo.

II. Os valores das métricas nos permitem observar o comportamento de cada nível textual em textos mais acessíveis e em textos menos acessíveis para que

possamos desenvolver uma “régua” comparativa – uma vez que, conforme anteriormente mencionado, essas métricas de textualidade não são voltadas especificamente para estimativas de complexidade; necessitam de uma interpretação de linguistas.

Na próxima seção, iniciaremos nossa análise pelos valores métricos do *corpus progressivo*.

8.1.2 Análise dos resultados das métricas dos textos do corpus progressivo

O quadro abaixo (Quadro 24) exhibe os valores das métricas dos textos do *corpus progressivo*, cujos valores nos auxiliarão a: I. Classificar cada índice a respeito de complexidade - uma vez que, por termos total controle dos textos uma vez que, propositalmente e estudadamente, aumentamos a progressão de dificuldade, podemos enxergar como cada índice se comporta diante das modificações adicionadas e; II. Desenvolver gráficos-linha que nos auxiliarão, visualmente, a compreender e interpretar o comportamento dos índices.

Quadro 24 - Valores das métricas do *corpus progressivo*

	SINTAXE					SEMÂNTICA E LÉXICO				
	NT	SS	SIMS	CR	IF	PS	IA	FPC	CLP	TT
CP1	0,66	6,55	0	81,59	100	2,33	183,00	645,00	100	100
CP2	93,45	100,00	100	100	100	3,50	183,00	626,50	100	33,3
CP3	89,80	100,00	100	100	100	3,17	183,00	626,50	100	33,3
CP4	14,92	100,00	100	100	100	2,92	183,00	616,75	100	70
CP5	42,47	100,00	60	99,46	86,45	2,55	209,00	615,88	100	68,8
CP6	50,80	100,00	68	96,41	95,42	2,57	225,33	600,92	100	63,6
CP7	57,53	100,00	66,8	92,79	98,31	2,88	225,33	601,93	99,86	68
CP8	71,90	100,00	64,6	94,84	96,91	2,75	225,33	592,81	99,95	64,5
CP9	64,80	100,00	58,3	91,77	91,72	2,82	225,33	592,94	99,93	68,6
CP10	22,66	100,00	58,3	98,78	91,72	2,82	225,33	592,94	98,96	68,6
CP11	19,49	99,55	30,1	97,50	90,85	2,82	225,33	592,94	100	70,3
CP12	20,33	98,34	24,5	99,61	89,60	2,82	225,33	592,94	100	70,3
CP13	80,78	92,22	17,1	8,38	88,48	2,92	226,00	591,88	99,96	79,5
CP14	62,55	76,11	14,1	16,35	85,35	2,90	274,13	587,63	99,94	68,8
CP15	51,60	75,17	13,2	13,79	73,59	4,12	286,22	594,62	99,34	62,5
CP16	72,24	77,94	12,9	13,57	77,55	3,93	294,30	594,81	98,75	67,5
CP17	42,47	42,86	9,2	9,34	57,66	3,34	329,35	581,70	91,47	59,3
CP18	48,80	37,45	8,8	8,23	56,35	3,32	310,76	587,10	57,53	57,2
CP19	20,05	10,56	6,1	5,59	39,20	3,33	349,90	564,01	84,38	53,6
CP20	18,94	3,51	5,1	6,55	39,03	3,11	349,12	562,74	84,85	54,1

Fonte: Elaboração própria

Legenda: PS: Polissemia | IA: Idade de aquisição | FPC: Familiaridade com palavras de conteúdo | CLP: Concretude Lexical | TT: Type-Token | NT: narratividade | SS: Simplicidade sintática | SIMS: Similaridade Sintática | CR: Coesão Referencial | IF: Índice Flesch

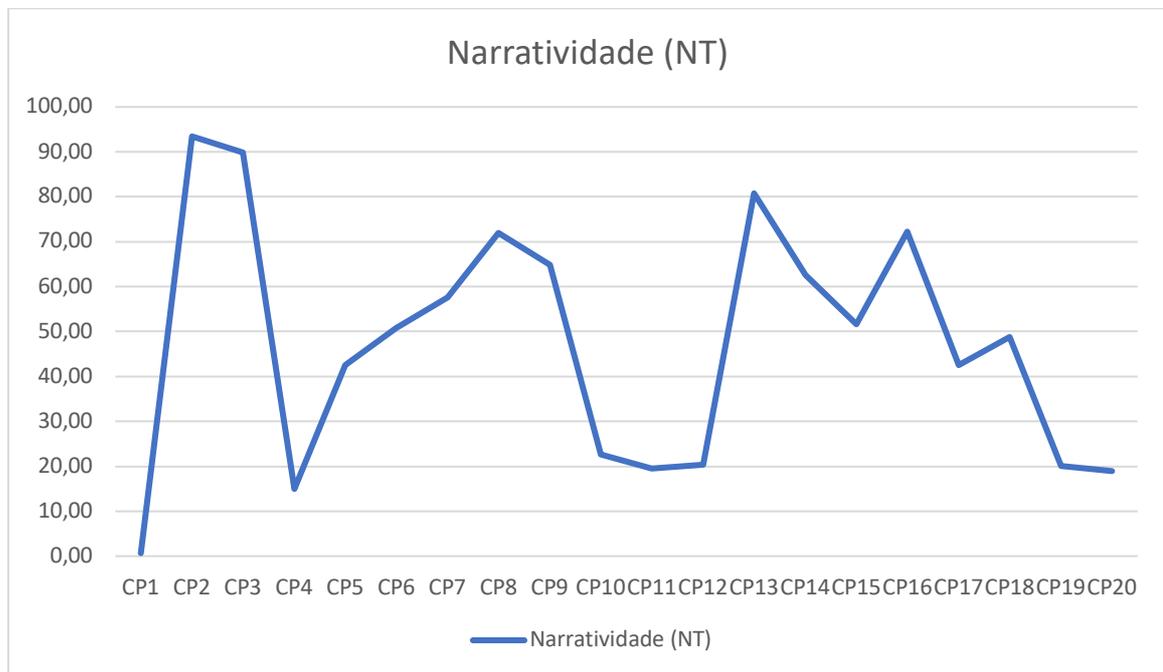
A partir dos dados métricos do *corpus progressivo*, desenvolvemos gráficos-linha que nos permite estudar como os valores se alteram na progressão dos textos. Vale lembrar que os textos estão apresentados na linha horizontal do gráfico em ordem de complexidade crescente (do CP1 ao CP20) e a linha representa o comportamento de cada uma das métricas trabalhadas.

Iniciaremos nossas observações a respeito do comportamento das métricas pela NT.

8.1.2.1 Comportamento da métrica Narratividade (NT)

O gráfico abaixo (Gráfico 1) nos permite observar o comportamento de NT à medida que a complexidade dos textos aumenta:

Gráfico 1 - Comportamento de NT dos textos do *corpus progressivo*



Fonte: Elaboração própria

É possível constatar pela linha indicativa de Narratividade que ela não apresenta comportamento constante. Entretanto, tal fenômeno é esperado, uma vez

que a estrutura sintática dos textos se altera, ora se distanciando e ora se aproximando de padrões de narrativas.

O texto CP1 apresentou valor zero de NT, uma vez que é composto apenas por um fragmento de frase (artigo e substantivo), e não uma oração.

É interessante observar que a maior queda desse índice foi do texto CP9 para o CP10 – onde houve modificação da ordenação das frases com o objetivo de afastar os sujeitos de seus pronomes e, assim, interrompendo um fluxo esperado em uma narrativa.

A elevação da métrica NT do texto CP12 para o CP13 indica a adição de uma frase coordenada por meio dos conectivos ‘*and*’ e ‘*but*’.

Já a queda quase constante entre os textos CP13 e CP20 pode ser resultado dos apostos inseridos com o propósito de aumentar o vocabulário e a extensão das palavras e de frases. Percebemos uma leve melhora entre os textos CP15 e CP16, onde foram realizadas modificações em pontuação e inserção de outras frases coordenadas.

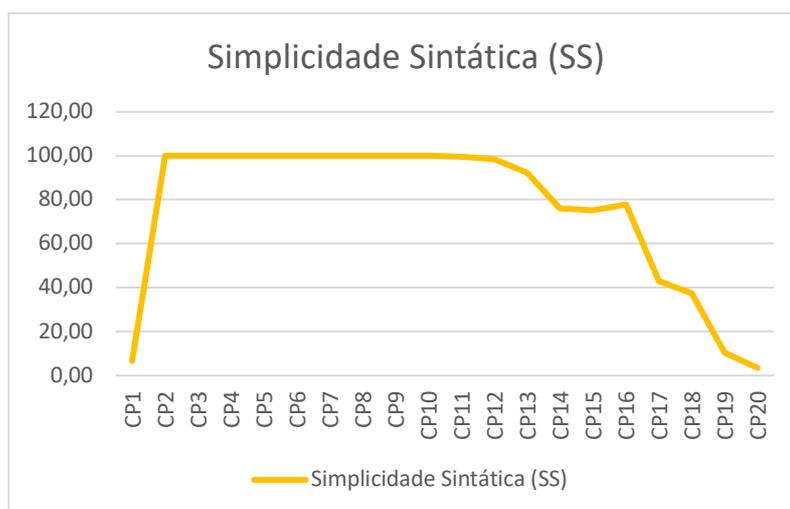
Também podemos perceber que o valor de NT dos textos CP19 e CP20 são considerados como indicativos de complexidade muito alta – fator que está atrelado à adição de apostos e à quebra de sequenciamento narrativo.

No que diz respeito à complexidade, então, a métrica NT não pode ser analisada sozinha. McNamara, et al. (2014) explicam que textos mais simples são aqueles mais próximos a narrativas, mas acreditamos que esse índice sozinho não dá conta de estimar complexidade com exatidão.

A seguir, analisaremos o comportamento da métrica SS no *corpus* progressivo.

8.1.2.2 Comportamento da métrica Simplicidade Sintática (SS)

Como podemos perceber pela linha que representa SS no gráfico a seguir (Gráfico 2), essa métrica apresenta um comportamento constante.

Gráfico 2 - Comportamento de SS dos textos do *corpus* progressivo

Fonte: Elaboração própria

Uma vez que o texto CP1 não é uma oração, sua valoração de SS é nula.

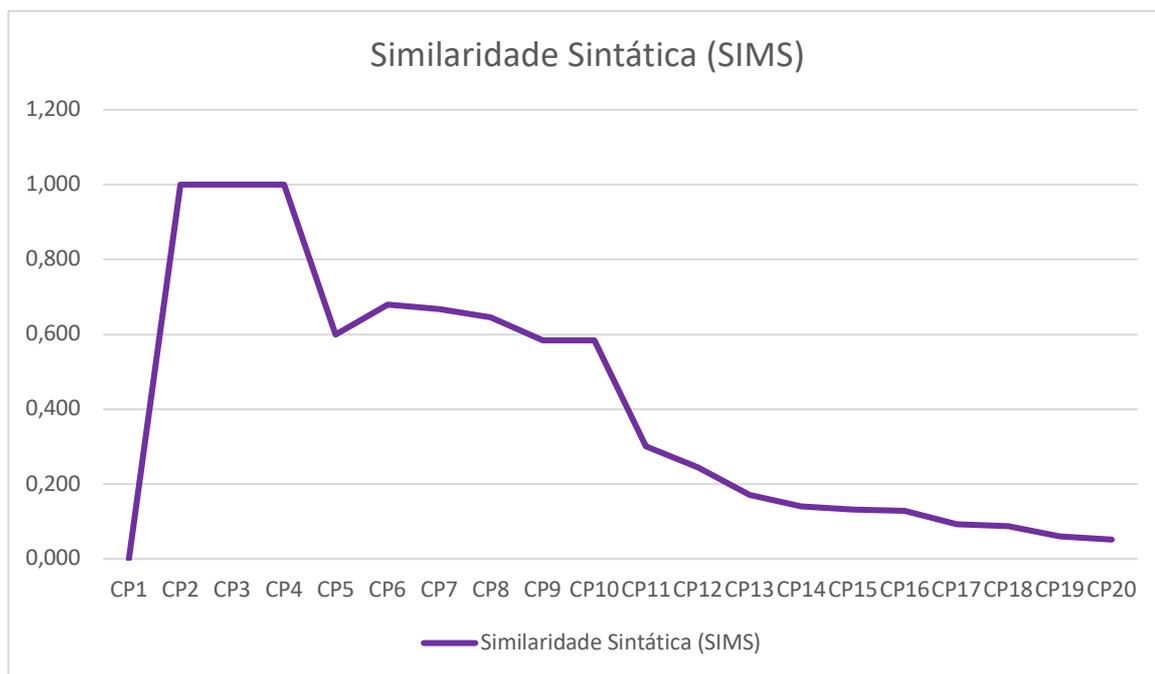
A partir do texto CP2, podemos observar uma queda constante nos valores de SS, sendo uma constância maior entre os textos CP2 e CP12 e uma queda muito perceptível entre os textos CP16 e CP20 – que pode ser justificado pelo fato de que a partir do texto CP16, a inserção de termos especializados é progressiva.

Portanto, para a utilização da SS como indicativo de complexidade, podemos concluir que textos sintaticamente mais acessíveis têm valoração na faixa entre 80 e 100 e que, à medida que esses valores diminuem, a complexidade sintática aumenta (como nos casos dos textos CP19 e CP20 que apresentam menores valores de SS por serem altamente terminológicos).

Na próxima subseção, estudaremos o comportamento da SIMS no *corpus* progressivo.

8.1.2.3 Comportamento da métrica Similaridade Sintática (SIMS)

Por meio do gráfico a seguir (Gráfico 3), podemos estudar o comportamento da SS em relação à progressão de complexidade dos textos do *corpus* progressivo:

Gráfico 3 - Comportamento de SIMS dos textos do *corpus* progressivo

Fonte: Elaboração própria

A análise do comportamento da Similaridade Sintática é simples: entre os textos CP1 e CP5, percebemos o valor zero para essa métrica. Isso ocorre porque esses textos ou são constituídos de apenas uma frase, ou são formados por pares de orações com estruturas sintáticas exatamente paralelas (ou seja, mesma formação). Sendo assim - conforme explicado anteriormente a respeito da SimS – o valor zero representa exatidão na similaridade sintática entre as frases. O mesmo fenômeno ocorre com o texto CP12.

Entretanto, é evidente que tal exatidão sintática é muito rara. Percebemos seis ocorrências no nosso *corpus* progressivo somente porque os textos foram propositalmente controlados para que houvesse paralelismo entre as orações.

Já os textos com maiores ocorrências de coordenação e subordinação, com maiores distinções entre a quantidade de palavras antes e depois dos verbos e com maior adição de apostos, os valores, obviamente, alteram-se.

Entre os textos CP6 e CP9, percebemos constância no comportamento da SimS. Isso ocorre porque ainda existem pares de frases com estruturas sintáticas paralelamente exatas, mas outras (poucas) que não possuem pares sintáticos.

Os valores de SimS decaem entre os textos CP9 e CP12 ao passo que mais frases sintaticamente semelhantes são adicionadas.

Do texto CP12 ao CP13, percebemos a maior diferença de valores de SimS; indicando que o texto CP12 apresenta exatidão em paralelismo sintático, enquanto o CP13, maior distinção na formação das frases – conforme explicado na seção de métricas desta tese.

A partir do texto CP14, percebemos determinada constância de progressão; o que reflete a adição de novas frases aos textos e novos processos de coordenação e subordinação que não seguem nenhum padrão estrutural sintático.

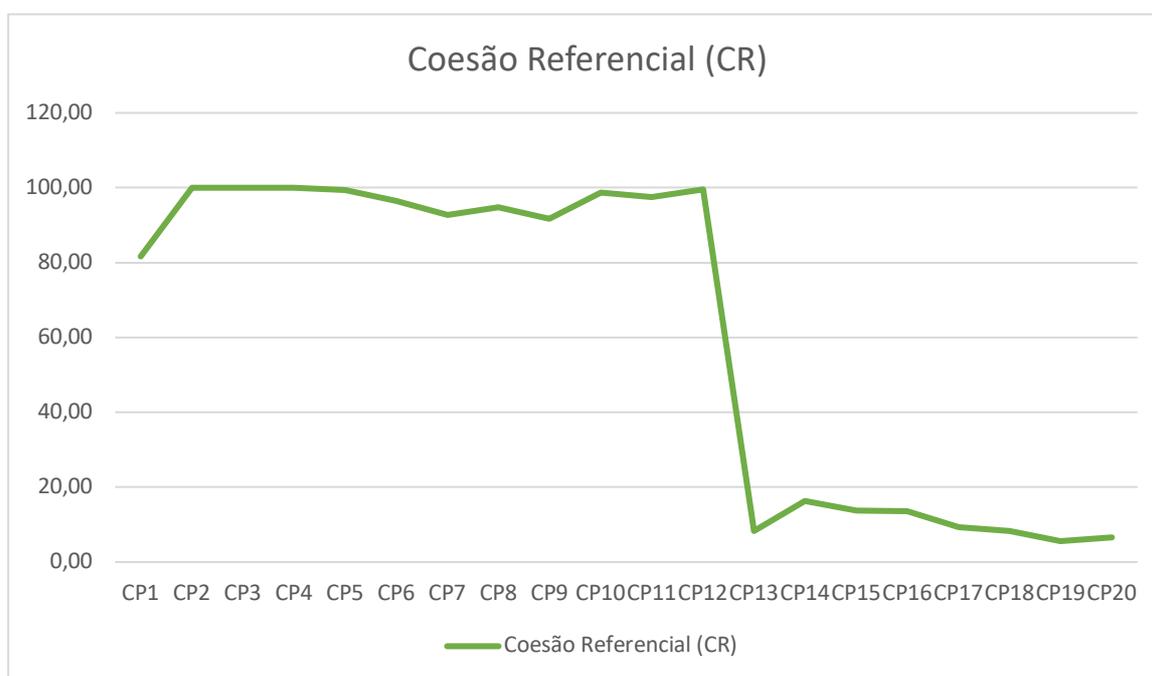
Sendo assim, a leitura do comportamento de SIMS pode ser resumida da seguinte forma: valores mais próximos de zero indicam maior similaridade sintática, enquanto valores mais próximos de um, maior diferença.

A seguir, apresentamos e discutimos o comportamento de CR nos textos do *corpus* progressivo.

8.1.2.4 Comportamento da métrica Coesão Referencial (CR)

Pela linha que representa a CR no gráfico abaixo (Gráfico 4), podemos perceber com facilidade o comportamento da métrica:

Gráfico 4 - Comportamento de CR dos textos do *corpus* progressivo



Fonte: Elaboração própria

Do primeiro ao segundo texto, percebemos um avanço da métrica. Isso ocorre porque no texto CP1 não temos nenhuma relação entre os elementos da frase. Já, no texto CP2, percebemos a coesão referencial entre o sujeito e a desinência verbal.

Do texto CP12 para o CP13 percebemos o maior declínio de CR – que se dá devido à intercalação de pronomes e da maior distância entre os pronomes pessoais de seus referentes.

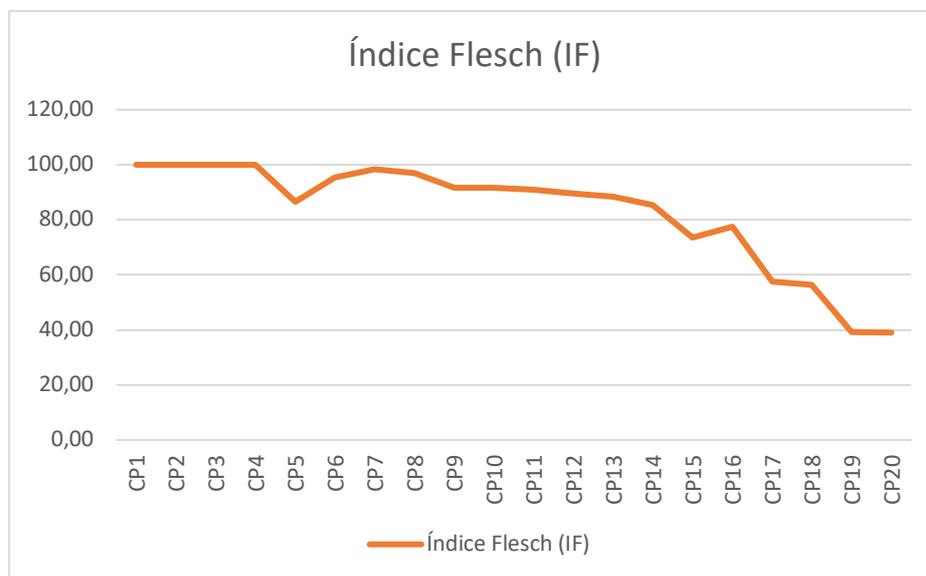
Do texto CP14 em diante percebemos um constante declínio (ainda que de forma mais discreta), que é consequência da maior extensão do texto que “empurra” os pronomes para mais longe dos seus sujeitos referentes devido às adições de apostos, adjuntos e outros elementos deslocados.

Portanto, a métrica CR nos demonstra que quanto mais possibilidades de referências possíveis (quando existem grandes quantidades de pronomes que poderiam se referir a diferentes sujeitos/objetos e quando as distâncias entre as palavras que se relacionam são grandes), menores são os valores da métrica.

Dando sequência a análise de comportamento das métricas no *corpus* progressivo, interpretamos, a seguir, o IF.

8.1.2.5 Comportamento da métrica Índice Flesch (IF)

Uma vez que o IF é a única métrica que foi desenvolvida especificamente para estimar complexidade e cujas orientações são pré-definidas, o comportamento da métrica era previsível: textos com maiores extensões teriam valores mais baixos de IF, que indicam maiores complexidades. O Gráfico 5 a seguir nos permite observar tal comportamento:

Gráfico 5 - Comportamento de IF dos textos do *corpus* progressivo

Fonte: Elaboração própria

Conforme o esperado, a partir do texto CP7 podemos perceber um declínio constante nos valores de IF – uma vez que a extensão das frases, das palavras e dos parágrafos também aumenta.

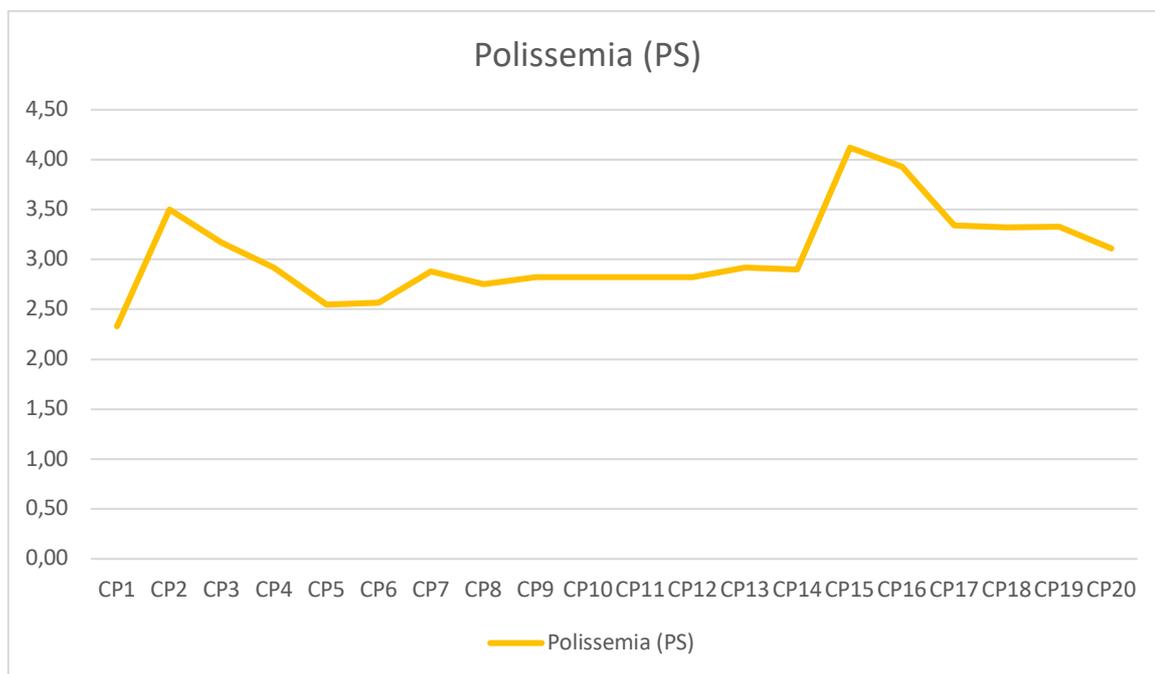
Embora possamos observar um aumento entre os textos CP4 e CP7, ao verificamos os esses textos individualmente percebemos que ainda que mais frases tenham sido inseridas, as palavras nas frases são curtas (como ‘*he*’, ‘*boy*’, ‘*cat*’ e ‘*dog*’), o que impacta diretamente na proporção do cálculo de IF, já que temos mais frases, mas redigidas com léxico de palavras curtas.

A relação entre IF e complexidade já é definida pelas diretrizes do próprio índice (uma vez que ela é a única métrica que é voltada primeiramente e especificamente para análise de complexidade textual): quanto mais baixos os valores de IF, mais complexos os textos são.

Para iniciar as observações dos comportamentos das métricas relacionadas ao léxico e semântica, começaremos nossa análise pela métrica PS na próxima seção.

8.1.2.6 Comportamento da métrica Polissemia (PS)

As indicações do comportamento da métrica de polissemia são extremamente claras por meio do gráfico visual (Gráfico 6): à medida que a complexidade dos textos aumenta, diminui o valor da métrica polissemia.

Gráfico 6 - Comportamento de PS dos textos do *corpus* progressivo

É importante destacar que o valor da métrica diminui, mas isso não significa menor quantidade de palavras polissêmicas. Ao contrário, menores valores de PS indicam maior quantidade de palavras polissêmicas e, conseqüentemente, maiores potenciais de complexidade textual.

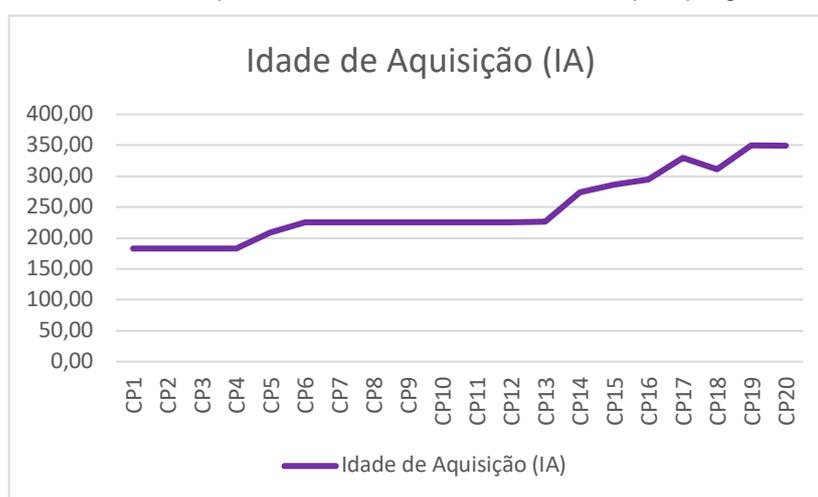
A PS também nos permite validar a necessidade de conduzir análises multimodais (BERBER SARDINHA, 2000) e corroborar a importância da verificação de um linguista. Se analisarmos o comportamento de PS para os textos CP19 e CP20, os mais complexos, acreditamos que eles indicam baixa complexidade em relação aos demais. Entretanto, como sabemos, esses dois textos são os mais complexos no nosso *corpus* progressivo. Tal fato (redução dos valores de PS) ocorre porque foram acrescentados aos textos uma grande quantidade de termos especializados. Embora esses termos aumentem a complexidade dos textos - conforme indicam Silva (2018), Finatto (2018), DuBay (2004) e tantos outros pesquisadores da área - é evidente que a presença de polissemia se reduziria. Uma vez que vocábulos são consideradas termos, seus valores semânticos, geralmente, são únicos (por serem especializados). Sendo assim, maiores quantidades de termos inseridos em um texto tendem a diminuir os indícios de presença de palavras polissêmicas.

A segunda métrica que analisaremos relacionada ao léxico e à semântica é a IA. Na seção a seguir, discutimos o comportamento dessa métrica.

8.1.2.7 Comportamento da métrica Idade de Aquisição

O gráfico abaixo (Gráfico 7) nos permite perceber que os crescimentos nos valores de IA ocorreram, principalmente, entre os textos CP4 e CP6, CP13 e CP17, CP18 e CP19.

Gráfico 7 - Comportamento de IA dos textos do *corpus* progressivo



Fonte: Elaboração própria

Entre os textos CP4 e CP6 foram adicionadas frases novas e, conseqüentemente mais vocabulários.

Entre os textos CP13 e CP17 também foram adicionadas maiores informações, bem como apostos explicativos e alguns termos especializados.

Já entre os textos CP18 e CP19 (intervalo que nos permite perceber o maior “salto” em relação ao comportamento de IA) foi justamente onde adicionamos uma grande quantidade de termos especializados e apostos a respeito desses termos – o que justifica o grande avanço de IA.

A pequena queda de valores percebida entre os textos CP17 e CP18 pode ser justificada pela inserção de frases com vocabulário simples e pelo fato de as novas palavras inseridas serem adjetivos e substantivos considerados simples.

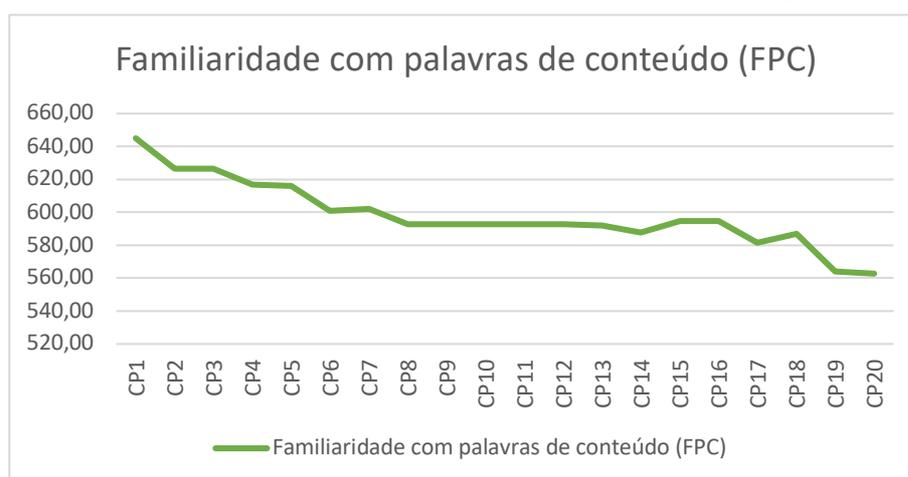
Logo, a relação entre IA e complexidade é clara: textos com mais palavras desconhecidas (ou seja, valores de IA mais altos) são indícios de maiores complexidades.

Uma métrica semelhante à IA e que pode ser utilizada para estimar complexidade lexical e semântica é a FPC, cujo comportamento nos textos do *corpus* progressivo será analisado a seguir.

8.1.2.8 Comportamento da métrica Familiaridade com Palavras de Conteúdo (FPC)

No gráfico abaixo (Gráfico 8), podemos perceber, com clareza, o comportamento da métrica FPC:

Gráfico 8 - Comportamento de FPC dos textos do *corpus* progressivo



Fonte: Elaboração própria

É interessante observar que essa métrica segue o caminho oposto aos valores de IA: textos mais simples apresentam valores mais altos e textos mais acessíveis, FPCs mais baixos.

A maior variação entre os valores se deu entre os textos CP18 e CP19, onde foram inseridos, propositalmente, uma grande quantidade de termos especializados. Por meio da terminologia inserida, podemos perceber um grande impacto nos valores de FPC.

Observamos, também, que embora os textos CP19 e CP20 tenham uma grande quantidade de termos, o impacto dessas palavras nos valores dessa métrica não é muito grande. Isso se justifica pela proporção entre palavras e termos. Ainda que existam muitos termos, os apostos adicionados para explicar essas unidades

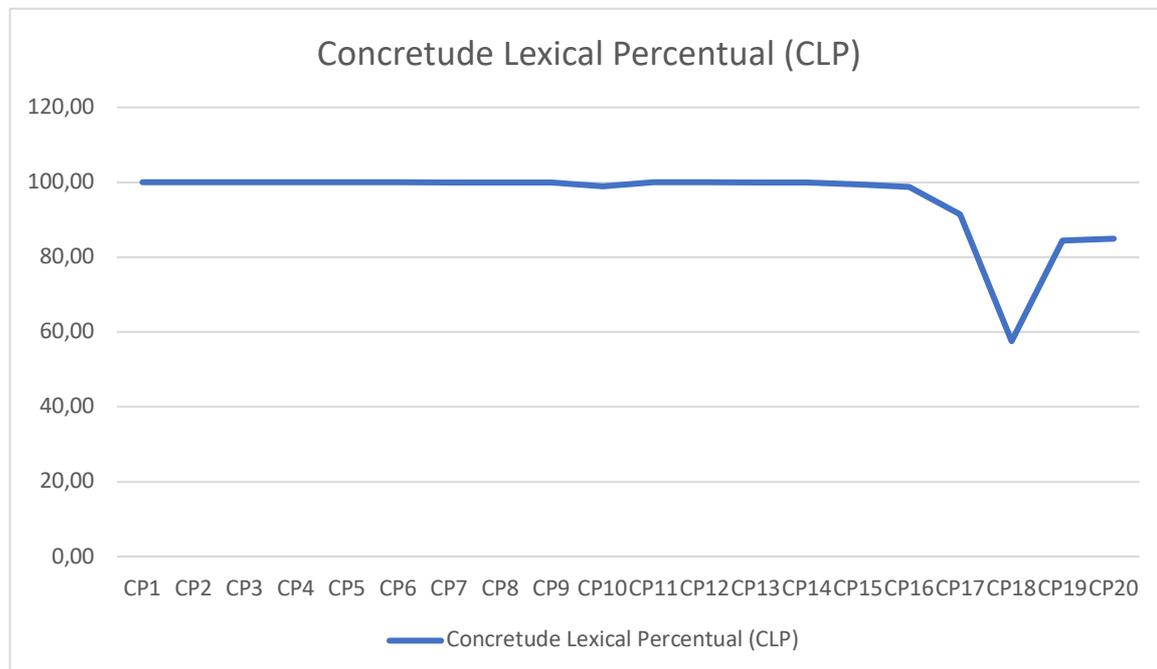
lexicais especializadas foram redigidos com palavras mais comuns. Logo, a proporção entre a quantidade de palavras comuns comparado a termos é, basicamente, 9:1. Dessa forma, percebemos que as valorações de FPC costumam variar na escala entre 500 e 600. Embora a métrica tenha sido originalmente proposta com valores entre 700 (palavras mais comuns) e 100 (palavras menos usuais), só poderíamos obter valores próximos a 700 se analisássemos uma lista de palavras simples. Da mesma forma, só conseguiríamos perceber valores próximos de 100 caso verificássemos os valores de FPC para uma lista de termos altamente complexos.

Na próxima subseção, descrevemos o comportamento da métrica CLP nos textos do *corpus* progressivo.

8.1.2.9 Comportamento da métrica Concretude Lexical (CLP)

O comportamento da métrica CLP pode ser interpretado pela linha do Gráfico 9 a seguir:

Gráfico 9 - Comportamento de SS dos textos do *corpus* progressivo



Fonte: Elaboração própria

Ao compreendermos que a CLP é uma métrica que indica a presença de palavras concretas nos textos, e que esses valores são proporcionais (ou seja, valores mais altos indicam maiores quantidades de palavras concretas), sabíamos de

antemão que nossos textos com alta presença de terminologias médicas (a partir do texto CP15) demonstrariam queda nos valores da métrica.

Conforme podemos observar pela linha, a partir do texto CP15 percebemos, de fato, que os valores de CLP são reduzidos em comparação ao intervalo entre os textos CP1 e CP14.

O aumento que podemos observar entre os textos CP18 e CP19 ocorre devido aos apostos que explicam os termos médicos do texto que foram adicionados ao CP19 e que foram redigidas com maior frequência de palavras concretas.

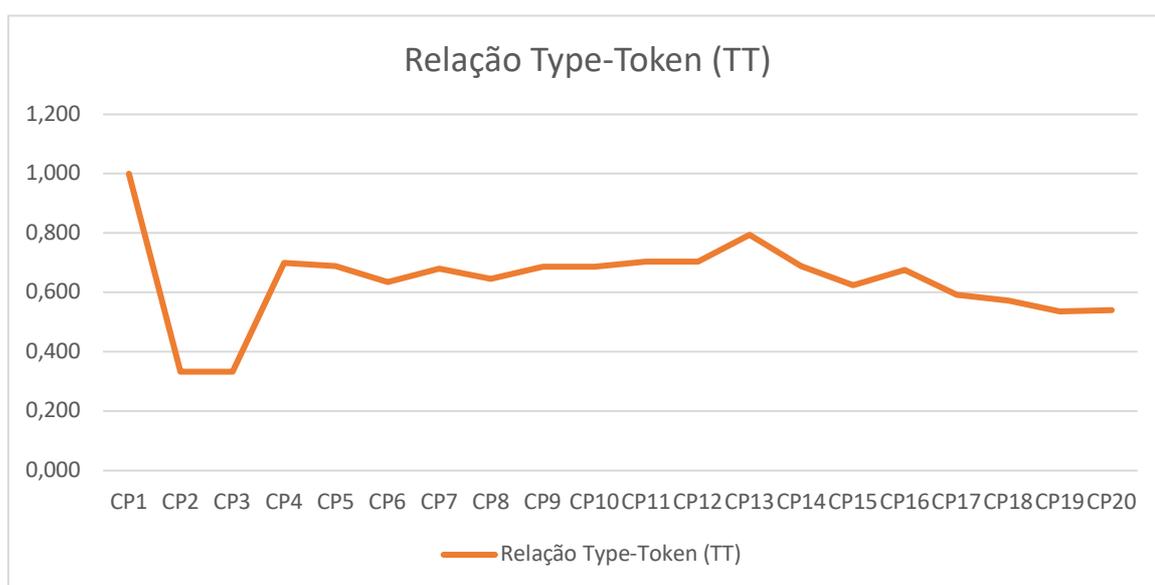
Ao estudarmos a métrica CLP sob a ótica da CT, chegamos à conclusão, então, que valores mais baixos dessa métrica são indícios de maiores potenciais de complexidade.

Finalizando a análise do comportamento das métricas no *corpus* progressivo, seguimos para o último índice de sintaxe: TT.

8.1.2.10 Comportamento da métrica Relação *Type-Token* (TT)

Uma vez que nossos textos do *corpus* progressivo apresentam, em determinados intervalos, adição de novas palavras, o comportamento da métrica – conforme observado no Gráfico 10 abaixo – foi conforme o esperado:

Gráfico 10 - Comportamento de SS dos textos do *corpus* progressivo



Fonte: Elaboração própria

A métrica TT, que lida com variação lexical, tem um comportamento esperado: entre os textos CP5 e CP10, percebemos pouca variação - uma vez que a composição das frases seguia o mesmo padrão e utilizava vocabulários repetidos.

Entre os textos CP10 e CP13, percebemos uma diminuição nos valores da métrica, resultado dos processos de coordenação que nos permitiu eliminar palavras.

Já a partir do texto CP13, podemos perceber um aumento muito rápido dos valores de TT, consequência da adição de novas informações e de muitos novos vocabulários de forma progressiva.

Nos textos CP19 e CP20, podemos observar um fenômeno interessante: os valores de TT reduzem em relação aos anteriores, passando a serem considerados como indicativos de complexidade baixa. Essa alteração corrobora o fato de que não podemos escolher apenas uma métrica para analisarmos complexidade textual; caso contrário, textos altamente complexos (como esses dois casos), poderiam ser considerados “simples” no que diz respeito à variação lexical (quando são, na verdade, altamente complexos).

Além do fato de não utilizarmos apenas uma métrica única para estudos de complexidade textual, também frisamos a importância da análise conduzida por um linguista – profissional que poderá interpretar os valores numéricos e, por meio de um contraste com suas perspectivas diante dos textos, validar as informações ou – como no caso dos valores de TT para os textos CP19 e CP20 - justificar indicações não tão precisas feitas pela máquina. No caso desse exemplo específico, embora os valores de TT tenham sido reduzidos para os dois textos, os demais valores métricos indicaram maiores complexidades. Dessa forma, ainda que pareça ter ocorrido uma variação lexical relativamente controlada, o valor e a complexidade semântica do léxico dos textos são muito mais complexos que dos anteriores.

Como interpretação de complexidade por meio da TT, corroboramos os valores apontados por Silva (2018): valores de TT mais baixos representam complexidades também mais baixas.

Ao finalizarmos a interpretação do *corpus* progressivo, apresentaremos, na próxima seção, a análise dos resultados dos índices do *corpus* comparativo. É importante observar que, além de auxiliar a interpretar os valores das métricas em relação à complexidade (conforme já explicado no capítulo 6 desta tese), esses

valores também são utilizados como pontos de comparação com as métricas do nosso *corpus* de estudo – sobre as quais discutiremos posteriormente.

Na próxima seção, apresentamos a análise das métricas dos textos simples do *corpus* comparativo.

8.2 ANÁLISE DOS RESULTADOS DAS MÉTRICAS DOS TEXTOS SIMPLES DO *CORPUS* COMPARATIVO

O quadro abaixo (Quadro 25) apresenta os valores dos textos simples do *corpus* comparativo, que nos auxiliarão na comparação com os valores do *corpus* de estudo. Uma vez que esses textos (LIS) são considerados, de antemão, simples, os valores de suas métricas nos permitirão observar se os materiais que são o foco da nossa análise (dos livros da coleção *English File*) têm resultados mais próximos a eles (e, conseqüentemente, indicam maior potencial de acessibilidade), ou aos textos complexos (LIC), sobre os quais falaremos posteriormente, (e, em consequência, maior potencial de complexidade).

Para a interpretação dos valores das métricas de LIS, uma vez que os textos não são o foco dos nossos estudos, utilizaremos as legendas de cores sem escala de complexidade. Ou seja, destacaremos em verde os textos considerados simples, em amarelo os textos de complexidade média e, em vermelho, de complexidade alta. Não representaremos maiores ou menores valores das métricas por meio de gradação das três cores (conforme faremos para os textos do *corpus* de estudo).

Ainda é válido lembrar que nossos apontamentos sobre a interpretação dessas métricas não é o foco desta tese. Nossa análise para as métricas do *corpus* comparativo tem a finalidade principal de orientar e organizar estatisticamente os valores métricos do *corpus* de estudo.

Quadro 25 - Valores das métricas de sintaxe dos textos simples do *corpus* comparativo

	SINTAXE					SEMÂNTICA E LÉXICO				
	NT	SS	SIMS	CR	IF	PS	IA	FPC	CLP	TT
LIS1	39,74	34,46	89	41,29	83,04	3,86	296,61	578,23	98,34	709
LIS2	57,93	93,45	93	7,21	71,48	4,18	313,46	580,70	9,01	687
LIS3	72,24	95,54	158	30,15	91,40	4,33	256,64	569,39	93,70	579
LIS4	63,68	66,28	152	32,28	88,01	4,47	272,79	586,21	72,57	747
LIS5	54,78	64,80	98	32,28	69,58	3,52	335,46	580,82	85,08	650
LIS6	90,99	98,81	171	56,36	93,99	4,04	216,66	590,42	70,19	479
LIS 7	69,15	82,38	141	11,31	77,43	4,12	232,98	589,93	87,70	756

LIS 8	69,15	98,81	171	65,54	100	3,58	235,56	592,84	92,22	520
LIS 9	91,31	94,52	170	90,66	78,07	3,68	304,74	565,38	73,57	556
LIS10	95,73	92,79	70	88,30	79,30	4,43	273,26	580,07	92,79	608
LIS11	23,27	91,31	240	37,07	86,76	5,07	246,83	591,74	41,68	521
LIS12	81,06	73,24	208	51,20	92,11	4,22	271,94	589,42	19,22	693
LIS13	91,31	85,31	127	71,23	94,40	3,71	257,75	566,60	98,12	600
LIS14	81,86	87,29	99	19,49	83,96	3,76	306,80	564,61	45,62	798
LIS15	91,15	69,50	96	47,21	90,60	4,53	272,26	574,37	99,13	861
LIS16	49,60	51,20	136	58,71	80,31	3,49	230,87	563,69	90,99	768
LIS17	89,97	45,62	80	66,28	85,19	3,63	254,04	578,91	91,31	696
LIS18	78,52	71,90	189	65,91	89,16	4,60	262,25	584,35	99,72	736
LIS19	79,39	33,00	86	87,49	85,27	4,10	216,56	578,57	98,30	685
LIS20	85,54	81,59	98	79,67	83,36	4,14	286,18	571,09	90,49	568
Média:	72,82	75,59	134	51,98	85,17	4,07	267,18	578,87	77,49	661

Fonte: Elaboração própria

Legenda: NT: Narratividade | SS: Simplicidade sintática | SIMS: Similaridade Sintática | CR: Coesão Referencial | IF: Índice Flesch | PS: Polissemia | IA: Idade de aquisição | FPC: Familiaridade com palavras de conteúdo | CLP: Concretude Lexical | TT: *Type-Token*

Pela interpretação dos valores dos textos simples do *corpus* comparativo (auxiliado pelo Quadro 19), podemos concluir o seguinte a respeito de cada métrica:

- I. **Narratividade (NT):** valores variam entre aproximadamente 23% e 95%, com uma média de 72% - fato que indica que os textos simples do nosso *corpus* comparativo são altamente aproximados de estruturas de narrativas e, pela interpretação dessa métrica, adequadamente acessíveis. 45% dos textos têm indícios de baixa complexidade por essa métrica, enquanto 50%, de média e 5%, de alta. Entretanto, ao observarmos os valores das métricas classificadas como indicativas de complexidade média, é possível verificar que a maior parte delas têm valoração no limítrofe entre a complexidade média e baixa;
- II. **Simplicidade Sintática (SS):** o valor médio de SS também sugere alta acessibilidade, uma vez que a média é de 75,59%, com apenas três casos com valores abaixo de 50% (LIS1, LIS17 e LIS19). Ao considerarmos os valores de SS por meio da tabela de interpretação de potencial de complexidade descrita anteriormente, observamos que a média dessa métrica para os textos LIS os caracteriza como tendo baixa complexidade, ou seja, de leitura fácil. Ainda que alguns textos tenham sido classificados como tendo complexidades médias e altas por essa métrica, é válido lembrar que a maior parte dos textos LIS são narrativas,

tipologia textual que, geralmente, apresenta distintas formações sintáticas para a construção das histórias e frequentes deslocamentos de elementos de suas posições canônicas com a finalidade de destacar informações (como, por exemplo, deslocado um adjunto adverbial para o início da frase com a finalidade de destacar um local ou tempo de um conto);

- III. **Similaridade Sintática (SIMS):** a média de valores de SIMS (134) pode ser classificada como indicativo de baixa complexidade. 55% das métricas são classificadas como indicativas de baixa complexidade, 40%, média e 5%, alta. A variação da métrica é entre 70 e 240. De forma similar ao que já comentamos sobre a métrica SS, o fato da maior parte dos textos LIS serem narrativas impacta nas formações sintáticas dos textos;
- IV. **Coesão Referencial (CR):** os valores de CR variam entre 7,21 e 90,66, sendo uma média de 51,98. Pela interpretação dessa métrica, 55% dos textos podem ser classificados como tendo baixa complexidade (bem como a média), 25%, complexidade média e 20%, complexidade alta. Uma vez que a maior parte dos textos LIS são narrativas, é possível que os resultados indicativos de complexidade média e alta sejam consequência da quantidade de personagens, locais e objetos que compõem a história e – por conseguinte - de pronomes necessários para retomá-los no decorrer das histórias;
- V. **Índice Flesch (IF):** a média de IF é de 85,15% (indicativo de alta acessibilidade), com variações entre 69,58% e 100%. Ainda que apenas um texto tenha sido classificado como indicativo de complexidade média (LIS5: 69,58), o valor de IF está no limítrofe entre as classificações de complexidade baixa e média. Sabendo que o IF está relacionado a extensões textuais, os textos LIS podem ser compreendidos como adequados em tamanhos de palavras e frases;
- VI. **Polissemia (PS):** a média de 4,07 de PS dos textos LIS pode ser classificada como indicativo de baixa complexidade. 55% dos textos são considerados simples, enquanto 45%, médios; sendo a variação de valores entre 3,49 e 5,07.

- VII. Idade de Aquisição (IA):** por serem primordialmente textos voltados para crianças, tínhamos uma hipótese *a priori* de que esses materiais apresentariam resultados indicativos de alta acessibilidade nas escolhas lexicais relacionadas à idade de aprendizagem (ou seja, que seriam textos nos quais haveria grandes quantidades de palavras mais simples para crianças pequenas e, conseqüentemente, mais simples). Essa hipótese foi corroborada, sendo a média de IA dos textos LIS classificada como complexidade baixa (267,18), 70% dos textos com valores que também indicam baixa complexidade e 30%, complexidade média – embora esses textos tenham resultados de IA no limítrofe entre as classificações de complexidade baixa e média;
- VIII. Familiaridade com Palavras de Conteúdo (FPC):** de forma similar à hipótese que levantamos *a priori* sobre a métrica IA, também esperávamos encontrar resultados de FPC que refletissem alta presença de palavras de conteúdo conhecidas no vocabulário do cotidiano. Entretanto, os resultados dessa métrica não foram tão positivos quanto os de IA, sendo 50% deles considerados como baixa complexidade e os outros 50%, média. Entretanto, ao analisarmos cada valor de forma isolada, podemos perceber que muitos deles estão mais próximos de complexidade baixa do que alta (ou seja, valores mais próximos de 580 do que de 565. A média dos valores de FPC (578,87) também é indicativo de complexidade média, mas muito próxima do valor que a classificaria como complexidade baixa (580);
- IX. Concretude Lexical (CLP):** a média de CLP é de 77,49% (classificado como complexidade média) e a variação entre 9,01 e 99,13%. Quanto aos percentuais de interpretação de complexidade, observamos 65% como indicativos de baixa complexidade, 15%, de média e 20%, alta. Ainda que tenhamos resultados que indiquem baixa acessibilidade em determinados textos, é preciso lembrar que os textos LIS são, em grande maioria, narrativas com alta frequência de palavras abstratas (que indicam sentimento ou ideias não concretas no mundo real); sendo assim, justifica-se os valores indicativos de acessibilidade média ou baixa;

- X. Relação Type-Token (TT):** com uma média de 661 (indicativo de complexidade baixa), os valores de TT dos textos LIS variaram entre 479 (maior acessibilidade) e 861 (menor acessibilidade). 45% dos textos podem ser classificados como tendo baixa complexidade, enquanto 35%, média e 20%, alta. Uma vez que a TT é relacionada à variação lexical, é surpreendente termos 45% dos textos classificados como altamente acessíveis. Silva (2018) aponta que em textos narrativos, geralmente há uma grande presença de variação lexical com para a manutenção tema-rema do decorrer dos enredos.

Em suma, uma vez que nosso principal objetivo com os textos LIS é apresentar referências para fins comparativos com os valores das métricas do nosso *corpus* de estudo, podemos confirmar pela interpretação dos números que é possível utilizar os resultados desses textos como indícios de baixa complexidade. Dado que a escolha dos textos foi feita de forma criteriosa observando seus níveis lexicais, semânticos e sintáticos, e que a seleção foi realizada pelo fato desses textos apresentarem características que nós, como linguistas, professores e pesquisadores de complexidade textual, entendemos como potencializadoras de acessibilidade, os resultados são válidos para que possam ser utilizados como parâmetros comparativos.

Na próxima seção, apresentaremos e discutiremos os valores dos índices dos textos complexos (LIC) do *corpus* comparativo.

8.3 ANÁLISE DOS RESULTADOS DAS MÉTRICAS DOS TEXTOS COMPLEXOS DO *CORPUS* COMPARATIVO

O quadro abaixo (Quadro 26) apresenta os valores dos textos complexos do *corpus* comparativo:

Quadro 26 - Valores das métricas dos textos complexos do *corpus* comparativo

	SINTAXE					SEMÂNTICA E LÉXICO				
	NT	SS	SIMS	CR	IF	PS	IA	FPC	CLP	TT
LIC1	10,20	37,83	57	24,20	32,95	3,52	421,67	555,40	1,19	656
LIC2	10,56	74,86	96	2,17	47,21	3,90	383,61	561,15	32,28	689
LIC3	5,82	56,75	107	72,91	5,36	3,08	422,89	546,94	42,86	656
LIC4	39,74	67,36	69	3,84	62,53	3,62	363,76	567,86	40,90	753

LIC5	3,51	69,85	118	10,56	39,43	3,19	372,09	551,46	72,24	773
LIC6	7,49	67,36	79	51,99	19,14	3,01	405,40	548,99	16,85	635
LIC7	5,05	16,60	59	62,55	11,50	3,10	422,39	545,28	28,43	757
LIC8	9,68	19,49	62	45,22	8,79	3,23	407,08	561,79	66,28	759
LIC9	7,64	5,59	43	38,97	8,58	3,29	398,26	562,75	87,29	836
LIC10	5,16	32,28	100	38,59	41,48	3,40	409,29	560,57	23,27	754
LIC11	87,29	62,55	109	7,08	68,69	3,78	344,57	576,12	32,28	789
LIC12	17,62	79,67	96	26,76	43,53	3,62	337,46	556,53	57,93	657
LIC13	52,39	43,64	59	4,36	59,69	4,17	359,69	575,42	47,61	730
LIC14	88,69	59,10	81	16,35	68,41	4,50	337,82	581,95	23,27	699
LIC15	40,90	33,36	43	9,85	47,85	3,85	368,93	576,06	70,54	739
LIC16	92,36	30,85	84	28,10	70,33	3,66	314,94	578,29	55,96	766
LIC17	24,20	50,80	81	3,29	56,19	3,60	341,68	562,99	61,41	783
LIC18	14,92	53,59	98	8,38	53,18	3,29	374,78	554,33	48,01	753
LIC19	55,96	1,70	32	28,43	57,24	3,44	272,96	572,00	93,70	803
LIC20	89,44	58,71	104	50,00	81,37	3,66	282,32	572,78	82,12	698
Média:	33,43	46,10	79	26,68	44,17	3,55	367,08	563,43	49,22	734

Fonte: Elaboração própria

Legenda: NT: Narratividade | SS: Simplicidade sintática | SIMS: Similaridade Sintática | CR: Coesão Referencial | IF: Índice Flesch | PS: Polissemia | IA: Idade de aquisição | FPC: Familiaridade com palavras de conteúdo | CLP: Concretude Lexical | TT: *Type-Token*

Visualmente, por meio das etiquetas de cor, podemos perceber que os valores das métricas dos textos LIC se diferem altamente dos resultados previamente comentados dos índices dos textos LIS. A análise dos valores dos textos complexos do *corpus* comparativo (Quadro 20), permite que possamos concluir a respeito das métricas:

- I. **Narratividade (NT):** os valores de NT dos textos LIC variam entre 3,51 e 92,36, sendo a média de 33,43. Além da média ter sido classificada como indicativo de alta complexidade, 60% dos textos também indicam que a NT pode ser uma característica de baixa acessibilidade, enquanto 20% a descrevem como indicativo de complexidade média e 20%, baixa. Visto que a maior parte dos textos LIC são científicos, era esperado que suas composições fossem distintas de narrativas;
- II. **Simplicidade Sintática (SS):** assim como a média de NT, a de SS também é classificada como indício de alta complexidade. A variação de valores dessa métrica foi entre 1,70 e 79,67%. Como indicativos de complexidade alta, os valores de SS apontaram 45% dos textos, 50% e 5% indícios de complexidade média e baixa, respectivamente. Em adição às indicações de NT, a métrica SS também corrobora alta

complexidade sintática desses textos (característica esperada por serem textos escolhidos devidamente por terem baixa acessibilidade);

- III. **Similaridade Sintática (SIMS):** a SIMS é outra métrica que corrobora a alta complexidade dos textos LIC em termos de sintaxe, posto que 20% dos valores indicam alta similaridade sintática entre todas as frases do texto (ou seja, menor complexidade), enquanto 40%, complexidade média e 40%, alta. Sendo assim, podemos interpretar que 80% dos textos apresentam variações médias e altas nas formações sintáticas, indicando potenciais de complexidade respectivos. A média de SIMS foi de 79, também sendo classificada como indicativa de complexidade alta;
- IV. **Coesão Referencial (CR):** similarmente às demais métricas de sintaxe, a CR também demonstrou alta complexidade a respeito desse nível textual, tendo uma média de 26,68, indicativa de baixa acessibilidade pelas diretrizes dessa métrica. Dentre todos os textos, 65% indicam baixa coesão referencial (consequentemente, maior potencial de complexidade), 15% sugerem complexidade média e 20%, baixa;
- V. **Índice Flesch (IF):** tendo um comportamento claramente oposto ao percebido nos textos LIS, os valores de IF de LIC indicam alta complexidade para a maioria dos textos (75%), média para 15% e baixa para 10%. A média da métrica é de 44,17, sendo classificada como indicativo de alta complexidade. É válido lembrar que o IF é a única métrica desenvolvida especificamente para fins de análise de complexidade, e um resultado de 44,17 está na segunda classificação mais complexa pelas orientações da métrica, conforme explica Goldim (2003);
- VI. **Polissemia (PS):** a análise dos valores de PS indica complexidade alta em 45% dos textos, média em 45% e baixa em 10%. Contudo, a média dos valores é de 3,55, cuja classificação é indicativa de complexidade média. O valor mais baixo de PS é 3,01, enquanto o mais alto, 4,50. Se compararmos com o valor mais alto e mais baixo dos textos LIS (5,07 e 3,49, respectivamente), percebemos que a diferença é relativamente alta (visto que essa métrica geralmente apresenta valores entre 3 e 5);
- VII. **Idade de Aquisição (IA):** conforme esperado, em virtude da natureza dos textos, a IA também indicou alta complexidade de léxico por meio

de uma média de 367,08. Pela interpretação de complexidade dessa métrica, 65% dos textos indicam alta complexidade, enquanto 30%, média e 5%, baixa. Ao lembrarmos que os textos são, primordialmente, científicos e acadêmicos, sabíamos, de antemão, que a IA indicaria alta complexidade (colaborando com o objetivo de coletar textos com baixa acessibilidade para que utilizarmos como referência em análise contrastiva com os valores dos textos do *corpus* de estudo, como apresentaremos posteriormente);

- VIII. Familiaridade com Palavras de Conteúdo (FPC):** seguindo o comportamento das métricas de léxico e semântica anteriores, a FPC também corrobora a alta complexidade dos textos LIC, com 65% deles sendo classificados como complexos, 30% indicativos de complexidade média e 5%, alta. A média de FPC (563,43) também pode ser classificada como índice de alta complexidade;
- IX. Concretude Lexical (CLP):** a média de CLP é de 49,22%, que indica que existe baixa concretude de léxico (e maior potencial de complexidade) em aproximadamente metade dos textos LIC. 55% dos textos indicam alta complexidade pela CLP, 30%, complexidade média e 15%, baixa. Tendo em mente que a CLP é uma métrica percentual, a complexidade nesse quesito é percebida não só pela distribuição dos valores nas escalas interpretativas, mas também por observarmos que um dos textos tem como CLP o valor de 1,19%. Ou seja, nesse texto (LIC1) a concretude lexical é extremamente baixa;
- X. Relação Type-Token (TT):** a indicação de variedade lexical, observada pelo TT, tem uma média de 734, sendo interpretada como índice de complexidade média. Com uma variação entre 836 (maior complexidade) e 635 (menor complexidade), os dados nos permitem concluir que 55% dos textos podem ser considerados complexos pela TT, 25% são classificados como tendo complexidade média e 20%, baixa.

Em síntese, da mesma forma que os resultados dos textos LIS nos permitiram validar suas métricas para formar nossas referências comparativas, os índices dos textos LIC também comprovam que eles têm traços de alta complexidade e, portanto,

são válidos para formar nossa escala contrastiva. Logo, nossa “régua comparativa” tem os valores de LIS como referências aproximadas de acessibilidade, enquanto os resultados de LIC como medidas comparativas de alta complexidade.

Na seção a seguir, apresentaremos os dados coletados sobre nosso *corpus* de estudo e discussões a respeito das interpretações de complexidade lexical, sintática e semântica desses textos para, ao final de cada análise métrica, apresentar uma comparação dos valores dos textos com a escala comparativa proporcionada pelos índices de LIS e LIC.

8.4 DESCRIÇÃO LINGUÍSTICA E PEDAGÓGICA DO *CORPUS* DE ESTUDO

Antes de realizarmos a análise das métricas textuais escolhidas para estimar complexidade, apresentamos, nesta seção, uma descrição global dos textos do *corpus* de estudo. Essa análise será realizada com base nas nossas experiências em sala de aula com esses textos e, também, como pesquisadores linguistas. Dessa forma, a descrição presente na seção não interpretará os valores dos índices textuais, mas nossa experiência e nossos conhecimentos como linguistas pesquisadores da área. Posteriormente, ao analisarmos cada uma das métricas, apresentaremos comentários mais detalhados não só das indicações de cada índice, mas também fundamentados por nossa experiência e posição como linguistas para casos particulares indicados pelos resultados métricos.

O principal objetivo dessa descrição é expor para a leitora⁶² da nossa tese nossas perspectivas sobre os conjuntos de textos de cada livro (*Elementary*, *Pre-intermediate* e *Intermediate*) e/ou apresentá-los para uma leitora que possa, porventura, desconhecê-los. Sendo assim, essa parte da pesquisa abordará os conjuntos de texto de forma globalizada, elencando características comuns aos textos de cada nível e, quando necessário, destacando textos que se sobressaiam. Posteriormente, ao apresentarmos e discutirmos as indicações das métricas textuais, abordaremos com maiores detalhes alguns dos textos e destacaremos trechos que exemplifiquem os resultados dos índices.

Iniciaremos a descrição pelos textos do livro *Elementary*.

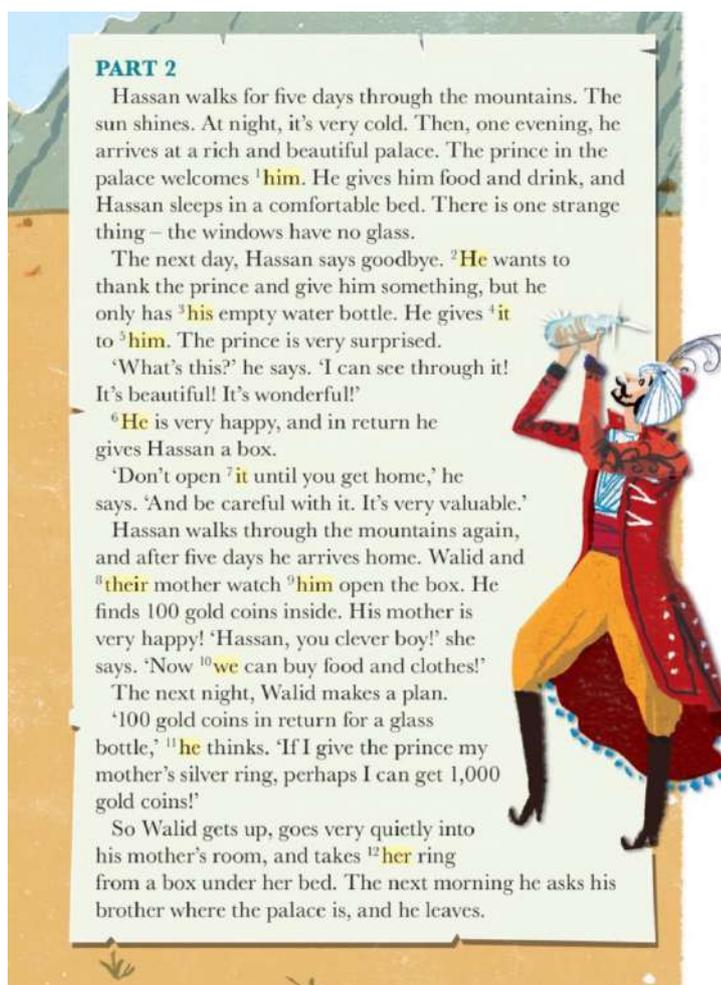
⁶² Embora saibamos que, em língua portuguesa, a generalização é feita no masculino, optamos por intercalar o uso de palavras masculinas e femininas quando estivermos nos referindo a grupos de pessoas de forma geral.

8.4.1 DESCRIÇÃO DOS TEXTOS DO LIVRO ELEMENTARY

Majoritariamente, os textos do livro *Elementary* têm características similares, sendo divididos entre narrativos e narrativos-descritivos.

No que diz respeito à extensão, os textos são relativamente curtos não só na sua totalidade, mas também na formação de frases e parágrafos. Um dos textos mais longos (E6, Figura 11 abaixo) é uma narrativa com a qual os alunos não tendem a ter dificuldade – embora seja mais extensa que os demais textos, a história segue um desenvolvimento lógico e utiliza vocabulário simples – além de ser sustentada por imagens.

Figura 11 - Exemplo de texto do livro *Elementary*



PART 2

Hassan walks for five days through the mountains. The sun shines. At night, it's very cold. Then, one evening, he arrives at a rich and beautiful palace. The prince in the palace welcomes ¹him. He gives him food and drink, and Hassan sleeps in a comfortable bed. There is one strange thing – the windows have no glass.

The next day, Hassan says goodbye. ²He wants to thank the prince and give him something, but he only has ³his empty water bottle. He gives ⁴it to ⁵him. The prince is very surprised.

'What's this?' he says. 'I can see through it! It's beautiful! It's wonderful!'

⁶He is very happy, and in return he gives Hassan a box.

'Don't open ⁷it until you get home,' he says. 'And be careful with it. It's very valuable.'

Hassan walks through the mountains again, and after five days he arrives home. Walid and ⁸their mother watch ⁹him open the box. He finds 100 gold coins inside. His mother is very happy! 'Hassan, you clever boy!' she says. 'Now ¹⁰we can buy food and clothes!'

The next night, Walid makes a plan.

'100 gold coins in return for a glass bottle,' ¹¹he thinks. 'If I give the prince my mother's silver ring, perhaps I can get 1,000 gold coins!'

So Walid gets up, goes very quietly into his mother's room, and takes ¹²her ring from a box under her bed. The next morning he asks his brother where the palace is, and he leaves.

Fonte: Latham-Koenig, Oxenden e Lambert (2018, p. 47)

Na realidade, acreditamos que a extensão dos textos deveria acompanhar a progressão das unidades, tendo textos mais curtos no início do livro e maiores de forma progressiva. Entretanto, o tamanho dos textos varia de forma não sistemática.

Sintaticamente, os textos não parecem seguir uma progressão de complexidade, visto que a partir do terceiro texto do livro *Elementary* já começamos a notar deslocamentos de elementos de suas posições sintáticas canônicas e de orações coordenadas e subordinadas. Uma das características que pode ser entendida como potencialmente complexa é a pontuação, visto que algumas regras são diferentes da língua portuguesa, especialmente em casos de textos com falas de personagens. Além disso, existem algumas inadequações de pontuações em algumas instâncias, especialmente quando adjuntos são deslocados para a frente das frases.

Ainda que textos apresentem formações sintáticas distintas e que deslocamentos de elementos de suas posições canônicas tenham objetivos específicos, acreditamos que para textos didáticos seria interessante observar a progressão e a forma como as alterações em ordem sintática são apresentadas. Em textos do livro *Elementary*, vemos exemplos de adjuntos deslocados para frente das frases, mas também para o meio. Seria interessante, contudo, alterar as posições dos elementos de forma gradativa e que acompanhasse o desenvolvimento lexical (visto que, em determinados momentos, a posição da palavra modifica sua função sintática e, conseqüentemente, a interpretação do texto).

Quanto às formações por coordenação e subordinação, ainda que naturais, também deveriam ser apresentadas de forma gradual e progressiva. Mesmo em textos iniciais, é possível perceber formações sintáticas que tornam frases e parágrafos relativamente longos para alunos em níveis iniciais (especialmente pelo fato de que extensão é uma das principais características de alta complexidade). Na imagem abaixo (Figura 12) podemos observar um exemplo de formação de frases que pode ser complexo para estudantes de ILA em níveis elementares, especialmente considerando que esse texto é parte da unidade quatro do livro *Elementary*:

Figura 12 - Exemplo de texto com sintaxe complexa no livro *Elementary*

What are their secrets?

What do the Greek island of Ikaria, the Nicoya peninsula in Costa Rica, and the island of Okinawa in Japan have in common? The answer is that, together with Loma Linda in California and Sardinia in Italy, they are the five so-called 'Blue Zones', the best places in the world to live if you want to have a long and healthy life. Many people there live until they are 100 or more. So what are their secrets?

Adapted from the British press

Fonte: Latham-Koenig, Oxenden e Lambert (2018, p. 35)

No fragmento ilustrado na imagem acima, percebemos a presença de aposto, de frases coordenadas, de deslocamento de elementos e de subordinação no mesmo parágrafo. Embora para nós, leitores proficientes, essas características não sejam difíceis, elas são potencialmente complexas para estudantes de ILA em níveis iniciais.

No que diz respeito ao léxico, alguns textos apresentam vocabulários relativamente complexos e que poderiam ser substituídos por outras palavras mais simples (ou talvez por reestruturação de frases) sem prejudicar o desenvolvimento de vocabulário dos alunos. Em determinados textos, há espaço para estratégias de simplificação, especialmente porque esse livro é voltado para alunos de níveis iniciais. Destacaremos trechos e palavras complexas na nossa análise das métricas de léxico e semântica, mas pela imagem a seguir (Figura 13), podemos ilustrar a presença de palavras complexas (como '*although*', '*guesthouses*', '*sledges*' etc.) no texto E15.

Figura 13 - Exemplo de texto com léxico complexo do livro *Elementary*

- A** For people who love nature, **Verana** in Puerto Vallarta in Mexico is an amazing spa hotel with beautiful views on all sides. It has ten guesthouses, and four new buildings with balconies above the trees, from where you can see the beautiful Bay of Banderas. The spa has an infinity pool and offers a variety of different spa treatments, as well as yoga classes. You can go whale-watching, fish for tuna, and explore the area on foot. This is the perfect place for your wedding, your honeymoon, or both. But remember, you can only get there by boat!
- B** Would you like to sleep in a mine? Then book the underground suite in **Sala Silvermine** in Sweden, and enjoy the world's deepest bedroom, 155 metres below the surface. Although the corridors are cold and dark, your rooms are warm (18°C), and the light comes from candles in beautiful silver candlesticks. You can also explore the caves and magical lakes with a guide. There are no other guests, so it's not for nervous people. Mobile phones don't work, but you have a radio for emergencies. In the morning they bring breakfast down to you. Perfect for romantic couples who like a bit of adventure.
- C** You don't need to travel to the Arctic Circle to spend time with Siberian huskies. At the **Husky Lodge** in the Swiss canton of Schwyz, dog-lovers can sleep in cabins, heated with wood fires, next to the dogs' kennels. During the day you can join in with their training runs. They pull sledges in winter and bikes and carts in summer. In the evening there's an excellent restaurant. And if you have to work, there are three rooms for small conferences and seminars. If you can't afford the cabins, there's a campsite too – though maybe only in summer!

Fonte: Latham-Koenig, Oxenden e Lambert (2018, p. 85)

Ainda que as palavras que podem ser consideradas complexas no livro *Elementary* não sejam extremamente difíceis (como seria no caso de terminologia especializada), elas ainda apresentam características que reduzem a acessibilidade e a compreensão global dos textos – tais como tamanho, tipo de palavra ou ortografia complexa. Conforme mencionamos na seção 2.4 desta tese ('Características Potencializadoras de Complexidade'), existem tipos de palavras que são consideradas potencialmente complexas, tais como advérbios, adjetivos e palavras extensas para falantes nativos, como indica o manual de acessibilidade do projeto *Plain Language* do governo dos Estados Unidos (mencionado anteriormente) e autores como DuBay (2004) e Silva (2018). Considerando que algumas classes de palavras são avaliadas como complexas para falantes nativos, é altamente provável que sejam tão (ou mais) complexas para estudantes de ILA.

Uma das principais características que podem ser destacadas como complexas (com base na nossa experiência em sala de aula) é a presença de palavras

que são parte do vocabulário estudado em unidades muito distantes de onde ocorrem - por exemplo, o alto uso de advérbios de modo no texto da unidade 3A, sendo que essa classe de palavras é introduzida apenas na unidade 11A (especialmente porque o livro é usualmente dividido em duas partes e as unidades anteriores à sétima são estudadas em um semestre, enquanto as posteriores, em outro semestre). Levando em consideração as observações sobre a complexidade de advérbios e a distância entre o texto da unidade 3A e o estudo de advérbios na 11A, seria possível adaptar o material para reduzir a presença advérbios. Embora saibamos que a presença dessas palavras não impede, totalmente, a interpretação dos textos, sua ausência não prejudicaria a manutenção linguística e, ao mesmo tempo, contribuiria para maior acessibilidade textual.

Em raros textos, glossários são apresentados e, quando o são, poucas são as palavras presentes. Uma vez que existem palavras que não podem ser substituídas por sinônimos mais simples em determinados momentos, acreditamos que uma frequência maior de glossários seja válida. Para alunos brasileiros, especialmente, acreditamos que o destaque para falsos cognatos seja interessante em alguns glossários (como para as palavras '*major*' no texto E8, '*novels*' no E12, '*fortune*' no E14 etc.). Além de facilitar a interpretação, é preciso lembrar que os textos podem ser atribuídos aos alunos como tarefas de casa e, dessa forma, se não houver a intervenção do professor, os falsos cognatos podem prejudicar a interpretação dos textos.

Por fim, outra observação importante a respeito dos textos do livro *Elementary* (e que também é válida para os demais livros) é que eles tendem a apresentar uma história relacionada à temática da unidade, mas utilizam poucas estruturas gramaticais e vocabulários que sejam o foco de cada capítulo. Acreditamos que os textos devem complementar os conteúdos lexicais e gramaticais das unidades e que devem ser utilizados como espaço de aplicação dos pontos que são trabalhados em cada unidade. Embora algumas palavras e exemplos de estruturas que serão estudados sejam assistematicamente apresentados nos textos, ainda há espaço e alternativas para oportunizar mais insumos nesses materiais.

Na próxima seção, descreveremos os textos do livro *Pre-intermediate* sob nossa ótica como docentes de ILA e como pesquisadores linguistas.

8.4.2 DESCRIÇÃO DOS TEXTOS DO LIVRO PRE-INTERMEDIATE

De forma similar aos do *Elementary*, a maior parte dos textos do livro *Pre-intermediate* são narrativos-descritivos, embora os descritivos sejam mais frequentes nesse nível do que no anterior.

A extensão dos textos varia com frequência e sem nenhuma progressão constante – ou seja, o tamanho dos textos não segue a sequência das unidades. Na unidade 11C, por exemplo (que é uma das unidades finais do livro), o texto principal é relativamente curto, conforme podemos observar na imagem abaixo (Figura 14):

Figura 14 - Exemplo de texto curto do livro *Pre-intermediate*

Adapted from the British press

Two women recently met by chance at Bremen University, in Germany. One was English, and one was Irish, and they were both on Erasmus scholarships. Nothing unusual there, except for one thing. The two girls look identical. Their hair is the same colour and length, they're the same age and size, and when you see them together, in the photo they put on social media, you would think that they were identical twins. In fact, Cordelia Roberts and Ciara Murphy are unrelated.

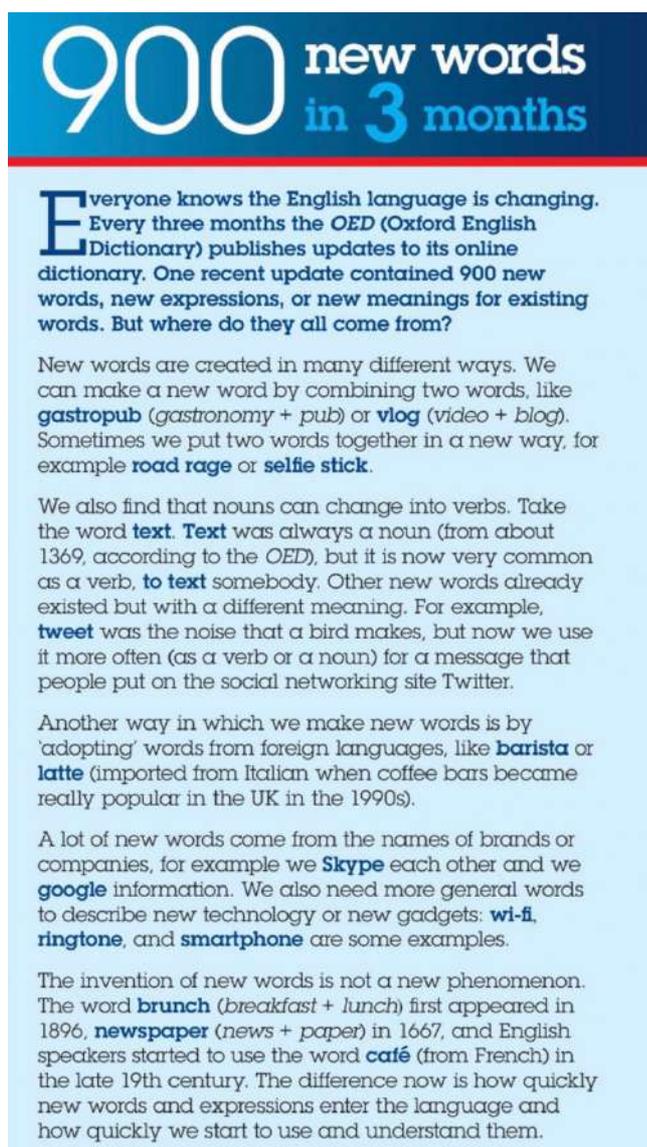
It seems that it is not uncommon for people who are unrelated to look almost identical. Niamh Geaney, from Dublin, and two friends were so interested in trying to find their 'twins' that they set up an online project called Twin Strangers. Very quickly, Niamh found a remarkably similar-looking stranger who lived just a few miles away. It's perhaps not so surprising, as both young women look typically Irish, with dark hair and very pale skin, but then Niamh found another lookalike – Luisa Guizzardi, who is from Genoa in Italy!

Glossary
Erasmus scholarships
 a programme which allows students from the European Union to study in another country

Fonte: Latham-Koenig, Oxenden e Lambert (2019, p. 90)

Já em alguns capítulos iniciais, são apresentados textos mais longos. Um exemplo está na parte de revisão após a unidade 4C, como ilustra a imagem a seguir (Figura 15):

Figura 15 - Exemplo de texto mais extenso do livro *Pre-intermediate*



900 new words in 3 months

Everyone knows the English language is changing. Every three months the *OED* (Oxford English Dictionary) publishes updates to its online dictionary. One recent update contained 900 new words, new expressions, or new meanings for existing words. But where do they all come from?

New words are created in many different ways. We can make a new word by combining two words, like **gastropub** (*gastronomy* + *pub*) or **vlog** (*video* + *blog*). Sometimes we put two words together in a new way, for example **road rage** or **selfie stick**.

We also find that nouns can change into verbs. Take the word **text**. **Text** was always a noun (from about 1369, according to the *OED*), but it is now very common as a verb, **to text** somebody. Other new words already existed but with a different meaning. For example, **tweet** was the noise that a bird makes, but now we use it more often (as a verb or a noun) for a message that people put on the social networking site Twitter.

Another way in which we make new words is by 'adopting' words from foreign languages, like **barista** or **latte** (imported from Italian when coffee bars became really popular in the UK in the 1990s).

A lot of new words come from the names of brands or companies, for example we **Skype** each other and we **google** information. We also need more general words to describe new technology or new gadgets: **wi-fi**, **ringtone**, and **smartphone** are some examples.

The invention of new words is not a new phenomenon. The word **brunch** (*breakfast* + *lunch*) first appeared in 1896, **newspaper** (*news* + *paper*) in 1667, and English speakers started to use the word **café** (from French) in the late 19th century. The difference now is how quickly new words and expressions enter the language and how quickly we start to use and understand them.

Fonte: Latham-Koenig, Oxenden e Lambert (2019, p. 37)

Ainda que não exista um padrão e uma progressão na extensão dos textos, nossa experiência em sala de aula aponta que, na verdade, tamanho não tende a ser um potencial de complexidade. Na realidade, acreditamos que, para otimizar a manutenção linguística, seria possível apresentar textos um pouco mais extensos em determinadas unidades e que fossem mais ricos em insumos relacionados aos tópicos gramaticais e lexicais das unidades (mesma crítica apresentada para o livro *Elementary*).

No que diz respeito ao léxico, há maiores frequências de palavras mais extensas e um pouco mais complexas em relação ao *Elementary* (mas não tão mais complexas que os textos do livro anterior). Textos com léxicos mais densos incluem,

com frequência, termos e palavras relacionadas a áreas específicas - como no texto P7 (*'barista', 'gastropub', 'vlog'*) ilustrado pela imagem acima (Figura 25) e no P11 (*'mentors', 'preachers', 'soundtrack'*), como demonstra a Figura 26 abaixo:

Figura 16 - Exemplo de texto com léxico complexo do livro *Pre-intermediate*

The woman who inspired *'I have a dream'*

Mahalia Jackson was a musical legend who helped to bring gospel music from the church to large audiences. She was Aretha Franklin's mentor, and in 1961 she was the first gospel singer to win a Grammy Award. She was also an important member of the Civil Rights Movement in the USA, and she was a close friend of Martin Luther King.

Mahalia often went with King on civil rights demonstrations and marches, including into the most hostile parts of the South, and she sang at the events. As a presenter from National Public Radio said, 'her voice became the soundtrack of the Civil Rights Movement'.

On 28th August 1963, there was a famous march in Washington against racism. Mahalia was with King, who was going to give a five-minute speech. Before the day itself, he and his advisers decided what he was going to say.

King began speaking to an audience of more than 250,000 people. But towards the end, he felt that his speech was not going well. Suddenly Mahalia shouted "Tell them about the dream, Martin. Tell them about the dream." She knew about it because she was at a previous event in Detroit, in June of that year, where King talked about his dream for African Americans. It was, in the words of King's adviser Clarence Jones, 'one of the world's greatest gospel singers shouting to one of the world's greatest preachers'. King looked at Mahalia. Then he threw away his written speech, and looked at the audience.

'I have a dream...I have a dream that my four little children will one day live in a nation where they will not be judged by the colour of their skin, but by the content of their character...'

And so he continued, and he gave one of the best-loved speeches in American history.

Glossary

Civil Rights Movement the campaign in the 1950s and 1960s to change the laws so that African Americans have the same rights as others

preacher a person who gives inspiring talks about religion

Adapted from a website

Fonte: Latham-Koenig, Oxenden e Lambert (2019, p. 53)

Por outro lado, alguns textos do livro *Pre-intermediate* possuem vocabulários relativamente simples ou, então, um balanço entre unidades lexicais complexas e outras mais conhecidas (característica que seria ideal para o nível), como podemos observar pelo fragmento do texto P4 a seguir (Figura 27):

Figura 17 - Exemplo de texto com léxico balanceado do livro *Pre-intermediate*



Fonte: Latham-Koenig, Oxenden e Lambert (2019, p. 21)

Já sobre a formação sintática, não é possível observar nenhum padrão de formação. Parâmetros de formação de estrutura sintática não parecem ter sido seguidos e os textos são compostos de forma livre (ou seja, sem busca de padronizar ou apresentar diferentes formações sintáticas de forma progressiva).

Ainda que não sejam considerados extremamente complexos em termos de sintaxe, os textos são redigidos com muitos apostos, deslocamentos de elementos, diferentes pontuações e, ao mesmo tempo, exigem do leitor maiores capacidade de relacionar elementos (como pronomes e seus referenciais) visto que, em alguns momentos, um pronome se refere a algo que é mencionado no título ou em alguma outra parte do texto, ou um sinônimo não tão conhecido é utilizado para retomar algo mencionado anteriormente. No fragmento do texto P12 ilustrado abaixo (Figura 28), por exemplo, o pronome *'that'* do primeiro parágrafo se refere à pergunta feita no título:

Figura 18 - Exemplo de texto que exige conhecimentos de coesão referencial do livro *Pre-intermediate*



Fonte: Latham-Koenig, Oxenden e Lambert (2019, p. 21)

Embora não seja tão complexo de compreender que o *'that'* se refere ao que foi perguntado no título do texto, com frequência em sala de aula, temos observado que essa característica é complexa para alguns estudantes. Obviamente, o fragmento apresentado acima é apenas para fins de ilustração; existem outros textos que “exigem” maiores habilidades de coesão referencial dos alunos – traço que pode ser potencialmente complexo.

Na nossa perspectiva como professores, acreditamos que o livro *Pre-intermediate* é – entre todos os demais da coleção – o menos coerente em manutenção de complexidade. Em determinados pontos do livro, os textos tendem a ser altamente acessíveis, mas em outros, complexos. Evidentemente, tal característica seria positiva considerando o nível para o qual os textos são propostos (mas somente se a progressão fosse coerente e crescente; ou seja, textos mais simples no início do livro, quando as habilidades dos alunos estão mais próximas do nível elementar e textos mais complexos do meio para o final – quando os estudantes têm habilidades mais próximas ao nível intermediário.).

Na próxima subseção, apresentamos uma breve descrição dos textos do livro *Intermediate*.

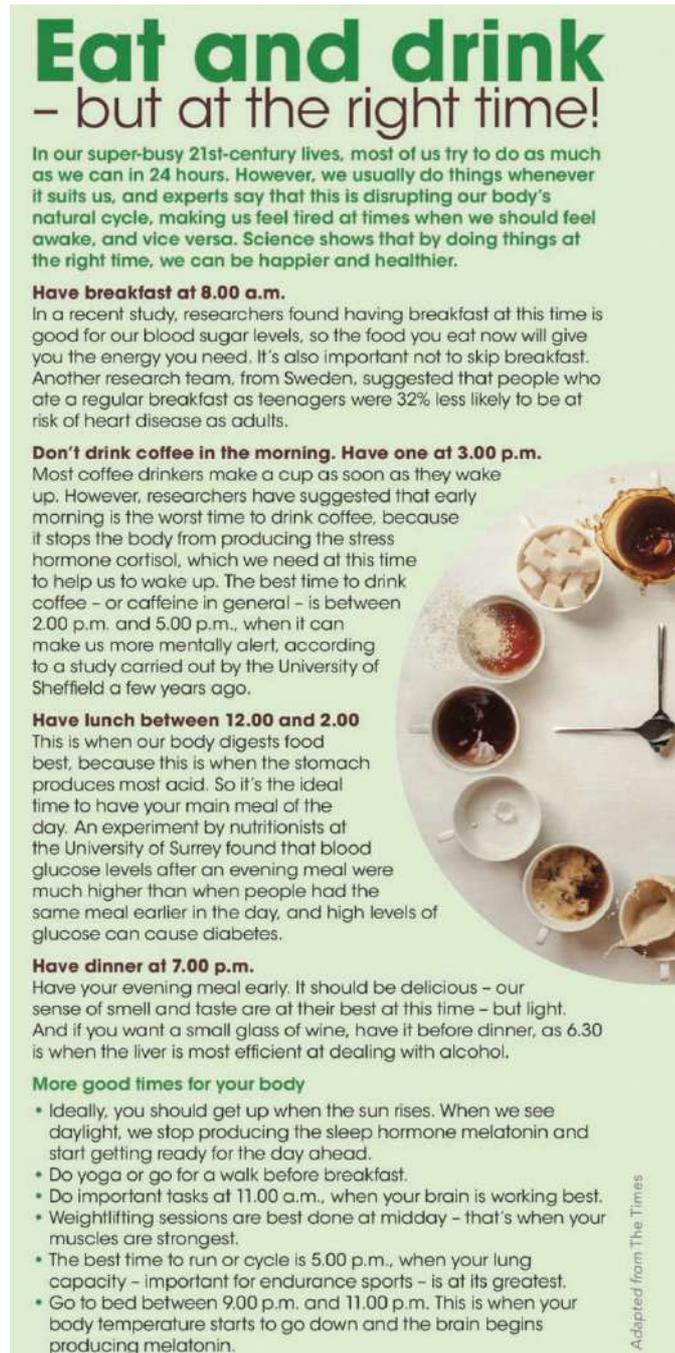
8.4.3 DESCRIÇÃO DOS TEXTOS DO LIVRO INTERMEDIATE

Embora, conforme mencionamos antes, não tenhamos como objetivo conduzir uma análise completa dos textos do livro *Intermediate*, consideramos importante apresentar uma breve descrição desses materiais. Mesmo que uma análise minuciosa não seja apresentada, os valores métricos dos textos e um estudo mais sucinto serão. Sendo assim, a descrição oferecida nessa seção poderá auxiliar a leitora a realizar uma interpretação mais detalhada das métricas, caso ela assim desejar.

De forma um pouco diferente dos livros anteriores, o *Intermediate* apresenta textos muito mais próximos do gênero descritivo do que do narrativo (embora ainda apresente alguns narrativos). Além disso, as temáticas são mais densas e especializadas e, em determinadas unidades, com linguagem e estrutura similares a outros gêneros textuais (como textos jornalísticos, acadêmicos etc.), ainda que adaptados.

Em termos de extensão, os textos do livro *Intermediate* aumentam razoavelmente (ao compararmos com os dois antecessores) desde a primeira unidade - embora, em determinados momentos, sejam mais curtos. Na imagem abaixo, podemos ter uma ideia do tamanho dos textos que, em geral, regulam com o I1 (Figura 19), texto presente na primeira unidade:

Figura 19 - Exemplo de extensão de texto do livro *Intermediate*



Eat and drink
- but at the right time!

In our super-busy 21st-century lives, most of us try to do as much as we can in 24 hours. However, we usually do things whenever it suits us, and experts say that this is disrupting our body's natural cycle, making us feel tired at times when we should feel awake, and vice versa. Science shows that by doing things at the right time, we can be happier and healthier.

Have breakfast at 8.00 a.m.
In a recent study, researchers found having breakfast at this time is good for our blood sugar levels, so the food you eat now will give you the energy you need. It's also important not to skip breakfast. Another research team, from Sweden, suggested that people who ate a regular breakfast as teenagers were 32% less likely to be at risk of heart disease as adults.

Don't drink coffee in the morning. Have one at 3.00 p.m.
Most coffee drinkers make a cup as soon as they wake up. However, researchers have suggested that early morning is the worst time to drink coffee, because it stops the body from producing the stress hormone cortisol, which we need at this time to help us to wake up. The best time to drink coffee - or caffeine in general - is between 2.00 p.m. and 5.00 p.m., when it can make us more mentally alert, according to a study carried out by the University of Sheffield a few years ago.

Have lunch between 12.00 and 2.00
This is when our body digests food best, because this is when the stomach produces most acid. So it's the ideal time to have your main meal of the day. An experiment by nutritionists at the University of Surrey found that blood glucose levels after an evening meal were much higher than when people had the same meal earlier in the day, and high levels of glucose can cause diabetes.

Have dinner at 7.00 p.m.
Have your evening meal early. It should be delicious - our sense of smell and taste are at their best at this time - but light. And if you want a small glass of wine, have it before dinner, as 6.30 is when the liver is most efficient at dealing with alcohol.

More good times for your body

- Ideally, you should get up when the sun rises. When we see daylight, we stop producing the sleep hormone melatonin and start getting ready for the day ahead.
- Do yoga or go for a walk before breakfast.
- Do important tasks at 11.00 a.m., when your brain is working best.
- Weightlifting sessions are best done at midday - that's when your muscles are strongest.
- The best time to run or cycle is 5.00 p.m., when your lung capacity - important for endurance sports - is at its greatest.
- Go to bed between 9.00 p.m. and 11.00 p.m. This is when your body temperature starts to go down and the brain begins producing melatonin.

Adapted from The Times

Fonte: Latham-Koenig, Oxenden e Lambert (2019, p. 7)

Ainda que mais longos que os textos dos livros anteriores, a extensão ainda é coerente para finalidades em sala de aula, permitindo que os alunos tenham tempo suficiente para leitura e releitura (caso precisem), visto que são utilizados em períodos de cinquenta minutos (e geralmente duplos, configurando uma hora e trinta minutos).

Outra característica comum entre os textos do *Intermediate* é que muitos deles são divididos por tópicos com subtítulos, ou com tópicos frasais – configuração que potencializa a acessibilidade, conforme explica Silva (2018). Além disso, muitos textos desse nível utilizam recursos como *bullet points* ou outros marcadores visuais que – também como aponta Silva (2018) – tornam a leitura mais fácil, como podemos perceber pelo exemplo apresentado acima (texto I1, Figura 19) e, também, pelo texto I13 abaixo (Figura 20):

Figura 20 - Exemplo de texto com marcadores visuais do livro *Intermediate*

What your profile photo says about you

Choosing a profile photo is a serious business. It will be the first thing old school friends, jealous exes, and even potential bosses see when they search for you online. Your image depends entirely on your choice of photo. Here are the most common types of profile photo on social media. Is yours one of them? And if so, what does it say about you?

A  **The portrait**
A clear, close-up photo.
What it says about you You are quite boring. If it's a selfie, you are quite annoying.

B  **The childhood photo**
A cute picture of you as a baby.
What it says about you You are the type of person who thinks that everything used to be better than it is now. You still listen to the same music, wear the same clothes, and love the same things you did at school, and you'll probably never change.

C  **The pet**
Your pet looking adorable.
What it says about you It depends on what kind of animal it is. Cat: You are a woman without a boyfriend. Dog: You are a man without a girlfriend. Snake: You are a teenage boy or death metal fan.

D  **The wedding photo**
Man, woman, dress, suit – you know, the usual.
What it says about you You want everyone to think that you are a grown-up. You don't go out and have a good time any more. No, you are married! Also, you don't feel you exist as an individual any more, and don't have any friends of your own.

E  **The family photo**
A photo of your children / baby.
What it says about you The main thing you have accomplished in your adult life is having children. You used to be fun and fabulous and have a lot of friends, but now all you talk about is nappies and children's TV.

F  **The popular culture reference**
A picture of a cartoon character, a movie poster, a book cover, a musical act, a celebrity, etc.
What it says about you You have no personality of your own. Your identity depends on your entertainment choices – television, music, sci-fi, literary, or other. You own at least two T-shirts with stupid slogans on them.

G  **The party photo**
You, often with other people, enjoying yourself at a party.
What it says about you You are young and stupid, and will be fired from at least one job for something you posted on Facebook. One day, you might regret this picture and replace it with a wedding picture, and then photos of your children.

H  **No photo at all**
An icon, not a photo.
What it says about you You are technologically incompetent and don't know what a jpeg is; or you think you're too busy to find a photo; or you think not having a photo is 'cool'. Having no profile photo is annoying for everybody else. Get a photo.

Fonte: Latham-Koenig, Oxenden e Lambert (2019, p. 60)

Quanto ao vocabulário, os textos apresentam maiores variedades lexicais (quando comparado aos antecessores) e maior densidade semântica, como no texto I9 com palavras muito específicas de áreas como esportes e Química que tendem ser complexas para os alunos, como podemos observar pelo fragmento desse texto abaixo (Figura 21):

Figura 21 - Exemplo de texto com léxico complexo do livro *Intermediate*

1 **S**port has always been full of superstitions, even for the best athletes in the world. Long-distance runner Mo Farah always shaves his head before a race. Judo star Kayla Harrison always wears the lucky socks that were a gift from her grandmother. Footballer Cristiano Ronaldo always steps onto the pitch with his right foot first. And tennis legend Serena Williams always bounces the ball five times before a first serve, and twice before a second serve.

2 **Do actions like this have any real effect on sports results?** According to a study at the University of Cologne, it seems that they can increase confidence and a sense of control, which might help athletes. The study showed that people who were given a golf ball that they were told was 'lucky' played better than those who used a 'normal' ball. And it isn't just sports. In another experiment, people who were allowed to carry their lucky charm performed better at memory tests than people without one. The researchers concluded that superstitions make people feel more confident, which makes them perform better.

Fonte: Latham-Koenig, Oxenden e Lambert (2019, p. 47)

Já em termos de sintaxe, os mesmos comentários que apresentamos sobre os textos do livro *Pre-intermediate* também são válidos para o *Intermediate*: não parece haver nenhuma coerência em progressão de complexidade e os textos são ricos em variadas formações sintáticas, deslocamentos, coordenações, subordinações, elementos de referenciação etc. Contudo, é válido lembrar que uma vez que esses materiais são voltados para alunos que são, supostamente, de um nível intermediário, a complexidade sintática pode ser um pouco maior do que nos livros anteriores.

De forma geral, os textos do livro *Intermediate* apresentam traços que podem ser considerados complexos mesmo para o nível para o qual ele é proposto. Embora não tenhamos a pretensão de analisar de forma detalhada os indicativos métricos desses textos, destacaremos, posteriormente, alguns pontos que merecem a atenção caso um professor deseje tornar esses textos um pouco mais acessíveis.

Em suma, acreditamos que nossas descrições dos livros podem ser utilizadas como forma de complementação às informações obtidas pelas análises das métricas (que apresentaremos na seção a seguir). Essas duas formas de estudo (semiautomática e holística) de acessibilidade/complexidade podem auxiliar pesquisadores que desejam analisar outros materiais, orientar professoras a respeito de práticas de análise de CT e AT e prover informações para docentes que trabalham com os livros da coleção *English File* e queiram adaptar seus textos para fins de acessibilidade (seja para aumentá-la ou diminuí-la).

8.5 ANÁLISE SEMIAUTOMÁTICA DO CORPUS DE ESTUDO

Para que a interpretação das métricas seja a mais detalhada possível, nesta seção faremos para o conjunto de textos de cada nível linguístico (*Elementary*, *Pre-intermediate* e *Intermediate*) as seguintes análises:

a. Da complexidade lexical-semântica, utilizando as métricas PS, IA, FPC, CLP e TT.

b. Da complexidade sintática, com base nos índices NT, SS, SIMS, CR e IF.

Nossa discussão sobre os indícios de complexidade será baseada nas diretrizes de cada métrica elegida, na organização dos valores em relação à complexidade estabelecidos por Silva (2018) e por meio dos nossos *corpora* progressivo e comparativo e num cruzamento de dados.

Iniciaremos nossa análise pelos textos do livro do *Elementary*.

8.5.1 RESULTADOS MÉTRICOS DOS TEXTOS DO LIVRO DE NÍVEL ELEMENTARY

O quadro abaixo (Quadro 27) exhibe os valores das métricas de léxico e semântica dos textos do livro *Elementary* coletados por meio do Coh-Metrix:

Quadro 27 - Valores das métricas de léxico e semântica dos textos do livro *Elementary*

	SEMÂNTICA E LÉXICO				
	PS	IA	FPC	CLP	TT
E1	4,22	311,18	579,96	95,82	578
E2	3,35	275,72	595,44	59,1	478
E3	4,03	303,48	593,69	45,22	490

E4	4,13	270,03	597,04	85,31	504
E5	3,67	298,08	586,95	90,15	530
E6	4,17	231,03	593,85	98,26	402
E7	3,75	281,18	583,49	77,34	513
E8	3,47	308,52	582,79	36,32	489
E9	4,05	269	591,14	50	430
E10	3,79	236,42	602,04	63,68	482
E11	3,64	268,6	580,16	89,25	543
E12	3,32	385,51	567,85	70,19	582
E13	3,5	266,12	580,33	76,11	567
E14	4,11	291,44	585,15	42,86	446
E15	3,34	314,75	578,53	92,92	596
E16	3,95	330,46	584,48	60,26	499
E17	4,3	280,5	582,34	72,57	502
E18	3,6	287,59	587,05	55,17	536
E19	4,08	284,15	588,4	80,78	462
E20	3,97	332,35	593,16	9,51	517
MÉDIA	3,82	291,31	586,69	67,54	507

Fonte: Elaboração própria

Legenda: PS: Polissemia | IA: Idade de aquisição | FPC: Familiaridade com palavras de conteúdo | CLP: Concretude Lexical | TT: *Type-Token*

Antes de interpretarmos os dados, é importante retomar que esses valores métricos, por si, não interpretam complexidade; é necessário que um analista de complexidade (geralmente um linguista com experiência com esses índices) realize a leitura e a interpretação indicativa de complexidade por meio desses números.

A análise será realizada da seguinte maneira: I. Destacando e comentando o texto cujos resultados de cada métrica indicam maior acessibilidade e o texto cujos valores indicam maior complexidade e; II. Apresentando exemplos relativos a cada índice que influenciaram positiva ou negativamente nos resultados.

Nossa discussão dos valores das métricas pode servir como orientações para análises detalhadas de todos os demais textos. Ou seja, apresentaremos exemplos de como os valores podem ser interpretados e quais são as ocorrências e os fenômenos relacionados a cada métrica de forma que professores possam seguir passos semelhantes para analisar não só os textos para os quais não apresentaremos detalhes, mas também outros materiais escritos em língua inglesa para estudantes de ILA. Cabe também lembrar que nos capítulos 9 ('Sugestões de Estratégias de Acessibilidade para os Textos do *Corpus* de Estudo') e 10 ('Orientações para

Acessibilidade Textual Linguodidática') desta tese, apresentamos sugestões pontuais de estratégias de acessibilidade para cada um dos textos destacados como potencialmente complexos pelas métricas e orientações de como utilizar cada índice para estimar complexidade e, a partir dos resultados, escolher estratégias de simplificação.

Iniciaremos nossa análise pela interpretação dos valores com relação à complexidade lexical-semântica, conforme a seção a seguir.

8.5.1.1 Análise da complexidade lexical-semântica dos textos E

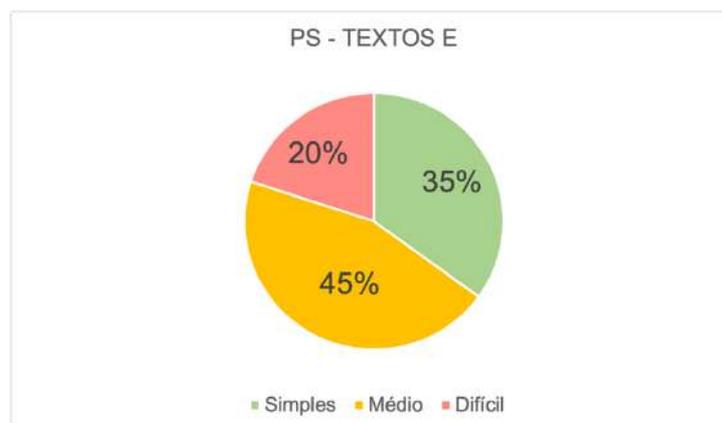
Nossa primeira análise diz respeito ao vocabulário presente nos textos dos livros didáticos. Como previamente mencionado e defendido por pesquisadores de Complexidade Textual (DUBAY, 2004; SILVA, 2018; FINATTO, 2018, entre outros), o léxico tende a ser o maior fator de complexidade em textos.

Se as pesquisas dos autores supracitados apontam que o léxico é um dos maiores complicadores de interpretação de texto tendo como base textos escritos na língua primária dos leitores, é evidente que esse fator também será responsável por aumentar ou diminuir a acessibilidade para leitores ainda não proficientes de uma língua adicional.

Iniciaremos nossa análise das métricas de léxico e semântica pelo índice PS, conforme a próxima subseção.

8.5.1.1.1 Análise da métrica PS dos textos E

O gráfico a seguir (Gráfico 11) apresenta o percentual dos textos classificados como simples, médio e difícil pelas interpretações de PS:

Gráfico 11 - Classificação percentual de complexidade de PS do livro *Elementary*

Fonte: Elaboração própria

Como podemos constatar, 35% dos textos podem ser classificados como simples pela métrica PS, enquanto 45% e 20% como médios e difíceis, respectivamente. Ainda que os números nos permitam classificar os potenciais de complexidade dessa forma, destacamos que para o livro *Elementary*, a polissemia deve ser observada com minuciosidade. Uma vez que polissemia está relacionada à compreensão de múltiplos possíveis significados de uma palavra, é necessário que, então, os leitores conheçam essas alternativas. Geralmente, um leitor em estágios iniciais de aprendizado de ILA conhece apenas os significados mais “comuns” das unidades lexicais (e não os demais). Portanto, ainda que o índice PS nos alerte sobre possíveis palavras que podem ser indicativos de complexidade, a observação de um linguista é essencial. Palavras como ‘*can*’ (poder e lata) e ‘*light*’ (luz, leve e claro), por exemplo, são ilustrações de polissemia cujos múltiplos significados podem ser conhecidos por um estudante em níveis iniciais. Por outro lado, palavras como ‘*wing*’ e ‘*bank*’, por exemplo, na maioria das vezes evocam apenas os significados mais comuns (no caso dos exemplos, ‘asa’ e ‘banco’, respectivamente; uma vez que outros significados desses exemplos – ‘ala’ e ‘margem’ - são vocabulários mais avançados). Assim, para a análise da métrica PS, destacaremos apenas as palavras que consideramos que poderiam evocar mais de um significado para um aluno em nível inicial.

Ao observarmos o quadro que exhibe as métricas (Quadro 27 acima), percebemos que o texto com melhor valor de PS (indicando, então, maior acessibilidade) é o texto E17 (PS: 4,3). Já o texto com valor mais baixo dessa métrica (indício de maior complexidade) é o E12 (PS: 3,32). No quadro abaixo (Quadro 28),

apresentamos as palavras polissêmicas de cada um desses textos que consideramos como possíveis de evocação de mais de um significado para o nosso leitor:

Quadro 28 - Textos do livro *Elementary* com exemplos de polissemia

TEXTO	EXEMPLOS DE POLISSEMIA
E17 (PS: 4,3)	<i>about, assistant, back, book, can, company, cook, dream, end, experiences, flat, left, like, play, read, rest</i>
E12 (PS: 3,32)	<i>bit, book, can, characters, novel, past, present, read, type</i>

Fonte: Elaboração própria

Embora o contexto possa auxiliar na ativação do significado correto das palavras polissêmicas e que nós – leitores proficientes – possamos pensar “não tem como confundir os significados possíveis dessas palavras”, não podemos, como já mencionado inúmeras vezes, supor que o contexto (seja pelo léxico antecedente e posterior ou pela ideia do texto) ou que pistas contextuais serão compreendidas pelo leitor. Sendo assim, os destaques de palavras polissêmicas servem como um alerta de possível complexidade para que professores possam atentar para a compreensão dessas palavras e/ou eliciar possíveis compreensões dos alunos.

No texto E17, por exemplo, que é sobre viagem e sobre atividades feitas pela personagem, a palavra ‘*book*’ pode evocar tanto o significado de ‘reservar’ como o de ‘livro’, assim como a palavra ‘*cook*’ pode direcionar ao significado de ‘cozinhar’ ou de ‘cozinheira’ e ‘*left*’ de ‘esquerda’ ou ‘deixar’.

Similarmente, algumas palavras do texto E1 (que foi classificado como tendo o segundo melhor resultado de PS) podem ter o mesmo efeito de evocação de múltiplos significados (exemplos nesse texto: ‘*building*’, ‘*can*’, ‘*great*’, ‘*park*’, ‘*plan*’, ‘*state*’, ‘*subway*’, ‘*tips*’). Entretanto, além desse material ter sido indicado como tendo o segundo melhor resultado de PS, ele é apresentado na primeira unidade do livro. Uma vez que – em teoria – essa seria a primeira leitura dos alunos de um texto em inglês, é possível que seus vocabulários sejam altamente limitados e, conseqüentemente, as unidades terem pouca probabilidade de impedir a compreensão devido à polissemia.

No texto considerado como mais complexo pela PS (E12), algumas palavras podem, de fato, direcionar para um significado que não seja pretendido. Uma vez que o texto está falando de histórias e publicações de detetives, palavras como ‘*characters*’, ‘*present*’ e ‘*type*’ podem – ainda que pouco provavelmente – serem

interpretadas como ‘caracteres’ (número de palavras), ‘apresentar’ e ‘digitar’ em vez dos significados reais: ‘personagens’, ‘presente’ e ‘tipo’.

Um exemplo de frase com uma palavra polissêmica que pode causar erro na interpretação no E12 é a primeira do texto: “*Millions of readers, like me, love the books of Gillian Flynn and Ruth Rendell...*”. Nesse caso, uma vez que os leitores desse texto são alunos em níveis iniciais e podem desconhecer as regras de pontuação (fato que também é comum mesmo na língua portuguesa), a palavra ‘*like*’ pode ser interpretada pelo seu sentido mais conhecido, ‘gostar’. Nessa ocorrência, o aluno poderia compreender a frase como “Muitos leitores gostam de mim [e] amam os livros de Gillian Flynn e Ruth Rendell...” em vez de “Muitos leitores, como eu, amam os livros de Gillian Flynn e Ruth Rendell...”.

Outro texto com baixos valores de PS, o E15, também apresenta palavras que possibilitam interpretações distintas das pretendidas. Visto que o texto é sobre hotéis incomuns e sobre atividades que podem ser feitas nos locais, palavras como ‘*book*’, ‘*fish*’, ‘*mine*’ e ‘*well*’ podem ser entendidas como – respectivamente – ‘livro’, ‘peixe’, ‘meu’ e ‘poço’ quando, na verdade, significam ‘reservar’, ‘pescar’, ‘mina’ e ‘bem’.

Conforme comentamos sobre a posição do texto E1 e o fato de o primeiro texto do livro ter sido classificado como segundo melhor resultado de PS, também é válido observar que o segundo texto desse nível (E2) apresentou valor dessa métrica, indicando alto potencial de complexidade. Entretanto, ao analisarmos o texto de forma holística, percebemos que as palavras polissêmicas são exemplos de unidades que raramente direcionariam para o significado não-pretendido pelo texto, já que essa ativação de outro significado exigiria domínio mais avançado de vocabulário. Exemplos de palavras com potencial polissêmico nesse texto são: ‘*minute*’, ‘*bank*’, ‘*shops*’.

Além dos textos já mencionados com valores de PS indicativos de complexidade, o texto E8 também foi destacado como potencialmente complexo. Entretanto, assim como os demais textos do *Elementary*, podemos observar que a polissemia dificilmente impedirá a compreensão do texto pelos motivos supracitados.

Por meio da interpretação e verificação dos valores de PS e seus indícios observamos que a métrica, por si só, tem pouco potencial para estimar complexidade a nível lexical e semântico. Ela serve, evidentemente, como uma forma de alerta para observarmos possíveis ocorrências de palavras que podem ser compreendidas de forma diferente à intencionada, mas – conforme citamos anteriormente – a

possibilidade de ativação de múltiplos significados está diretamente relacionada ao maior domínio de vocabulário que, por sua vez, usualmente é indício de maior proficiência linguística. Ao mesmo tempo, um aluno que tenha maior proficiência linguística irá, geralmente, compreender o significado esperado das palavras polissêmicas pelo contexto comunicativo na qual estão inseridas.

Ainda assim, é importante destacar que (conforme já mencionamos) em processos de análise de acessibilidade textual, nunca devemos ignorar possíveis indicadores de complexidade por acharmos que, para nós, essas características não interfeririam na interpretação de um texto. Silva (2021) define o conceito que nomeia como “Princípio da Alcançabilidade” que, resumidamente, é a ideia que deve permear análises de complexidade e adoções de estratégias de simplificação e que orienta que para um público de leitores heterogêneos, devemos considerar todos os possíveis cenários e indícios de complexidade e, caso só possamos apresentar uma versão de um texto com o objetivo de ser acessível, esse texto deve ser redigido ou editado tendo em mente um leitor com a menor proficiência e com mais dificuldades de leitura. Sendo assim, se um texto for acessível para esse leitor, também será para outros com maiores proficiências e menores dificuldades.

Evidentemente, ao adotarmos estratégias de simplificação para fins didáticos, temos espaço para apresentar uma versão de texto para cada aluno (caso necessário) com estratégias que se adequem às suas necessidades. Logo, a parte do princípio da alcançabilidade proposto por Silva (2021) que devemos manter em mente é a consideração de todas as possibilidades possíveis de traços de complexidade, sem supor que se for simples para nós, será simples para todo tipo de leitor.

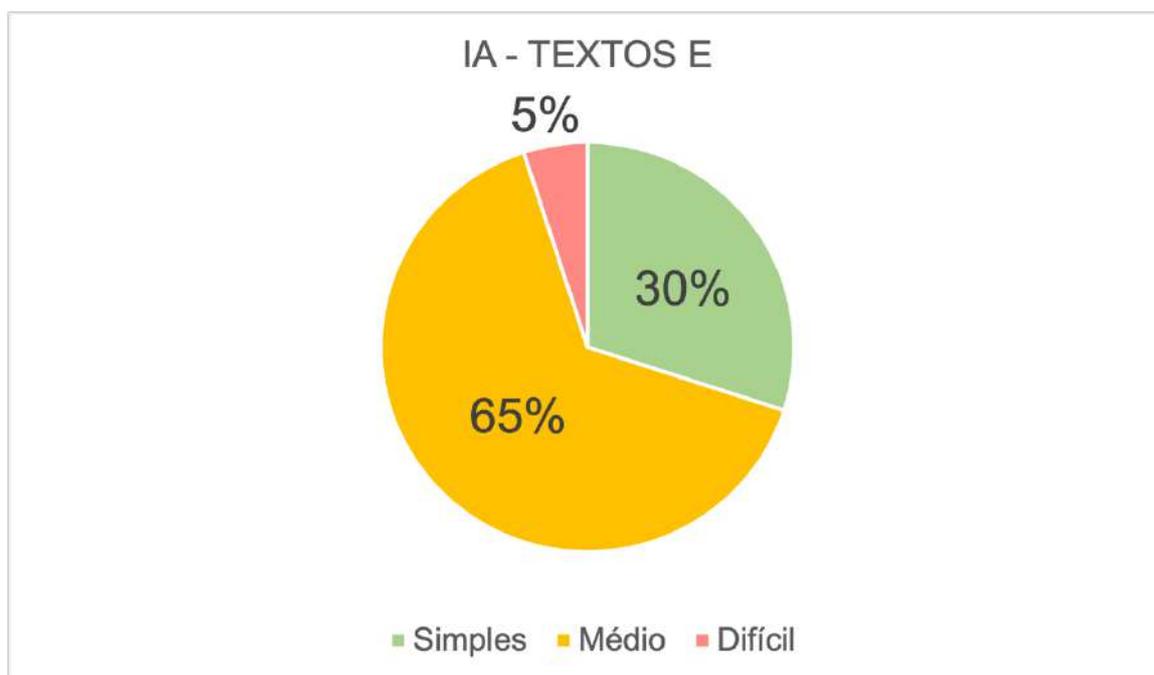
Ainda que a PS não deva ser utilizada sozinha para estimar complexidade lexical e semântica, seus valores servem, também, para destacar textos que exigem atenção para polissemia. Mesmo que se diante de resultados baixos de PS constatarmos que as palavras possivelmente polissêmicas presentes no texto não serão, de fato, potencializadoras de complexidade para nosso aluno, a verificação de textos “alertados” pela PS como tendo resultados indicativos de alta complexidade por polissemia ainda é válida com a finalidade de verificar se o contexto e/ou a posição sintática de unidades polissêmicas colaboram para compreensão ou inferência do significado correto.

Passamos, a seguir, para a interpretação da métrica IA.

8.5.1.1.2 Análise da métrica IA dos textos E

Para prosseguirmos com a análise de complexidade lexical e semântica, apresentamos, a seguir, o Gráfico 12 que exhibe o percentual de textos que são considerados simples, médios e difíceis pela métrica IA:

Gráfico 12 - Classificação percentual de complexidade de IA do livro *Elementary*



Fonte: Elaboração própria

Por meio do gráfico, podemos constatar que a interpretação de IA (que calcula o quão cedo ou o quão tarde as palavras são geralmente adquiridas de forma natural na língua inglesa e cuja valoração cresce proporcionalmente à complexidade), podemos observar que 30% dos textos podem ser considerados como tendo complexidade baixa, 65%, média e 5%, alta.

É importante destacar que o índice IA atribui pontos às palavras tendo em mente a “aquisição” de léxico de um falante nativo. Como nosso estudo está voltado para aprendizes de inglês como língua adicional, acreditamos que o nível de dificuldade deve ser considerado um pouco acima dos valores apresentados pela métrica.

Mesmo com a ressalva do parágrafo anterior, acreditamos que – de fato – os números sejam coerentes com nossa experiência em sala de aula. Ainda que existam vocabulários controlados em cada unidade (ou seja, léxico trabalhado

especificamente naquela unidade), seria praticamente impossível que os textos fossem redigidos utilizando somente palavras já aprendidas; sendo assim, é natural que exista um balanço entre a complexidade das palavras.

O texto considerado mais simples pelas classificações de IA é o E6 (IA: 231,03) e o mais complexo, o E12 (IA: 385,51) – lembrando que os valores de IA se comportam da seguinte forma: valores mais altos indicam maiores complexidades.

Por meio da tabela com os resultados das métricas, é interessante observar três detalhes: 1. Os textos mais simples estão na metade inicial do livro e; 2. O segundo texto mais complexo é o último do livro.

O texto E6, cujos resultados de IA indicam maior acessibilidade, também foi classificado pela TT como o mais acessível entre os textos do *Elementary*. Esse texto é composto por uma grande quantidade de palavras simples e concretas e que geralmente são aprendidas e utilizadas desde a primeira infância (como ‘*mother*’, ‘*house*’, ‘*brother*’). Nossa experiência em sala de aula também demonstra que a acessibilidade desse texto é alta, não só pelo léxico, mas de forma global.

Outro exemplo de indícios de acessibilidade pela IA foi o texto E10 (segundo melhor resultado dessa métrica). Nele, também podemos observar palavras do cotidiano e que são naturalmente presentes no dia a dia de uma criança, como ‘*room*’, ‘*night*’, ‘*morning*’, ‘*dinner*’, ‘*tired*’, ‘*bed* etc.

Ao fazermos uma breve consulta em listas de palavras aprendidas em cada faixa etária da infância, como em estudos apresentados por Graham et al. (1993) – que desenvolveram listas de palavras cuja ortografia é aprendida em cada ano escolar - podemos perceber que uma grande quantidade de palavras dos textos destacados como acessíveis pela IA é conhecida por crianças entre o primeiro e quinto ano do ensino que equivale ao fundamental do sistema brasileiro.

Já o texto E12, classificado como complexo pela IA (e pela TT como sendo o segundo texto com maior complexidade no quesito variação lexical), é o texto que tem uma grande quantidade de nomes próprios (como ‘*Agatha*’, ‘*Baskersvilles*’, ‘*Doyle*’) – fato que justificaria valores de IA não tão apropriados.⁶³ Mesmo que os nomes próprios possam ter influenciado os valores de IA, destacamos que essas unidades tendem a ser facilmente interpretadas por estudantes em razão da ortografia. Para essas

⁶³ As listas de palavras de cada um dos textos podem ser consultadas nos apêndices desta tese.

palavras, o importante é o conhecimento de serem nomes, sem que exista uma necessidade de relacionar significante e significado (já que são atribuídos a pessoas).

Ainda que os nomes presentes no texto E12 possam ter influenciado nos valores de IA, existem algumas palavras que podem ser complexas para um aluno desse nível (como *'nearly'*, *'early'* e *'novels'*, por exemplo).

No caso da primeira palavra (*'nearly'*), temos a possibilidade de substituição por *'approximately'* que, além de facilitar a compreensão, também serve para manutenção de vocabulário (visto que o significado pode ser facilmente inferido por um leitor falante de português devido à similaridade com a palavra 'aproximadamente'). Além disso, a utilização da palavra *'approximately'* no lugar de *'nearly'* é mais adequada, porque o texto E12 é apresentado na unidade 8 do livro e advérbios de modo são estudados apenas na unidade 11. Somando-se ao fato da distância entre o texto no qual a palavra é apresentada e a unidade na qual essa classe de palavras é aprendida, DuBay (2004) e o programa Plain Language (conforme já mencionado), advertem para o uso de advérbios porque, segundo esses pesquisadores, essa classe de palavras é considerada indicativa de complexidade.

O segundo exemplo de palavra potencialmente complexa no texto E12 (*'early'*) também pode ser complexa pelos mesmos motivos de *'nearly'*. Ao analisarmos o contexto dessa palavra, percebemos que ela não é essencial para o texto. O mesmo sentido poderia ser apresentado sem o uso da palavra, como apresentamos abaixo:

Frase original: *Many of the best early detective stories were short stories.*

Alternativa: *Many of the best detective stories were short stories.*

Ainda como exemplo de palavra potencialmente complexa no texto E12 (indicado como o texto mais complexo pela IA), temos a palavra *'novels'*. Essa palavra pode ser considerada complexa por ser um falso cognato para a língua portuguesa. Se o aluno não questionar ou o professor não atentar para a explicação dessa unidade lexical, ela pode ser compreendida como *'novela'*, alterando o significado do texto.

Outro texto com indicação de alta complexidade por IA é o E20 (segunda maior complexidade). Nele, podemos perceber uma certa quantidade de nomes próprios (como *'Bridget'*, *'Jane'*, *'Alan'*), mas também palavras que podem ser consideradas mais complexas, tais como *'atmosphere'*, *'definitely'*, *'exciting'*, *'however'*, *'obviously'* etc. Ainda que tais palavras sejam exemplos muito específicos, é válido lembrar que

uma vez que os textos são apresentados para alunos iniciais, algumas orientações de acessibilidade textual para falantes nativos devem ser observadas (já que esses fatores são complexos para nativos, também serão para não-nativos), tais como o uso de adjetivos e advérbios longos e de conjunções complexas; especialmente quando existem alternativas de simplificação.

Para exemplificar, considere a seguinte frase do texto E20: “*The atmosphere is great, perfect for a Friday or Saturday night.*”.

Embora a palavra ‘*atmosphere*’ seja um exemplo não tão complexo, não podemos inferir que um aluno iniciante compreenderá o significado dessa unidade como “lugar” e não como o cognato “atmosfera” da língua portuguesa. Sendo assim, essa palavra pode ser considerada um exemplo de léxico que pode ter um alto potencial de distração.

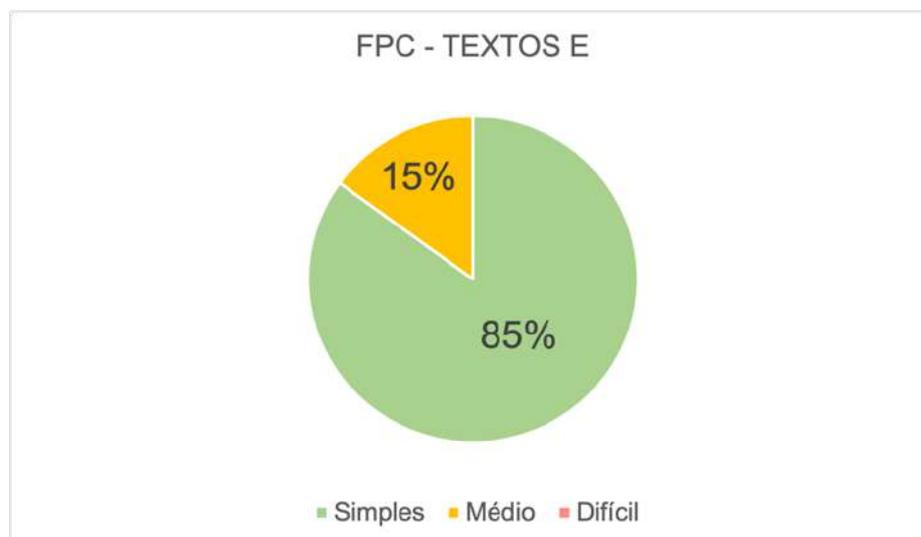
Com base na interpretação dos valores de IA, podemos compreender que o léxico presente nos textos do livro *Elementary* merece atenção em algumas instâncias uma vez que 70% deles apresenta complexidade média e alta. É válido lembrar que a lista-base da métrica IA, conforme apresentado por McNamara et al. (2014), foi analisada tendo em mente a idade de aprendizado dos vocábulos por falantes nativos. Logo, se 70% dos textos apresentam indicações de IA que indicam complexidade média e alta para leitores que têm o inglês como primeira língua, é válido considerar que existem apropriações que possam ser realizadas para leitores não-nativos de ILA. Para isso, as marcações no nosso quadro de valores das métricas dos textos E podem nos orientar para quais merecem atenção detalhada quando forem apresentados para os leitores do nosso estudo.

Continuando com a análise das métricas de léxico e semântica, apresentamos, na próxima seção, os valores de FPC.

8.5.1.1.3 Análise da métrica FPC dos textos E

Prosseguindo com nossa análise sobre as métricas de léxico, o gráfico abaixo (Gráfico 13) apresenta as classificações de complexidade de FPC – que estima a familiaridade com palavras de conteúdo.

Como podemos observar, as interpretações de FPC são positivas. 85% dos textos são considerados como simples, enquanto 15% como médios. Nenhum texto foi considerado complexo.

Gráfico 13 - Classificação percentual de complexidade de FPC do livro *Elementary*

Fonte: Elaboração própria

O texto considerado mais simples é o E10 (FPC: 602,04) e o menos acessível (embora classificados como tendo complexidade média), o E1 (FPC: 579,96).

O texto mais acessível pela interpretação dos valores de FPC (E10) também foi apontado por IA como o segundo texto de maior acessibilidade (conforme previamente mencionado) e apresenta uma grande quantidade de palavras comuns a crianças desde a primeira infância.

Visto que já discutimos a simplicidade do texto E10, destacamos o segundo melhor resultado de FPC como forma de ilustrar a métrica: o texto E4 (FPC: 597,04). Nesse texto, podemos observar vocábulos simples e que constam na lista de Graham et al. (1993) classificados entre a primeira e quinta série, tais como 'wake', 'live', 'breakfast', 'sun', 'fruit', 'juice' etc.

Por ser um texto que descreve a rotina de um personagem, o E4 é composto por verbos e substantivos (palavras de conteúdo) que são não só cotidianos para falantes nativos desde a primeira infância, mas também frequentemente apresentados em unidades iniciais para aprendizes de ILA. Sendo assim, os valores de FPC indicam baixa complexidade.

Obviamente, pela posição do texto E4 no livro (no segmento de revisão posterior à unidade 4), algumas palavras podem ser desconhecidas para alunos em níveis iniciais. Entretanto, esse texto está alocado em uma unidade posterior à apresentação e prática do *Present Simple* e de exercícios e outros exemplos que apresentam a maior parte dos verbos e substantivos presentes nesse texto.

Já o texto considerado como mais complexo pela FPC (E1) – curiosamente o primeiro do livro – é justificado ao realizarmos uma análise holística: há uma grande quantidade de nomes próprios (como *Rockefeller*, *Empire State*, *Coney Island*). Tal característica valida, mais uma vez, a necessidade da análise de um linguista. Em sala de aula, tais unidades lexicais não seriam impeditivas de aprendizagem não só pela grafia indicar nomes próprios, mas também pelo apoio visual do texto (conforme Figura 32 a seguir):

Figura 22 - Parte do texto E1 com auxílio de recursos visuais como disposto no livro



Fonte: (LATHAM-KOENIG, OXENDEN e LAMBERT, 2018, p. 22)

Se analisarmos o segundo texto mais complexo por FPC (E15), verificamos que ele também apresenta alguns nomes próprios (como '*Verana*', '*Puerto Vallarta*', '*Huskey Lodge*') mas, ao mesmo tempo, traz algumas palavras que podem ser complexas para um leitor aprendiz de ILA em níveis iniciais, tais como '*guesthouses*', '*balconies*', '*whale-watching*', '*underground*', '*candlesticks*', '*kennels*', '*sledges*' etc. Pela posição do texto no livro e pelo vocabulário complexo que, em algumas

instâncias, não permite uma substituição simples, acreditamos que existam muitas adequações a serem realizadas ou, ainda, alteração desse texto por outro.

Uma vez que somente três textos apresentaram valores baixos de FPC, poucas são as necessidades de destaques relativas à métrica.

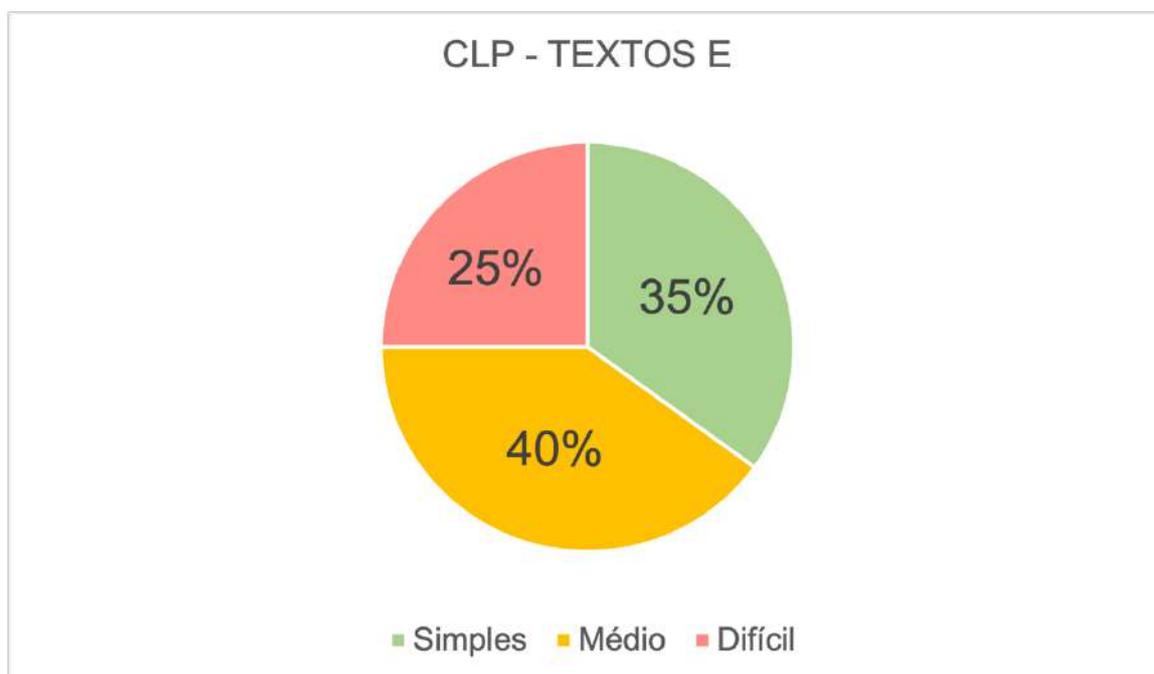
Em síntese, a interpretação dos valores de FPC serve como indício de que a composição dos textos é adequada no quesito familiaridade com palavras de conteúdo, mas que é possível um olhar mais atento ao vocabulário de determinados textos a fim de adotarmos estratégias de simplificação lexical para casos mais particulares – como no caso de alunos disléxicos - se necessário pelo julgamento do professor. Como exemplo, seria possível aplicar a estratégia de transformação de substantivo para locução substantiva a fim de substituir a palavra ‘*slaughterhouse*’, presente no texto E8, por ‘*place where animals are killed to make food*’; e a estratégia de substituição lexical de frequência para substituir o advérbio ‘*however*’ (também presente no texto E8) por ‘*but*’ – especialmente porque o primeiro advérbio é indicado pelo CEFR como apropriado para ser ensinado no nível pré-intermediário ou a palavra ‘*dogs’ kennels*’ (texto E15) por ‘*dogs’ houses*’ (embora ‘*kennel*’ e ‘*house*’ não tenham exatamente o mesmo significado).

Com base nessas observações sobre o FPC, enfatizamos que ainda que essa métrica indique que a maior parte dos textos apresenta vocabulário familiar e, ao mesmo passo, certa acessibilidade, ainda é possível aprimorar a escolha lexical sem prejudicar a manutenção linguística. Talvez, por exemplo, a palavra ‘*however*’ possa ser compreendido como “+2” em vez “+1” na fórmula “*c + 1*” de insumos de Krashen (1987), indicando que, possivelmente, esse insumo seja não um passo mais avançado que os conteúdos atuais do estudante, mas dois passos; ou seja, complexo.

Prosseguindo com a análise das métricas de léxico e semântica apresentamos, a seguir, a métrica CLP.

8.5.1.1.4 Análise da métrica CLP dos textos E

O gráfico abaixo (Gráfico 14) nos indica que 25% dos textos podem ser classificados como difíceis, enquanto 40%, médios e 35%, fáceis.

Gráfico 14 - Classificação percentual de complexidade de CLP do livro *Elementary*

Fonte: Elaboração própria

Diante de resultados de CLP que classificam muitos dos textos entre complexidades média e alta, precisamos atentar para aqueles destacados como em amarelo e vermelho no quadro das métricas, principalmente ao lembrarmos que os textos em foco no momento fazem parte do livro *Elementary*. Logo, 65% dos textos terem sido classificados dessa forma nos alerta que pode haver indícios relacionados à concretude das palavras que sugerem possível dificuldade de leitura.

Obviamente, assim como enfatizamos para as demais métricas, a CLP sozinha não é suficiente para definir com total precisão a complexidade do léxico dos textos, mas em conjunto com as demais, permite construirmos um panorama mais completo sobre possíveis elementos de dificuldade.

Como de costume, destacamos o texto com melhor classificação, que para a CLP foi o E6 (CLP: 98,26%) e o texto cujo valor do índice indicou maior complexidade, o E20 (CLP: 9,51%).

Ao observarmos o quadro das métricas de forma global, podemos perceber que o texto E6 também é destacado como o mais simples pela IA e por TT, e que o E20 é considerado o segundo mais complexo por IA.

O texto E6, conforme já observado na análise das demais métricas, é uma leitura geralmente considerada simples pelos alunos. Essa narrativa conta a história

de um rapaz que viaja pelo deserto para tentar encontrar recursos para ajudar a família e que encontra um palácio no qual é bem recebido e recebe moedas de ouro do príncipe. Ao analisarmos a lista de palavras dessa história, podemos constatar que a maior parte delas representa, de fato, uma ideia concreta e que poucas são as unidades lexicais que se referem a algo abstrato no mundo real. O texto traz muitas palavras descrevendo objetos (*'bottle', 'bed', 'house', 'coins', 'coffee'*) e poucas cuja relação significante-significado seja relativamente abstrata (*'beautiful', 'help', 'live'*). Ainda que o texto tenha palavras com significados não-concretos, em uma escala de abstrativismo, acreditamos que elas não exijam a evocação de conceitos difíceis de serem compreendidos especialmente por que os conceitos não são estranhos à língua portuguesa.

Outro texto com valor alto de CLP é o E1 (CLP: 95,82%), de forma similar ao E6, apresenta vocabulário com conceitos amplamente concretos – uma vez que descreve atividades e lugares em Nova York. No texto, encontramos com mais frequência palavras como *'building', 'camera', 'car'* do que léxicos abstratos (como *'cheap'* e *'traditional'*).

Além dos textos E6 e E1, outros que apresentaram valores excelentes de CLP (que, lembrando, é uma métrica cujos valores são percentuais) foram: E4, E5, E11, E15 e E19, sendo que com exceção do E15 (que resultou valor baixo de PS), todas as demais métricas desses textos indicaram complexidade baixa e média (e, majoritariamente, os índices com valores destacados como indicação de complexidade média foram valores próximos do limite entre complexidade média e baixa).

No que diz respeito à concretude das palavras, podemos afirmar o seguinte:

- No texto E4 (CLP: 85,31%) são muitas as palavras relacionadas a objetos e lugares do cotidiano (como *'sun', 'house', 'car'*);
- O texto E5 (CLP: 90,15%) descreve lugares e atividades para fazer em Londres e, também, é rico em palavras que expressam ideias concretas (como *'camera', 'garden', 'birds'*);
- No texto E11 (CLP: 89,25%) podemos observar palavras concretas para descrever atividades paranormais em hotéis, como *'room', 'man', 'lights'*). Entretanto, ainda que o texto seja considerado acessível pela CLP, existem unidades lexicais que podem ser complexas para nossos leitores (como *'haunted', 'suddenly', 'murdered'*);

- O texto E15 (CLP: 92,92%) também tem a temática de hotéis. Contudo, devido ao fato de que a temática subjacente não é atividades paranormais (como no texto E11), o vocabulário é mais comum e concreto, sendo rico em palavras que descrevem locais e objetos (como *'trees'*, *'pools'*, *'radio'*);
- No texto E19 (penúltimo do livro, com valor de CLP de 80,78%) lemos uma entrevista com o ator Ian McKellen. Para formular resposta às perguntas, o ator utiliza léxico frequentemente concreto, como *'music'*, *'bed'* e *'books'*, por exemplo. Além dos indícios de baixa complexidade pela CLP, nossa experiência com esse texto em sala de aula nos permite destacar que sua organização estrutural aumenta a acessibilidade. Visto que o texto é formado por perguntas e respostas, os tópicos são fáceis de serem identificados e as informações são claras. Tal organização estrutural é altamente indicada na literatura a respeito de AT, inclusive pela organização *Plain Language* e por Silva (2018).

O texto E20, considerado mais complexo pela CLP, é o último texto apresentado no livro. Essa posição permitiria um pouco mais de complexidade em relação aos anteriores, embora não valores tão baixos quanto o da métrica em questão. Ao realizarmos a leitura desse texto, encontramos altas frequências de palavras abstratas (como advérbios e outros elementos dêiticos) e baixa frequência de palavras concretas (em relação às abstratas). Além disso, as palavras concretas desses textos tendem a se repetir, criando uma proporção de quase três palavras abstratas para uma concreta. Exemplos de palavras abstratas nos textos são *'here'*, *'near'*, *'very'*, *'quite'*, *'fun'*.

Se analisarmos detalhadamente todos os textos para encontrarmos as palavras abstratas presentes neles, podemos perceber que poucos são os casos que, para leitores proficientes, poderiam ser considerados verdadeiramente potenciais de complexidade. Entretanto, levando em consideração que nossos leitores são aprendizes de ILA em estágios iniciais, é importante atentar para as indicações de CLP dos textos considerados como tendo complexidade média ou alta. Por esse motivo, é importante apresentarmos as seguintes observações sobre textos com valores baixos de CLP:

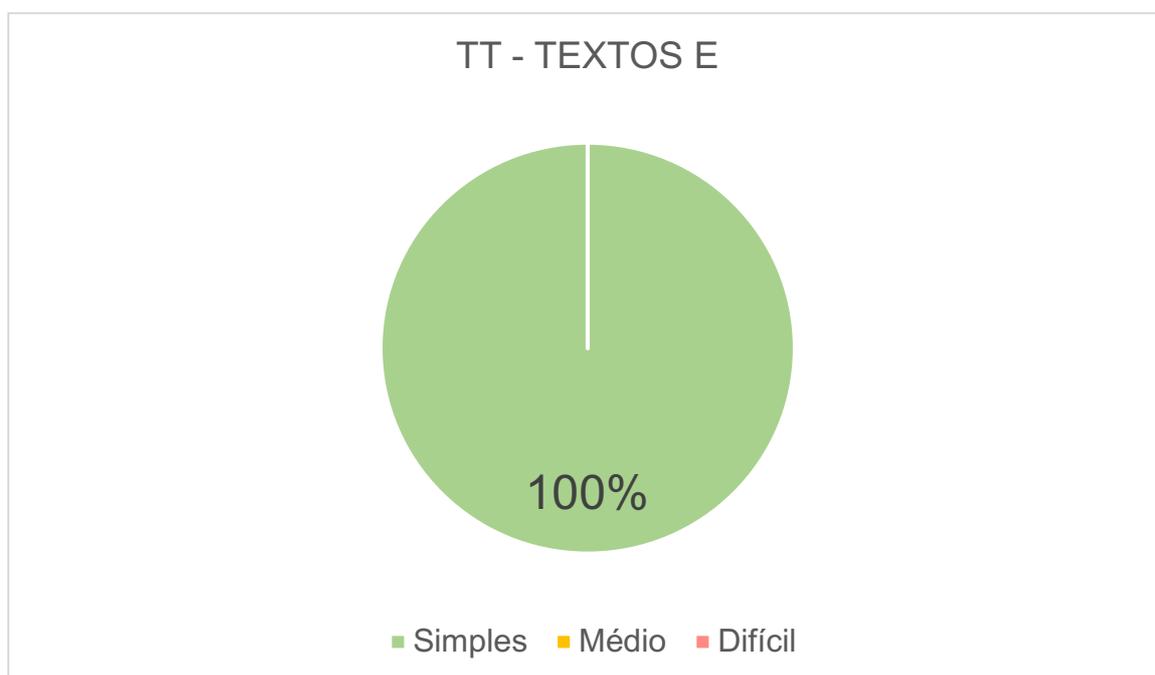
- O texto E8, embora tenha resultados baixos de CLP (36,32%) não tende a ser considerado complexo em sala de aula por razões de concretude lexical. Alguns termos nesse texto são complexos para alunos em níveis iniciais (como *'slaughterhouse'*, *'portrait'* e *'major'*), mas as palavras abstratas são, na grande maioria, adjetivos conhecidos;
- No texto E9 (CLP: 50%), embora o valor de CLP tenha indicado uma proporção 1:1 de palavras concretas por abstratas, não tende a ser considerado complexo, uma vez que as palavras abstratas também são, em grande maioria, adjetivos comuns. Outras palavras (como *'bit'*, *'suddenly'* e *'between'*) podem ser fatores de complexidade não só por serem abstratas, mas especialmente por serem palavras que são foco de estudo em unidades posteriores à unidade na qual o texto é apresentado;
- O texto E14 (CLP: 42,86%) apresenta frequentes palavras abstratas, uma vez que a temática é relacionada ao misticismo. Alguns exemplos são *'fortune'*, *'here'*, *'soon'*, *'well'* e *'love'*.

Por meio da análise de CLP, constatamos que os valores dessa métrica servem, como já mencionamos, como indicativos de quais textos merecem nossa atenção. É válido, entretanto, que os professores analisem as palavras abstratas em textos com valores baixos de CLP para analisar se essa falta de concretude é um fator potencializador de complexidade ou se, ainda que algumas unidades sejam abstratas, o léxico seja conhecido dos alunos.

Por fim, para concluir a observação dos comportamentos métricos de léxico e semântica, passamos para a análise de TT.

8.5.1.1.5 Análise da métrica TT dos textos E

Conforme podemos observar pelo gráfico abaixo (Gráfico 15), todos os valores de TT dos textos E classificam a complexidade relativa à diversidade lexical como baixa.

Gráfico 15 - Classificação percentual de complexidade de TT do livro *Elementary*

Fonte: Elaboração própria

Nossa experiência com os textos do nosso *corpus* de estudo nos permitia concluir, de antemão, que, na prática, a extensão dos textos tem sido adequada, enquanto a variação lexical, em algumas instâncias, pode ser um fator de alerta. É interessante analisarmos diversidade lexical tendo dados numéricos que nos permitam observar não só quantidade de palavras únicas (*types*) e de palavras totais de cada texto (*tokens*), mas também se há uma progressão dessas quantidades (fenômeno que é esperado, uma vez que cada unidade do livro tem por objetivo ser um pouco mais avançada que a anterior).

Por intermédio da análise dos *types*, podemos ter uma ideia da quantidade de novas palavras que podem estar sendo apresentadas para os estudantes (embora saibamos que a quantidade de *types* não irá refletir com total fidedignidade que essas palavras sejam novas para os estudantes e que não tenham sido apresentadas já em textos anteriores). Por meio da observação dos *tokens*, conseguimos observar se a progressão de extensão textual (por meio de quantidades de palavras) está sendo progressiva).

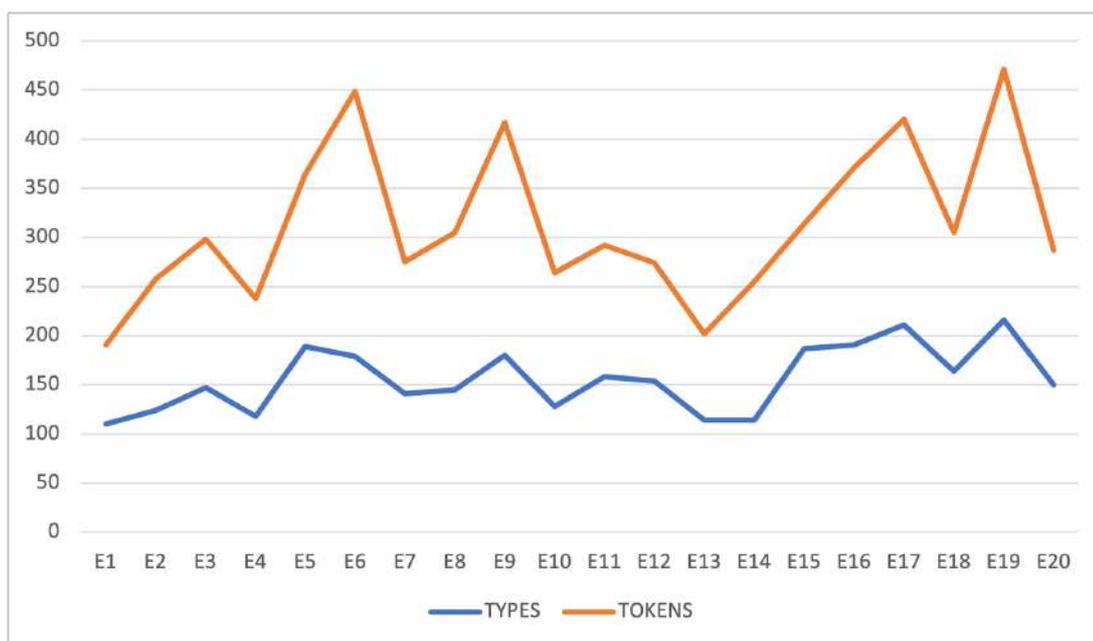
Portanto, para que possamos analisar os objetivos propostos nos parágrafos acima, o quadro abaixo (Quadro 29) apresenta a quantidade de *types* e de *tokens* de cada um dos textos do livro *Elementary*:

Quadro 29 - *Types e Tokens* dos textos do livro *Elementary*

TEXTO	TYPES	TOKENS
E1	110	191
E2	124	258
E3	147	298
E4	118	238
E5	189	364
E6	179	449
E7	141	275
E8	145	305
E9	180	417
E10	128	264
E11	158	292
E12	154	274
E13	114	202
E14	114	255
E15	187	314
E16	191	371
E17	211	420
E18	164	305
E19	216	471
E20	150	287

Fonte: Elaboração própria

Com a finalidade de tornar a análise mais visual, o gráfico abaixo (Gráfico 16) apresenta os valores do quadro acima por meio de linhas:

Gráfico 16 - *Types e Tokens* dos textos do livro *Elementary*

Fonte: Elaboração própria

Ao analisarmos o Gráfico 3 e o Quadro 22, podemos fazer as seguintes constatações:

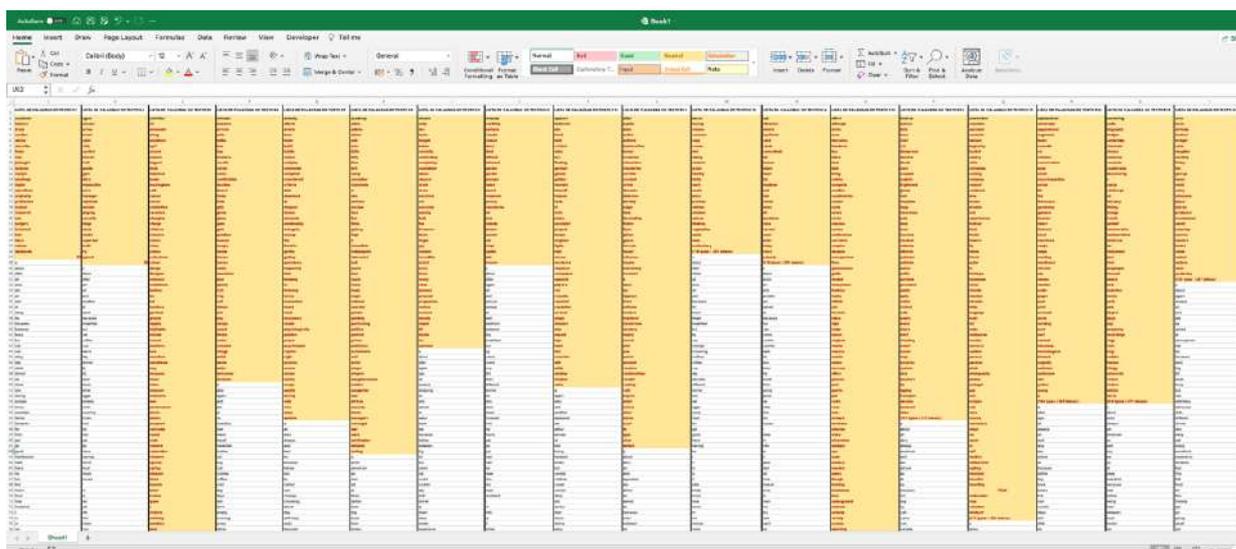
I. Há uma grande diferença da quantidade de palavras em cada texto (*tokens*) entre o primeiro texto e o último; fato que demonstra coerência de progressão;

II. A progressão não é constante. Entretanto, a maior redução de palavras está no meio do gráfico (que representa textos que são apresentados no meio do livro), o que torna coerente com a proposta do *English File*, que divide cada livro em duas partes (inclusive vendendo também de forma separada) e propõe a utilização de uma parte em cada nível de ensino. Sendo assim, a redução de *tokens* na posição onde ocorre pode indicar a preocupação da editora com a extensão dos materiais que serão utilizados para alguns alunos como os primeiros contatos textuais ao utilizarem a segunda parte do livro em um nível “novo”;

III. A quantidade de *types* também é progressiva, ainda que inconstantemente, o que pode indicar avanço na apresentação de novos vocabulários.

Ao colocarmos as listas de palavras no Excel, podemos desenvolver uma fórmula que destaca as células que contenham dados únicos. Ou seja, é possível destacar e contar as palavras que não se repetem entre todos os textos (palavras novas), conforme ilustra a imagem abaixo (Figura 23):

Figura 23 - Ilustração da seleção de unidades lexicais novas por meio do Excel



Fonte: Elaboração própria

A partir dessa seleção e contagem podemos, então, averiguar quantas e quais palavras novas foram introduzidas em cada texto e, assim, analisar a progressão de *inputs* de vocabulário novo. O quadro abaixo (Quadro 30) demonstra quantas palavras novas/únicas foram introduzidas por cada texto:

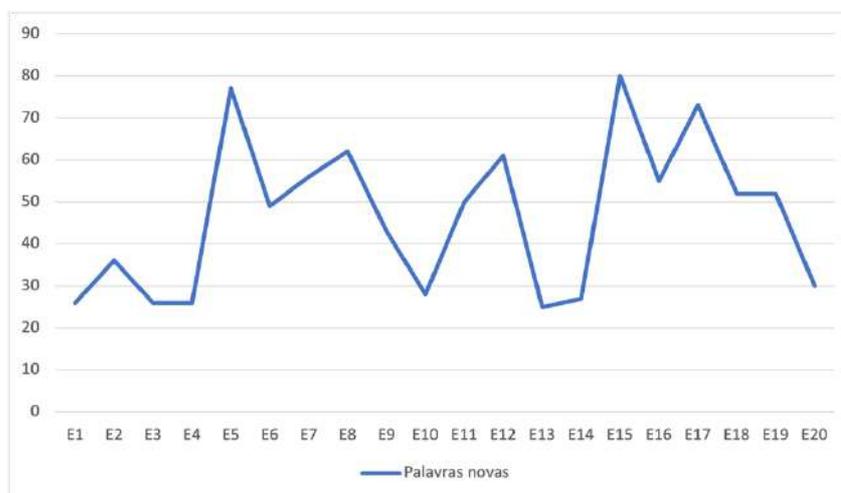
Quadro 30 - Palavras novas introduzidas por cada texto do livro *Elementary*

TEXTO	QUANTIDADE DE PALAVRAS NOVAS/ÚNICAS
E1	26
E2	36
E3	26
E4	26
E5	77
E6	49
E7	56
E8	62
E9	43
E10	28
E11	50
E12	61
E13	25
E14	27
E15	80
E16	55
E17	73
E18	52
E19	52
E20	30

Fonte: Elaboração própria

O quadro acima (Quadro 30) nos auxilia a observar que mais do que 25 novas palavras são apresentadas por texto, chegando até 80 novos vocábulos. Obviamente, a quantidade de palavras novas introduzidas em um texto por si só não é suficiente para estimar complexidade, mas podemos refletir que quanto maior for o número de palavras novas, maior é a necessidade de conhecimento de vocabulário para a interpretação de um texto. Acreditamos, por nossa experiência como docentes, que acima de 20 ou 30 novas unidades lexicais possam ser consideradas demasiadamente excessivas e se tornam, conseqüentemente, fatores de complexidade.

A coleta dos dados a respeito da quantidade de palavras novas apresentadas por cada texto também nos permite desenvolver um gráfico para observar se há progressão na quantidade de vocabulário novos apresentadas, conforme o Gráfico 17 abaixo:

Gráfico 17 - Quantidade de palavras novas introduzidas por texto do livro *Elementary*

Fonte: Elaboração própria

Como o gráfico acima auxilia a perceber, não há constância na apresentação de palavras novas; mas é interessante observar que existem determinados períodos que a disparidade entre quantidade de palavras novas é muito grande, como entre os textos E3 (26 palavras novas) e E5 (77), 11(50) e 13(25), 13 (25) e 15 (80), e 17(73) e 20(30) – o que nos leva a indagar se a quantidade de palavras novas não se torna um fator atenuante da complexidade e se a falta de constância na quantidade de vocábulos novos introduzidos não prejudica o insumo (c+1) necessário para a manutenção linguística. Tentaremos observar e responder essas hipóteses por meio da análise das métricas a seguir.

Por meio dos valores de TT, podemos perceber que os textos com maiores variações lexicais (maiores valores de TT⁶⁴) foram o E15 (TT: 596) e o E12 (582), enquanto os textos com menores variações (menores valores de TT) foram o E6 (TT: 402) e o E9 (430). A média dos valores dessa métrica foi de 507 que, de acordo com nossa metodologia classificatória, indicaria uma média de complexidade baixa.

No quadro abaixo (Quadro 31), apresentamos trechos dos textos destacados acima:

Quadro 31 - Trechos de textos do livro *Elementary* para ilustrar resultados de TT

TEXTO	TRECHOS
E15 (TT: 596)	<i>You don't need to travel to the Arctic Circle to spend time with Siberian huskies. At the Husky Lodge in the Swiss canton of Schwyz, dog-lovers can</i>

⁶⁴ Lembrando que os valores originais de TT variam nas casas decimais entre 0 e 1. Para fins de facilitar a leitura, multiplicamos esse índice por mil para arredondar os valores.

	<i>sleep in cabins, heated with wood fires, next to the dogs' kennels. During the day you can join in with their training runs. They pull sledges in winter and bikes and carts in summer. In the evening there's an excellent restaurant. And if you have to work, there are three rooms for small conferences and seminars. If you can't afford the cabins, there's a campsite too - though maybe only in summer!</i>
E12 (TT: 582)	<i>Many of the best early detective stories were short stories. Edgar Allan Poe (1809-1849) was probably the author of the first detective story. Arthur Conan Doyle (1859-1930) created perhaps the most famous detective, Sherlock Holmes. Father Brown was also a detective who was nearly as popular as Holmes; he was created by G.K. Chesterton (1874-1936).</i>
E6 (TT: 402)	<i>They live in a small house in the desert, near the mountains, a long way from a town or village. Their house only has two rooms. They live with their mother. Hassan wants to help his mother, and he decides to leave home. His mother gives him bread, and water in a glass bottle. Hassan walks through the mountains again, and after five days he arrives home. Walid and their mother watch 'him open the box. His mother is very happy!</i>
E9 (TT: 430)	<i>I went to a hotel with some friends. Then a few minutes later, I could hear people outside celebrating. It was awful - I didn't get out until almost 1.00 in the morning. And on top of it all, my friends didn't realize that I wasn't there! I wore a lovely dress, but after dinner we got a taxi back to the hotel and I put on boots and a warm jacket. But then just before midnight people came out again and danced and drank till about 5.00 in the morning. I thought it was time to start the countdown to midnight, I saw that people were all at the bar. I looked at my watch and I realized it was about two minutes slow.</i>

Fonte: Elaboração própria

No trecho do texto E15 (considerado o texto com maior variação lexical) podemos perceber que essa alta presença de léxico diversificado é justificado porque o texto tem como tema 'hotéis diferentes' e, além de apresentar os nomes próprios do hotel, também ilustra as razões pelas quais esses hotéis são diferentes, desde componentes arquitetônicos até presença de eletrônicos, atrativos etc.

O texto E12, classificado em segundo lugar como apresentando maior complexidade lexical indicado por variação, também apresenta uma característica peculiar que influenciou nos resultados da métrica: o texto narra a história de alguns detetives.

O fato de que os textos destacados têm alta variação lexical por apresentarem uma grande quantidade de nomes próprios comprova a essencialidade de uma análise de um linguista diante de indicações métricas. Se diante da indicação de alta variedade lexical nesses textos (pelos indícios das métricas) não tivéssemos atentado

a analisá-los individualmente, teríamos acreditado que existe um léxico altamente diversificado quando, na verdade, a complexidade não é tão alta.

No trecho do texto E6 (que apresentou a menor variação lexical e, por consequência, menor complexidade) nos auxilia a perceber que o texto é uma narrativa com uma grande repetição de palavras (como demonstrado pelas palavras em negrito no trecho exibido no Quadro 31). Nossa experiência em sala de aula com essa história também nos demonstra que os alunos não tendem a ter dificuldade com essa leitura.

Além das palavras destacadas no trecho do texto E6 disposto no Quadro 31, existem mais de vinte palavras que se repetem constantemente durante o texto, como *'glass', 'bottle', 'water', 'food', 'desert', 'sky', 'clothes', 'gold', 'coins'* etc.

O texto E9 é composto de três pequenas histórias distintas sobre experiências na noite de Ano Novo. Esse texto também apresenta pouca variação lexical, uma vez que muitas palavras se repetem entre as narrativas (por terem o mesmo tema) como podemos notar pelo trecho destacado no Quadro 31 apresentado anteriormente.

É importante observar, contudo, que no texto E9 temos a presença de uma palavra homógrafa sendo utilizada com dois significados distintos, *'watch'* (assistir e relógio). Ainda que a palavra seja usual e tenha sido apresentada em unidades anteriores à aparição do texto, a presença dela corrobora – mais uma vez – a necessidade de observação de um linguista. Ainda que o texto tenha sido indicado como tendo baixa variedade lexical, a métrica TT não consegue distinguir diferentes significados diante de homografia. Logo, faz-se necessária a confirmação de profissionais que possam encontrar pontos que necessitem de atenção.

Embora a quantidade de *types* e *tokens* sirva como indicativo da necessidade de observação da variedade lexical em alguns textos (com valores de TT mais altos), acreditamos que a variação do vocabulário dos textos do *Elementary* não seja, necessariamente, o maior fator de complexidade. Existem outros pontos relacionados ao léxico que merecem mais atenção (quanto ao uso de palavras complexas em textos posicionados em pontos iniciais do livro quando existiriam alternativas, como no caso de *'however'* e *'but'*, por exemplo).

É importante lembrar, ainda, que cabe ao professor desenvolver estratégias de desenvolvimento lexical e alertar seus alunos que para realizar uma leitura eficiente com a finalidade de encontrar informações e ideias globais (uma vez que os alunos estão em níveis iniciais de ILA), não é preciso que saibam o significado de todas as

palavras. O docente precisa saber determinar quais as palavras-chave da unidade e quais merecem mais atenção para cada momento.

Além disso, é importante lembrar que embora os textos apresentem uma grande quantidade de palavras distintas, muitas dessas palavras fazem parte do vocabulário controlado apresentado pela unidade e seu objetivo de inserção nos textos é oportunizar a prática. Sendo assim, é essencial lembrar que para termos uma análise mais precisa sobre possíveis traços de complexidade lexical e semântica, precisamos considerar não só a TT, mas as outras métricas previamente analisadas.

Com o intuito de finalizarmos a análise das métricas de léxico de forma mais global, comparamos, a seguir, os valores das métricas dos textos do livro *Elementary* aos resultados dos textos LIS e LIC.

8.5.1.1.6 Análise contrastiva das métricas de léxico e semântica de E com os de LIS e LIC

Nesta seção, apresentaremos um quadro (Quadro 32) no qual os valores das métricas dos textos E entre os valores médios dos textos simples e complexos do *corpus* comparativo. Sendo assim, para cada uma das métricas, a coluna à esquerda dispõe a média dos textos LIS, a coluna do meio exibe os valores dessa métrica dos textos E, e a coluna da esquerda, a média desse índice para os textos LIC.

O objetivo principal do Quadro 24 é permitir uma melhor visualização do comportamento das métricas de léxico dos nossos textos de estudo. Uma vez que os textos são partes do livro que se destina a estudantes de níveis iniciais (*Elementary*), acreditamos que o ideal seria que os valores das métricas se aproximem dos valores dos índices dos textos LIS (textos simples do *corpus* comparativo).

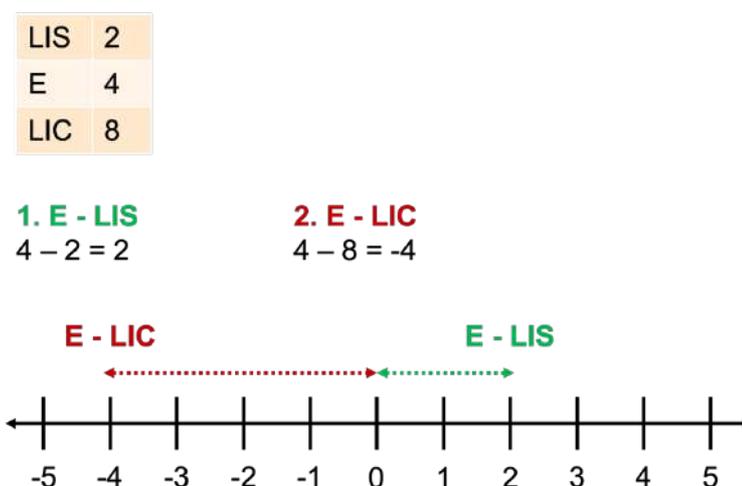
Com a finalidade de apresentar dados verossímeis, para que possamos verificar se os valores das métricas do texto E se aproximam mais de LIS ou de LIC, realizamos os seguintes procedimentos lógicos com o apoio do Excel:

1. E – LIS;
2. E – LIC;
3. Entre os valores de (E-LIS) e (E-LIC), o valor mais próximo de zero representa a maior proximidade. Ou seja, se (E-LIS) tiver um valor mais próximo de zero do que (E-LIC), isso significa que o texto em questão tem

valor métrico mais próximo à média de valor dos textos LIS do que da média dos textos LIC.

A fim de ilustrar de forma mais acessível a fórmula lógica adotada, considere o exemplo da imagem abaixo (Figura 24):

Figura 24 - Ilustração dos procedimentos lógicos para determinação de aproximação de valores métricos



Fonte: Elaboração própria

No exemplo ilustrado pela imagem acima, a suposta métrica analisada tem valor 4 para E, 2 para LIS e 8 para LIC.

Ao realizarmos a subtração do valor de E pelo de LIS, obtemos 2. Já o índice E menos LIC, -4.

Sendo assim, ao colocarmos esses valores em uma linha numérica, podemos perceber que a seta pontilhada em vermelho é maior do que a seta pontilhada em verde. Isso significa que a distância entre (E – LIC) e zero é maior do que a distância entre (E – LIS) e zero. Logo, o valor dessa suposta métrica para o texto E é mais próximo ao valor médio dessa mesma métrica dos textos simples do *corpus* comparativo.

Para que possamos tornar a interpretação mais visual, destacaremos em verde, no quadro abaixo (Quadro 32), os índices que forem mais próximos aos valores de LIS, e em laranja os índices mais próximos aos valores de LIC. Além disso, para facilitar a comparação, eliminamos as casas decimais (fator que não impactará na interpretação dessa etapa da pesquisa):

Quadro 32 - Comparativo de métricas lexicais-semânticas entre os textos do *corpus* comparativo e os textos do livro *Elementary*

	PS LIS	PS E	PS LIC	IA LIS	IA E	IA LIC	FPC LIS	FPC E	FPC LIC
E1		4,22			311,18			579,96	
E2		3,35			275,72			595,44	
E3		4,03			303,48			593,69	
E4		4,13			270,03			597,04	
E5		3,67			298,08			586,95	
E6		4,17			231,03			593,85	
E7		3,75			281,18			583,49	
E8		3,47			308,52			582,79	
E9		4,05			269			591,14	
E10	Valor médio	3,79	Valor médio	Valor médio	236,42	Valor médio	Valor médio	602,04	Valor médio
E11	PS:	3,64	PS:	IA:	268,6	IA:	FPC:	580,16	FPC:
E12	4,07	3,32	3,55	267,18	385,51	367,08	578,87	567,85	563,43
E13		3,5			266,12			580,33	
E14		4,11			291,44			585,15	
E15		3,34			314,75			578,53	
E16		3,95			330,46			584,48	
E17		4,3			280,5			582,34	
E18		3,6			287,59			587,05	
E19		4,08			284,15			588,4	
E20		3,97			332,35			593,16	
Média		3,82			291,31			586,69	

	CLP LIS	CLP E	CLP LIC	TT LIS	TT E	TT LIC
E1		95,82			578	
E2		59,1			478	
E3		45,22			490	
E4		85,31			504	
E5		90,15			530	
E6		98,26			402	
E7		77,34			513	
E8		36,32			489	
E9		50			430	
E10	Valor médio	63,68	Valor médio	Valor médio	482	Valor médio
E11	CLP:	89,25	CLP:	TT:	543	TT:
E12	77,49	70,19	49,22	661	582	734
E13		76,11			567	
E14		42,86			446	
E15		92,92			596	
E16		60,26			499	
E17		72,57			502	
E18		55,17			536	
E19		80,78			462	
E20		9,51			517	
Média		67,54			507	

Fonte: Elaboração própria

Por meio do quadro acima e das etiquetas de cor representando proximidade, observamos que a maior parte dos valores métricos dos textos E se aproximam da média dos textos mais simples. Sendo que, dos valores métricos de E:

- a. 50% dos textos apresentam valores de PS que se aproximam da média dos textos simples, enquanto 50%, dos complexos. Entretanto, a média de PS dos textos E é mais próxima ao valor da média de LIS indicando que, os textos tendem a ter traços de polissemia mais similares aos textos simples;
- b. 85% dos valores de IA se aproximam dos textos simples, sendo apenas 15% mais aproximados dos textos LIC. Isso indica que o léxico presente nos textos E é, majoritariamente, mais similar ao dos textos mais simples do *corpus* comparativo – fato que também se corrobora ao verificarmos que a média de IA dos textos E também se aproxima da média de LIS. Além disso, a distância entre a média de IA dos textos E para a média dos textos LIS é muito superior à distância percebida entre o valor médio de E e dos textos LIC;
- c. 95% dos valores de FPC são mais próximos de LIS, somente 5%, de LIC. Somada à informação a respeito de IA, essa estatística aumenta os indícios de que os textos E apresentam vocabulários indicativos de alta acessibilidade uma vez que seus valores de FPC também são mais próximos aos dos textos simples. A média dos valores de FPC de E também são mais próximos à média de LIS (embora a distância entre os números não seja tão diferente quanto a de IA);
- d. 60% dos resultados de CLP indicam proximidade à média de LIS, enquanto 40%, de LIC. Sendo assim, a maior parte dos textos E apresenta vocabulários com concretude lexical similares aos dos textos simples. Entretanto, é importante destacar que um percentual de 40% de resultados próximos aos valores de LIC pode ser um indício para que possamos observar os textos E com cuidado no que diz respeito à concretude, conforme explicado anteriormente;
- e. 100% das indicações de TT enquadram os textos E como próximos aos textos simples do *corpus* comparativo, sugerindo que em termos de variedade lexical, os textos E estão apropriados.

Evidentemente, apenas essa aproximação não é suficiente para que possamos enfatizar que os textos são acessíveis. Esse procedimento metodológico comparativo se soma aos passos anteriores da nossa análise.

Por fim, nossa análise a respeito do léxico dos textos componentes do livro do nível *elementary* nos permite concluir que:

- I. A presença de palavras polissêmicas é relativamente ampla, ainda que seja preciso levar em consideração o fato de que múltiplas significações são geralmente evocadas por leitores mais avançados (uma vez que para conhecer diferentes significados de uma unidade lexical é preciso vocabulário mais extenso). Logo, ainda que a PS indique potenciais de complexidade, é necessário observar quais palavras de fato podem causar dualidade de interpretação (conforme explicado na nossa análise a respeito dessa métrica);
- II. A complexidade das palavras presentes nos textos do nível *Elementary* são adequadas de acordo com a métrica IA; ou seja, são unidades lexicais que são geralmente aprendidas em idades iniciais da infância. Ainda que seja importante destacar que esse índice é calculado tendo em vista um leitor nativo da língua inglesa; ou seja, é necessário que, para aprendizes de ILA, a métrica IA seja observada como ponto de partida (corroboração de aprendizado de palavras pode ser realizada por uma análise de vocabulários trabalhados anteriormente à unidade onde o texto é apresentado);
- III. A maior parte dos textos é escrito com palavras consideradas familiares (como indicado pela FPC), mas ainda há espaço para adoção de estratégias de simplificação no nível lexical, uma vez que existem palavras que podem ser substituídas e que não se adequam ao insumo "+1" para manutenção linguística por estarem muito afastados do nível para o qual são orientados pelo CEFR - como no caso da palavra *'however'* e de algumas unidades lexicais que estão presentes com conteúdo a serem ensinados em unidades muito distantes de onde se encontram no texto (como advérbios que poderiam ser substituídos por palavras mais simples, por exemplo *'nearly'* por *'almost'*).

- IV. Grande parte dos textos apresenta vocabulário que permite uma evocação de significado concreto, fato que facilita a compreensão uma vez que o leitor pode construir uma imagem de cada “conceito” com maior facilidade. Ainda que a métrica CLP aponte que existem textos de complexidade média e alta, podemos observar que muitas palavras abstratas não são conceitos que podem ser considerados fora da capacidade de um leitor adulto, mesmo que em níveis iniciais;
- V. A variedade lexical é relativamente ampla em determinados textos (destacados, principalmente, pela métrica TT), que pode significar maior complexidade, mas;
- VI. Ainda que a variedade lexical seja relativamente ampla, a presença de vocabulário vasto pode ser explicada pelo uso das palavras de controle a serem aprendidas pela unidade.

Em suma, acreditamos que o nível lexical dos textos analisados está adequado, mas que ainda existe a possibilidade de adequação de vocabulário (em algumas instâncias específicas) para evitar que o insumo para o aprendizado seja “+2” (um passo além da proposta de “+1” de Krashen (1987); característica que pode aumentar o filtro afetivo e prejudicar a progressão linguística. Dessa forma, professores que desejam tornar seus textos mais acessíveis no quesito léxico (seja para alunos com dificuldades ou para alunos que não sejam neurotípicos), podem utilizar a organização de etiquetas de complexidade da nossa tabela e os procedimentos de análise como apresentamos anteriormente a fim de encontrar possíveis elementos e características que possam ser potencialmente complexos.

Na próxima subseção, analisaremos as métricas indicativas de complexidade sintática dos textos do livro *Elementary*.

8.5.1.2 Análise da complexidade sintática dos textos E

Para completar nossa análise sobre os textos presentes no livro *Elementary*, passamos agora para a interpretação das métricas referentes à sintaxe, iniciando pela apresentação desses valores.

O Quadro 33 abaixo exhibe os valores das métricas de sintaxe dos textos do livro *Elementary*, cuja interpretação será conduzida a seguir.

Quadro 33 - Valores das métricas de sintaxe dos textos do livro *Elementary*

	SINTAXE				
	NT	SS	SIMS	CR	IF
E1	23,27	89,62	122	8,08	81,07
E2	84,61	88,3	205	24,2	87,31
E3	93,45	31,21	114	60,64	84,14
E4	98,21	55,17	131	87,08	84,1
E5	39,36	72,57	81	10,38	72,05
E6	89,44	87,49	155	28,43	92,92
E7	64,06	65,91	77	33,36	72,54
E8	42,47	22,06	116	68,79	72,3
E9	94,52	57,93	108	64,43	87,54
E10	93,7	91,47	133	30,15	91,37
E11	46,02	65,17	100	49,2	76,37
E12	24,2	33,36	79	54,38	58,47
E13	88,49	52,39	56	35,2	86,84
E14	83,89	93,57	99	15,87	93,44
E15	44,04	58,71	90	4,46	73,02
E16	90,99	62,17	94	11,9	83,91
E17	87,49	85,31	130	26,43	82,27
E18	73,89	62,93	80	5,94	73,75
E19	75,8	69,5	77	33	76,51
E20	87,29	67	123	12,3	81,12
MÉDIA	71,26	65,59	109	33,21	80,55

Fonte: Elaboração própria

Legenda: NT: Narratividade | SS: Simplicidade sintática | SIMS: Similaridade Sintática | CR: Coesão Referencial | IF: Índice Flesch

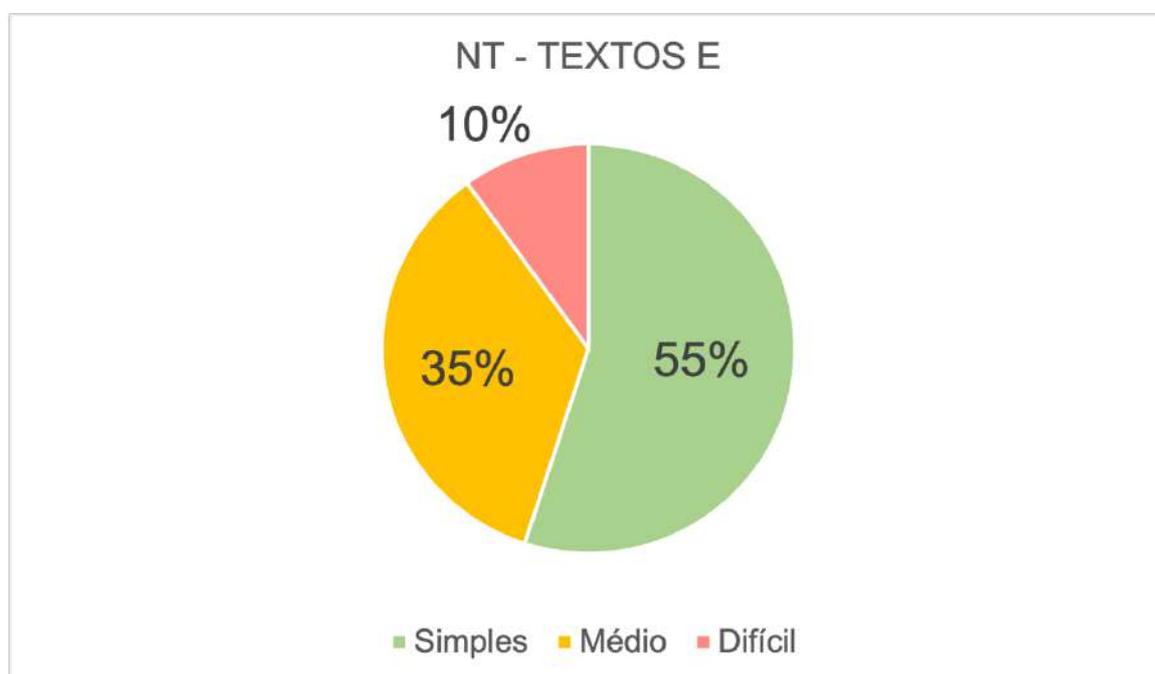
Para auxiliar a estimar o potencial de complexidade sintática dos nossos textos, iniciaremos a análise pela interpretação dos resultados da métrica NT, que verifica o quanto um texto é similar ou diferente de narrativas (tipologia textual frequentemente simples, conforme explicado no capítulo 6.2 *Métricas de Leitabilidade*).

8.5.1.2.1 Análise da métrica NT dos textos E

De acordo com orientações de complexidade textual - tanto das observações de McNamara et al. (2014) quanto de diversas outras referências sobre o tema - a leitura tende a ser mais simples em textos que são similares a narrativas, ou seja, com

estruturação que permita uma leitura que flua de forma parecida com uma história que tenha enredo, personagens, cenários etc. Como forma de estimar essa semelhança, utilizamos a métrica NT (conforme já apresentado na seção de métricas desta tese). Para que possamos analisar os resultados desse índice, iniciaremos pelo gráfico abaixo (Gráfico 18) que nos mostra o percentual de textos que podem ser classificados como simples, médios e difíceis por meio da nossa organização de valores de NT:

Gráfico 18 - Classificação percentual de complexidade de NT do livro *Elementary*



Fonte: Elaboração própria

Conforme podemos observar pelo gráfico acima, pouco mais do que a metade dos textos podem ser considerados altamente semelhantes a narrativas (55%) e, conseqüentemente, acessíveis, enquanto os demais se classificam entre complexidades médias (35%) e difíceis (10%). É importante lembrar que esses textos são apresentados no livro *Elementary* e, portanto, não deveriam apresentar índices que indiquem alta complexidade.

O texto considerado mais simples pela NT foi o E4, com índice de 98,21% de semelhança entre ele e uma narrativa (por, de fato, ser uma narrativa). A leitora pode lembrar que esse texto também indicou alta acessibilidade textual por métricas relacionadas ao léxico e à semântica, conforme comentado anteriormente.

Por ser uma história contada em primeira pessoa, o texto apresenta um desencadeamento lógico simples de ser seguido e estruturas com repetições de sujeitos e verbos. No decorrer da leitura, é possível imaginar, com facilidade, o enredo, o personagem e as atividades, como pode ser exemplificado pelo trecho inicial: “*I wake up at about 1.00 p.m., and the first thing I do is go outside. I live in Ibiza and I like having breakfast in the sun. I usually have fruit juice, eggs, fruit, and tea. I never drink coffee.*”

Conforme já demonstramos pelo Gráfico 18 apresentado acima, mais da metade dos textos obteve como resultado valores de NT considerados acessíveis, mas é muito interessante observar que esses resultados não são apenas parte da classificação “simples” da complexidade estimada pela métrica, mas são valores extremamente altos desse índice (com variação acima de 84%).

Para que possamos observar algumas características importantes, no quadro abaixo (Quadro 34) apresentamos trechos de textos com bons resultados de NT e tecemos alguns comentários sobre características que contribuíram pelo aumento desse índice:

Quadro 34 – Trechos de exemplo de textos com resultados de NT positivos do livro *Elementary*

TEXTO	TRECHOS	COMENTÁRIOS
E9 (NT: 94,52)	<i>First, we went to a really expensive restaurant for dinner. It was early, about 6.00, because that's when people have dinner in Iceland.</i>	Narrativa curta. Frequência alta de conectores temporais. Grande presença de verbos no passado e infinitivo.
E10 (NT: 93,7)	<i>The inspector came in. 'Mr Travers died between midnight last night and 7.00 this morning,' he said.</i>	História com diversos personagens. Sequência lógica clara. Presença de diálogos. Personagens e referência aos personagens evidentes.
E3 (NT: 93,45)	<i>I get up between 6.00 and 6.30 a.m., I get to work at 7.00, and my meetings usually start at 7.30. After that. I don't have a break. I have lunch at my desk.</i>	Texto em primeira pessoa. Sequência lógica clara. Marcadores temporais lógicos. Sujeitos claros e repetidos.
E16 (NT: 90,99)	<i>When I eat out with my husband or friends, we usually share, because food portions are huge and we don't want all that food.</i>	Narrativa dividida em tópicos. Frases e parágrafos curtos. Sujeitos claros.
E6 (NT: 89,44)	<i>They get up, and they have sweet black coffee for breakfast. Then Hassan works in</i>	Narrativa simples. História rica em detalhes. Sequência lógica clara.

	<i>the fields - but Walid just sits and looks at the hot, empty desert and the hot, empty sky.</i>	Presença de diálogos. Personagens e referência aos personagens evidentes.
E13 (NT: 88,49)	<i>The good thing about this plan for me is that I spend less time and money buying food, and I also never need to worry if my diet is healthy or not (because I know it is).</i>	Narrativa dividida entre terceira pessoa e primeira pessoa. Alta frequência de verbos.
E17 (NT: 87,49)	<i>During my studies, I worked in the Canary Islands and in Argentina and I studied for two semesters in the UK. After I finished, I decided that I didn't want to work for the rest of my life.</i>	Narrativa em primeira pessoa. Sequência lógica clara. Frequência alta de conectores temporais. Alta frequência de verbos.
E20 (NT: 87,29)	<i>We went yesterday for an early dinner with our three-year-old daughter and they really made us feel at home - and they produced a delicious birthday cake for her.</i>	Narrativa dividida em tópicos. Personagens claros. Frases e parágrafos curtos. Alta frequência de verbos.
E2 (NT: 84,61)	<i>In the UK women have a lot of freedom. In my country, if I want to go to the bank I take my brother and he does everything for me.</i>	Texto narrativo. Sujeitos claros. Alta presença de formação canônica. Sequência lógica clara. Frequência alta de verbos.

Fonte: Elaboração própria

Ao mesmo tempo que uma grande quantidade de textos resultou em valores altos de NT, também podemos observar textos que indicam possível complexidade por meio dessa métrica. O texto E1 (que, curiosamente, é o primeiro apresentado no livro) é classificado como tendo o resultado mais baixo de NT, 23,27%. Além dele, o texto E12 também apresenta um número considerado como indicativo de alta complexidade por essa métrica, 24,20%. Diante desses valores, é importante observarmos esses textos com mais atenção.

Ainda que o texto E1 tenha sido apontado como sendo o mais complexo pela NT ele é, na verdade, simples. A métrica pode ter considerado o texto como distante de uma narrativa porque ele é escrito em pequenas dicas no imperativo, sem que exista conexão temporal e, principalmente, com a omissão de um sujeito (“*Explore the city on foot. Don't rent a car in New York City.*”). O destaque do texto E1 ser considerado como mais complexo pela métrica em análise nos aponta, mais uma vez, a importância de realizarmos uma análise humana diante dos indícios das métricas.

O texto E12, por outro lado (segundo mais complexo por NT) apresenta alguns traços que podem, de fato, classificá-lo como complexo pela NT, dentre os quais:

relações não tão evidentes entre sujeito e predicado, grande presença de substantivos compostos nas posições de sujeito e de objeto e diversos deslocamentos de posições sintáticas canônicas, como no trecho “*In 1916, in her book *The Mysterious Affair at Styles, Agatha Christie introduced a new type of detective novel, and a new detective, Hercule Poirot.*”*

Então, ainda que tenhamos textos com valores de NT não tão altos, podemos perceber que – em alguns casos – outros fatores podem aumentar a acessibilidade desses textos. Grande parte dos textos apresentados no livro *Elementary* são verdadeiramente narrativas e, conseqüentemente, os valores desse índice tendem a ser altos.

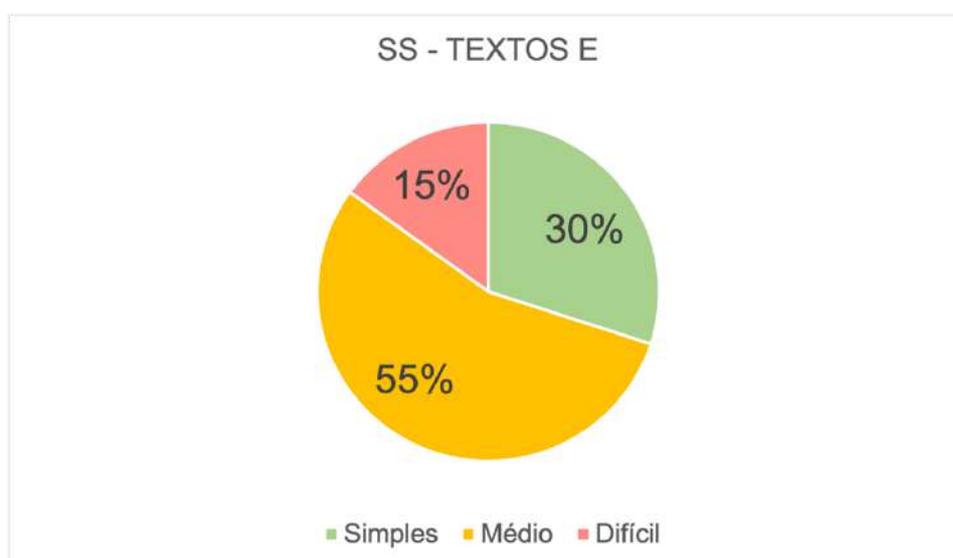
Existem estratégias que podem ser adotadas para aumentar os valores de NT (e conseqüentemente a acessibilidade textual), sobre os quais falaremos posteriormente.

Uma vez que apenas uma métrica não é suficiente para estimar complexidade em nenhum nível textual de forma acurada, também observaremos os apontamentos da métrica SS (Simplicidade Sintática).

8.5.1.2.2 Análise da métrica SS dos textos E

Iniciaremos a análise de SS pelo gráfico abaixo (Gráfico 19), que organiza os textos entre suas interpretações de complexidade:

Gráfico 19 - Classificação percentual de complexidade de SS do livro *Elementary*



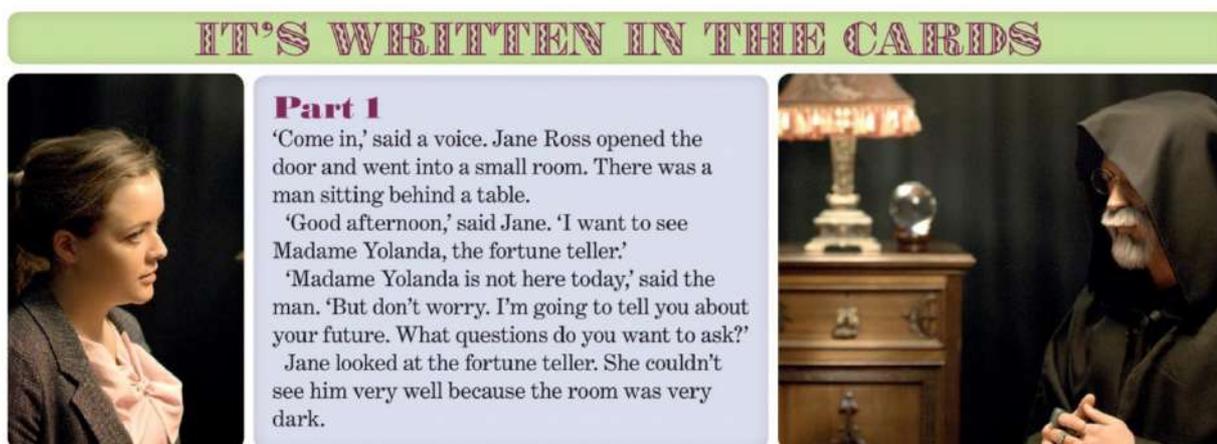
Fonte: Elaboração própria

O gráfico de SS nos demonstra uma pequena diferença em relação à NT. De acordo com essa métrica, 30% dos textos são considerados simples, 55%, médios e 15%, difíceis. Os valores de SS nos permitem observar que o texto mais simples é o E14 (SS: 93,57%) e o mais complexo, o E8 (SS: 22,06%).

Ao estudarmos o texto E14 (mais simples de acordo com a SS), podemos perceber que ele é composto em forma narrativa e organizado primordialmente em forma de diálogo. Essa história é escrita por meio de frases curtas, sem muitas variações sintáticas e sem deslocamentos de estruturas de suas posições canônicas. É possível observar constância entre a formação sintática das frases.

A imagem abaixo (Figura 25) é parte da história como apresentada pelo livro:

Figura 25 – Exemplo de SS no texto E14 do livro *Elementary*



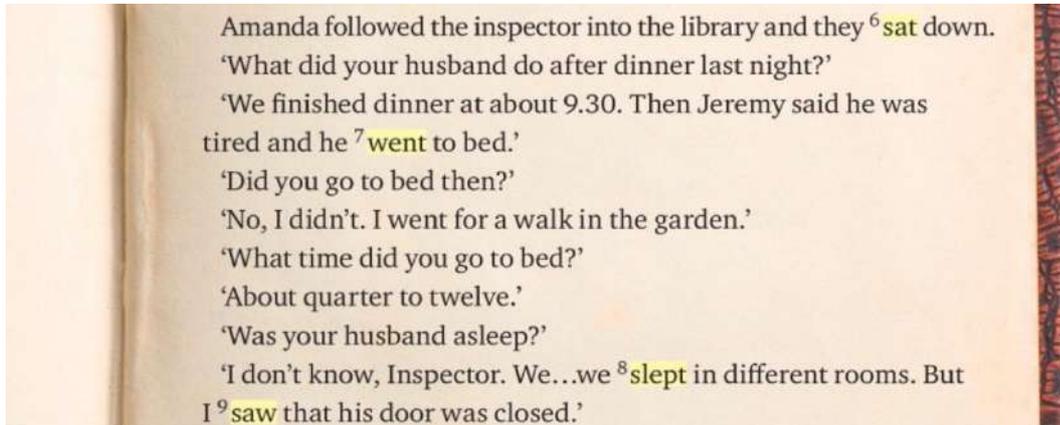
Fonte: Latham-Koenig, Oxenden e Lambert (2018, p. 83)

Como é possível perceber, os diálogos são curtos e simples de serem entendidos. Entretanto, é importante destacar que ainda que a formação sintática das frases não pareça ser um elemento com alto potencial de complexidade, a pontuação exibida pelo texto pode ser complexa para um aluno brasileiro, uma vez que as regras de uso de aspas e de pontuação se diferem da língua portuguesa. Além disso, outro detalhe que pode tornar a compreensão global mais difícil é o fato de que o texto é dividido em três partes, sendo a primeira e terceira partes apresentadas por escrito, enquanto a segunda, por áudio.

Outro texto que merece destaque é o E10, que foi classificado como o segundo mais simples pela SS (SS: 93,7%). Ele também é uma história narrativa com muitos diálogos e que apresenta uma sequência lógica visível e estruturas semelhantes.

Grande parte do texto é formada por perguntas e respostas diretas (como ilustrado pela Figura 26 abaixo), sem muitas formações sintáticas distintas e sem muitos deslocamentos de elementos de suas posições canônicas:

Figura 26 – Exemplo de SS no texto E10 do livro *Elementary*

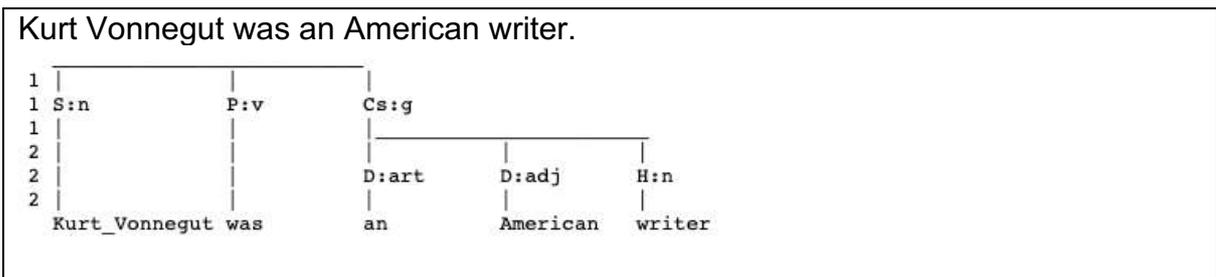


Fonte: Latham-Koenig, Oxenden e Lambert (2018, p. X)

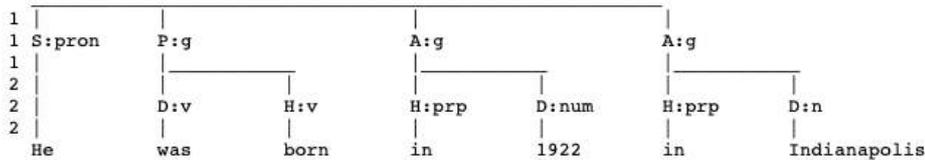
Por serem textos narrativos curtos, ambos textos apontados pela SS como mais simples adotam determinados padrões nas formações das frases e, ao realizarmos a análise sintática das frases, podemos determinar uma alta presença de estruturas que seguem a forma canônicas (sujeito – verbo – objeto), fato que torna a leitura mais simples, conforme indica Silva (2018), DuBay (2004) e outros pesquisadores da área.

O texto E8, por sua vez, apontado como sendo o mais complexo pela SS é uma narrativa biográfica com grande variação de estruturas. O texto é uma apresentação da história de três celebridades, e a formação sintática das frases varia não apenas entre cada personagem, mas também em cada parágrafo. Ao realizarmos a análise sintática das três primeiras frases do texto, como forma de ilustrar essa alta variação, observamos três formações distintas, conforme a imagem abaixo (Figura 27):

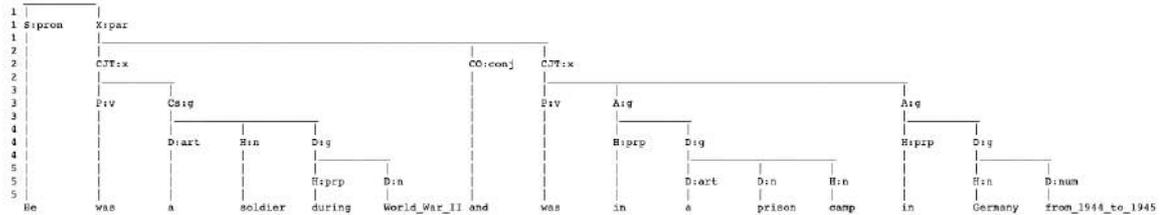
Figura 27 - Análise sintática das três primeiras frases do texto mais complexo por SS



He was born in 1922 in Indianapolis.



He was a soldier during World War II, and was in a prison camp in Germany from 1944 to 1945.



Fonte: Elaboração própria por meio da plataforma visl.sdu.dk

Como podemos ver, as formações das três frases são:

- Frase 1: Sujeito, verbo de ligação, complemento (predicativo do sujeito);
- Frase 2: Sujeito, verbo de ligação, complemento, adjunto, adjunto;
- Frase 3: Sujeito, verbo de ligação, complemento, adjunto, conjunção, verbo de ligação, adjunto, adjunto.

Além disso, ao realizarmos a leitura do texto tendo como objetivo analisar sua sintaxe, notamos que a formação sintática é inconstante em todas as frases do texto e que a quantidade de palavras que ocupam cada função também é altamente variável, justificando os valores baixos de SS.

O texto E3, segundo mais complexo por SS, também apresenta estruturas lexicais distintas, frequentes deslocamentos de elementos, grande presença de adjuntos e de frases relativas. O fragmento a seguir ilustra o constante deslocamento de funções sintáticas: "After that, I don't have a break. I have lunch at my desk. I often do two operations a day, and I also have lectures and more meetings. At home, I have dinner with my son. Between 9.30 and 11.30 p.m., I do research and I watch the news on TV."

Em suma, poucos são os textos cujos valores de SS indiquem altíssima complexidade. É válido lembrarmos, conforme já explicamos, que a natureza dinâmica de uma língua permite diversas configurações sintáticas. Esse fato ocorre, principalmente, em narrativas, posto que o deslocamento de elementos de suas

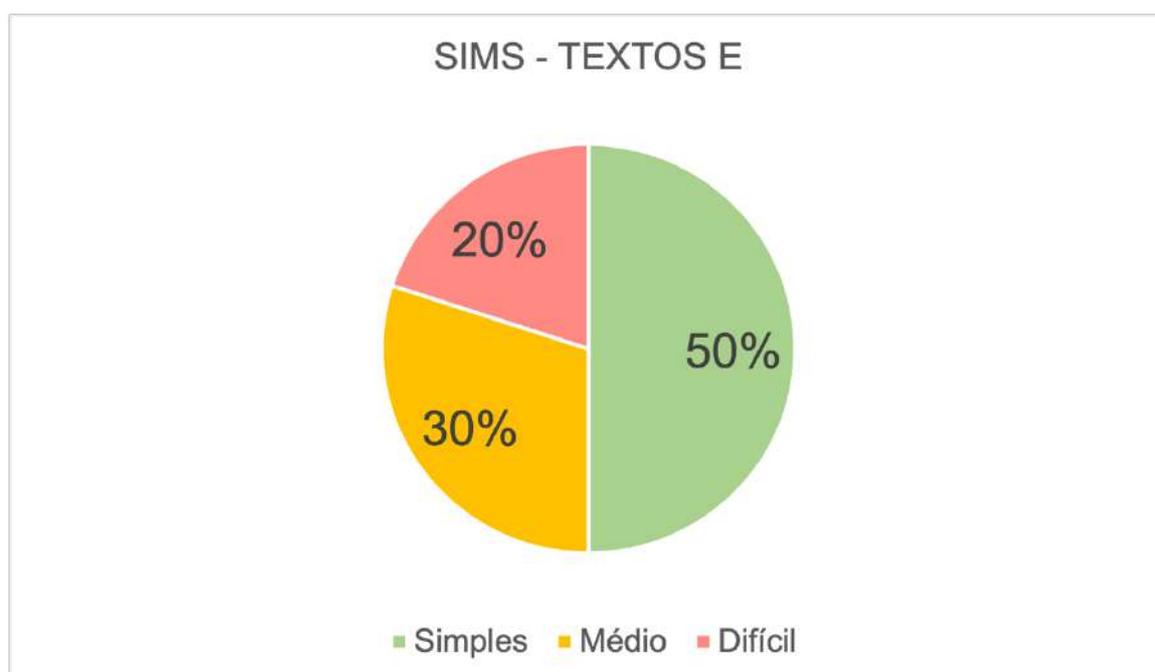
funções canônicas é frequente (por diversas finalidades) e que é muito comum termos distintas formas de coordenação e subordinação de frases.

Para entendermos ainda mais os potenciais de complexidade da sintaxe dos textos do livro *Elementary*, nossa análise se direciona para a observação dos valores de Similaridade Sintática (SS).

8.5.1.2.3 Análise da métrica SIMS dos textos E

O gráfico abaixo (Gráfico 20), classifica quantos textos resultaram em valores que os classificam como simples, médios ou difíceis para a SIMS:

Gráfico 20 - Classificação percentual de complexidade de SIMS do livro *Elementary*



Fonte: Elaboração própria

Ao analisarmos os valores de SIMS por meio da nossa classificação de complexidade, relatamos que 50% dos textos são considerados simples, 30% médios e 20% complexos (valor alto para textos do livro *Elementary*). O texto que indica menor complexidade por essa métrica é o E2 (SIMS: 205) e o mais complexo, o E13 (SIMS: 56).

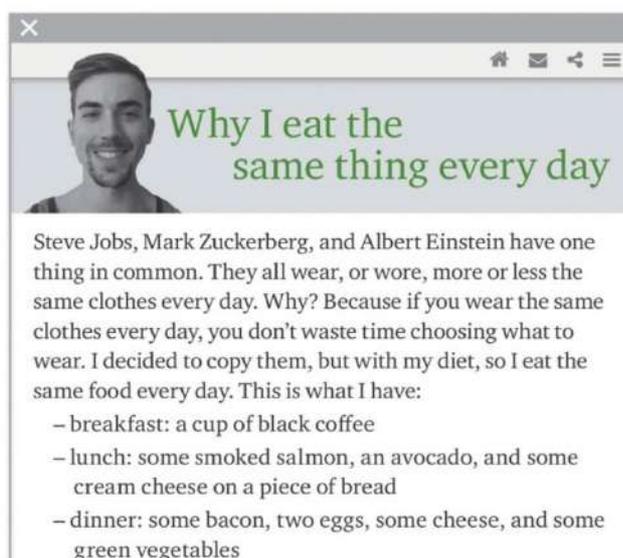
Mesmo sem realizar uma análise sintática do texto E2, podemos perceber a similaridade sintática entre as frases muito claramente. O texto conta pequenas histórias sobre cinco personagens e suas impressões sobre o Reino Unido. A primeira

frase de cada personagem é uma repetição sintática perfeita (“*He lives in Newcastle. She lives in Liverpool. She lives in Chorlton. She lives in London. She lives in Manchester.*”). Grande parte das frases seguem a mesma estrutura sintática (sujeito, verbo, objeto) e os poucos deslocamentos de elementos de suas posições canônicas também tendem a se repetir, como entre as frases: “*In my country, if I want to go to the bank I take my brother and he does everything for me. Here I do everything I like the weather here.*” onde podemos perceber que as palavras que exercem a função de adjunto foram deslocadas para a mesma posição (início da frase), ainda que existam inadequações em pontuações no fragmento utilizado como exemplo.

Outro texto que merece destaque pelo valor de SIMS é o E6 (SIMS: 155), que é uma narrativa relativamente longa (em comparação aos demais textos), mas com pouca variação nas formações sintáticas. Por ser uma narrativa, a presença de estruturações em formações canônicas (SVO) é alta, como no fragmento “*They have no money and they have no animals. Their house only has two rooms.*”. Quando existem deslocamentos, a posição também é repetida (como no posicionamento de adjuntos no início da frase), como no exemplo “*At night, it's very cold. (...) The next day, Hassan says goodbye. (...) The next night, Walid makes a plan.*”.

Por outro lado, o texto E13, considerado o mais complexo pela métrica SIMS, apresenta uma característica peculiar: ele começa por um parágrafo, é interrompido por uma lista e retoma para a formação de parágrafo, como podemos observar pela imagem abaixo (Figura 28):

Figura 28 – Texto mais complexo por SIMS do livro *Elementary*



The good thing about this plan for me is that I spend less time and money buying food, and I also never need to worry if my diet is healthy or not (because I know it is). Now, you're probably thinking, 'he's sure to get tired of eating the same food every day'. Well, after about a year, I can promise you that I'm not. I change my meals a bit by having different kinds of cheese and using different herbs and spices. I also change the vegetables that I have for dinner.

Adapted from a website

Fonte: Latham-Koenig, Oxenden e Lambert (2018, p. 70)

Além de formações sintáticas distintas, a lista no texto E13 não tem sujeitos e as frases não são orações, características que influenciam nos valores de SIMS. Além das diferentes configurações sintáticas, a quantidade de palavras que ocupa cada função também é bastante distinta. Na frase “*Steve Jobs, Mark Zuckerberg, and Albert Einstein have one thing in common.*” podemos computar seis substantivos próprios na posição de sujeito, enquanto na frase “*The good thing about this plan for me is that I spend less time and money buying food...*” temos uma frase com oito palavras como sujeito.

O texto E19 (segundo mais complexo por SIMS) é redigido na forma de perguntas e respostas. Uma vez que as respostas direcionam a formação das perguntas, é natural que a similaridade sintática entre as frases seja menor, especialmente porque há grande variação na quantidade de palavras que exercem cada função, alta frequência de deslocamentos de elementos de suas posições canônicas e muitas ocorrências de apostos. Além disso, existem frases que não são orações (“*Perhaps William Shakespeare.*”) e outras que utilizam o infinitivo como substantivos (“*e.g. to sing well, to play the piano, and to speak foreign languages.*”) que geram configurações sintáticas distintas e únicas – aumentando a quantidade de estruturas diferentes.

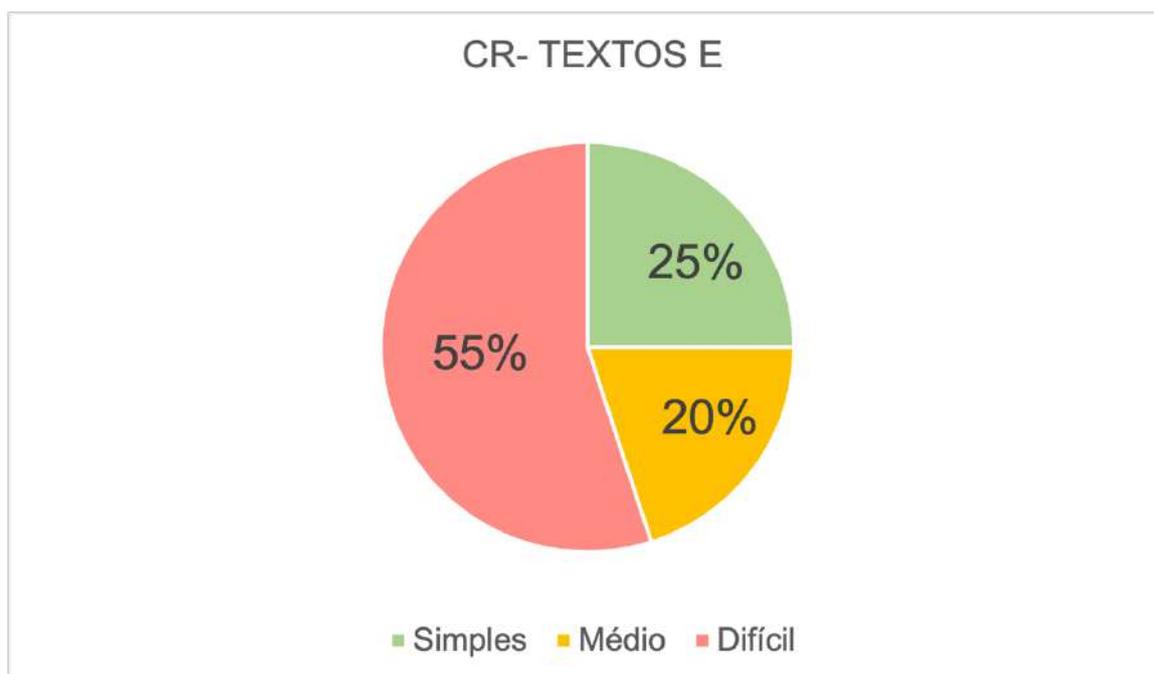
Com a finalidade de observarmos outra característica sintática que pode ser potencialmente complexa para nossos alunos, passamos agora para nossas observações sobre a métrica CR.

8.5.1.2.4 Análise da métrica CR dos textos E

Ao analisarmos o percentual de classificação de complexidade de acordo com a métrica CR, o gráfico abaixo (Gráfico 21) nos auxilia a observar que (de forma

contrária às demais métricas de sintaxe) a maior parte dos textos são classificados como difíceis (55%), enquanto 20% como médios e 25%, simples. Para que possamos entender esses valores, vamos analisar os textos com melhores e piores resultados de acordo com esse índice.

Gráfico 21 - Classificação percentual de complexidade de CR do livro *Elementary*



Fonte: Elaboração própria

Tendo um resultado de 87,08% de CR, o texto E4 é considerado o texto mais simples pela métrica. Ao retornarmos ao início da nossa análise sobre a complexidade sintática dos textos do livro *Elementary*, podemos lembrar que esse texto também foi considerado como o mais acessível de acordo com as classificações de NT. A leitura do texto nos permite observar com facilidade a razão pela qual a coesão referencial é tão alta: a maior parte do texto tem como sujeito a primeira pessoa, ou seja, o uso pronominal é dêitico, logo, sem necessidade de retomada de substantivos por meio de um pronome. A imagem abaixo (Figura 29) apresenta o texto na íntegra, cujo título é “*A Life in the Day – The Superstar DJ David Guetta*”:

Figura 29 – Texto mais simples por CR do livro *Elementary*

I wake up at about 1.00 p.m., and the first thing I do is go outside. I live in Ibiza and I like having breakfast in the sun. I usually have fruit juice, eggs, fruit, and tea. I never drink coffee. After breakfast, I answer my emails for an hour, then I go to the gym.

I never listen to music in the house, or even in the car, because music is my job. On a typical day I spend two or three hours in my studio, then another four hours at a nightclub. My work starts in the evening. I usually have dinner in a restaurant, and then I go to the club. I try to have a normal life, but my job isn't normal. I arrive at a club like a secret agent – I go in through the back door and Security takes me to the stage.

I finish work at 4.00 in the morning. Security takes me out, and then I go home. After about four hours playing music I'm very excited. My manager says, 'Go home and sleep', but that's impossible. First I need to calm down. When I get home I have a cup of tea, brush my teeth and say, 'Thank you for this wonderful life'. I am 47 now, but I want to do this when I'm 60 or 80. I want to do this forever.

Adapted from the British press

Fonte: Latham-Koenig, Oxenden e Lambert (2018, p. 37)

Como é possível observar, o uso do pronome pessoal em primeira pessoa “força” a referenciação na instância (ativado pela relação lógica e mental entre o tipo de texto – narrativa em primeira pessoa – e o nome do narrador). Poucas são as instâncias em que o sujeito é alterado (como nos casos de ‘*Security*’ e ‘*My manager*’) e, quando ocorrem, não há continuidade e uso de outras unidades lexicais para referenciação a essas pessoas.

O segundo texto com melhor valor de CR apresentou quase 20% de distância entre o primeiro. Com um percentual de 68,79%, o texto E8 é indicado pela métrica como o segundo mais acessível. Esse texto havia sido apontado pela métrica SS como o mais complexo. Entretanto, ainda que o texto possa apresentar grandes variações nas estruturações sintáticas, é possível que a coesão referencial seja relativamente adequada. Ao analisarmos o texto, podemos perceber que: I. A

quantidade de pronomes é alta e repetitiva; II. Os pronomes não estão muito distantes de suas referências; III. Há uso de outras palavras (adjetivos e substantivos) que também fazem referência aos sujeitos; IV. O pronome 'he' é utilizado para fazer referência a duas pessoas; V. Quando há troca de sujeitos no mesmo parágrafo, o pronome é diferente e, conseqüentemente, não apresenta dificuldade na seleção do referencial correto; VI. A concordância verbal também auxilia na referenciação, embora existam erros de pontuação.

Os fragmentos abaixo, onde destacamos em negrito os pronomes e em sublinhado os referentes, auxiliam-nos a ver tais observações:

Fragmento 1:

*Kurt Vonnegut was an American writer. **He** was born in 1922 in Indianapolis. **He** was a soldier during World War II, and was in a prison camp in Germany from 1944 to 1945. **His** first novel was Piano Player, published in 1952, but **it** wasn't a success.*

Fragmento 2:

*Adele is a British singer and songwriter. **She** was born Adele Laurie Blue Adkins in London in 1988. When **she** was a child **she** was much more interested in music than in schoolwork.*

Com base nos fragmentos do texto E8 apresentados acima e nas observações supracitadas, percebemos que, de fato, a coesão referencial (seja entre pronomes e substantivos, ou entre substantivos e adjetivos/locuções adjetivas, ou entre sujeito e verbo) é alta. Na realidade, a análise holística (e a experiência em sala de aula com o texto) indica que embora o valor de CR tenha sido 68,79%, os alunos não costumam apresentar dificuldades em encontrar a coesão entre as referências nesse material.

Já os textos E15 e E18, por outro lado, apresentam valores extremamente baixos de CR, tendo o primeiro 4,46% e o segundo, 5,94%. Ao observarmos o primeiro texto, percebemos que a relação entre palavras não ocorre apenas por pronomes pessoais, mas também por pronomes relativos. Além disso, o texto apresenta muitas instâncias de generalização (como na utilização de 'people', 'you' e 'guests') e alternância constante entre sujeitos (diferenciando, na mesma frase, não apenas

pronomes, mas também concordância verbal, como podemos perceber pelo fragmento abaixo:

“Would you like to sleep in a mine? Then book the underground suite in Sala Silvermine in Sweden, and enjoy the world's deepest bedroom, 155 metres below the surface. Although the corridors are cold and dark, your rooms are warm (18°C), and the light comes from candles in beautiful silver candlesticks. You can also explore the caves and magical lakes with a guide. There are no other guests, so it's not for nervous people. Mobile phones don't work, but you have a radio for emergencies. In the morning they bring breakfast down to you. Perfect for romantic couples who like a bit of adventure.”

Na primeira frase, o sujeito é referenciado por ‘you’ (que é um pronome dêitico e se ativa somente na instanciação). Em sequência, por serem frases no imperativo, o sujeito (que embora continue sendo ‘you’) é omissso. Logo, temos mudanças de sujeitos ‘the corridors’, ‘your room’ e ‘the light’ seguido de uma nova frase com ‘There are’ e, logo em seguida, ‘it's not for nervous people’. Nesse exemplo, podemos perceber que alguns pronomes apontam para algo que não está necessariamente no texto (como o ‘it’) na frase anterior, e que nem sempre os sujeitos são facilmente identificados.

Fenômenos similares ocorrem no texto E18: existe uma generalização muito frequente por meio de palavras como ‘somebody’, ‘people’, ‘you’ e grandes variações de sujeito com a utilização de pronomes que podem causar dificuldade em referenciação, como podemos perceber pelo fragmento abaixo:

“Sending an email is so easy these days, but some years ago, when you needed to send something to somebody urgently and the post was too slow, you sent a fax. When they were introduced in the 1970s people thought they were a technological miracle. You can print something and send it over the phone in seconds! Magic!”

Apenas nesse parágrafo, podemos verificar um exemplo de presente participio sendo utilizado como sujeito da frase (‘sending an email’), generalização (‘you’ e ‘people’), utilização do pronome ‘they’ para retomar um substantivo (‘fax’) em uma oração anterior (não na mesma oração) que, além da referência ser feita entre duas

orações, o pronome pessoal *'they'*, utilizado para plural, está retomando uma ideia apresentada na forma singular (*'a fax'*). Ainda que leitores proficientes possam compreender que o pronome se refira às máquinas de fax, não podemos inferir que um aluno em nível *elementary* também seria facilmente capaz de fazer tal relação.

Outro fenômeno interessante de ser observado no parágrafo acima do texto E18 é a utilização do pronome *'something'* que exige a referência a uma ideia não mencionada na frase ou parágrafo, uma vez que essa palavra “aponta” para algo não especificado. Além disso, na continuação da frase, o pronome pessoal *'it'* se refere à *'something'* que, conseqüentemente, exige do leitor uma referência dupla: entre o *'it'* e o *'something'* e entre o *'something'* e algo não especificado.

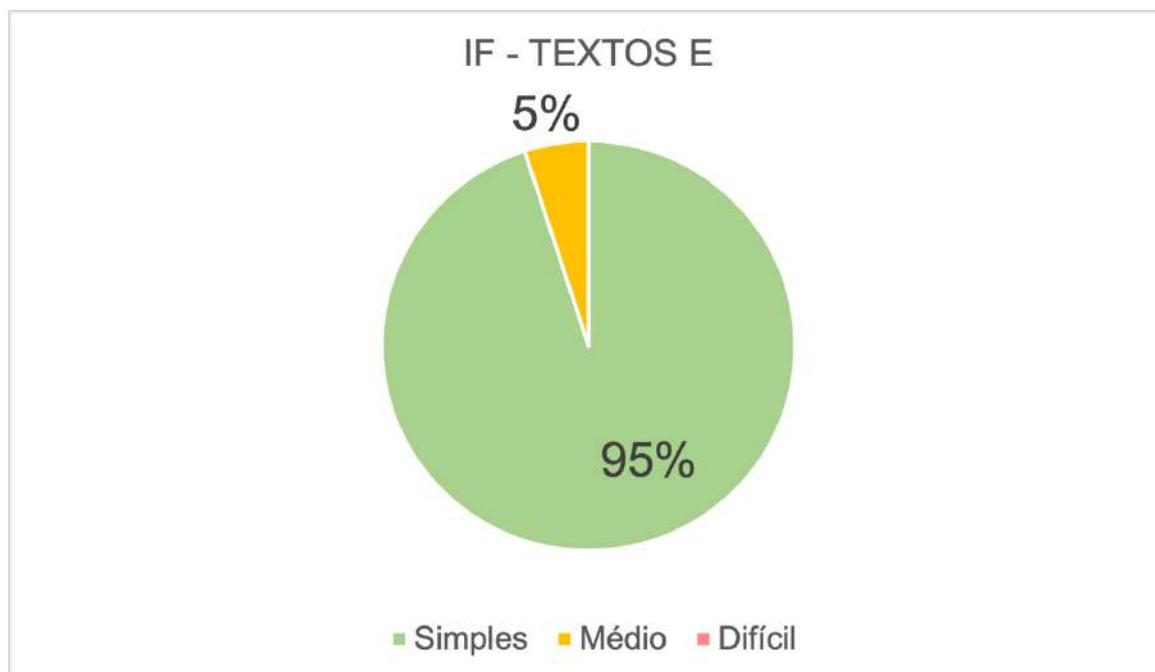
Embora os textos E15 e E18 não tenham sido os únicos com valores baixos de CR, os exemplos apresentados acima servem de ilustração para que possamos encontrar ocorrências onde as referências (seja entre pronomes e substantivos ou entre significantes e significados) podem não ser tão claras. Sendo assim, diante dos destaques em vermelho na tabela cujos valores métricos dos textos foram apresentados, um professor pode atentar para a adequação da coesão referencial quando julgar necessário.

Por fim, ainda que não seja uma métrica estritamente de sintaxe, vamos analisar uma métrica muito utilizada para análises de complexidade textual, o IF.

8.5.1.2.5 Análise da métrica IF dos textos E

É importante iniciarmos nossa análise de IF com o lembrete de que essa métrica estima complexidade com base na extensão das palavras e das frases. Ainda que essa seja a única métrica presente nesta tese que tenha sido originalmente desenvolvida com a finalidade de estimar complexidade, autores e pesquisadores de AT enfatizam que pesquisas que tenham por objetivo estimar complexidade não devem utilizar somente o IF, uma vez que, como já mencionado, um texto pode ser considerado menos acessível não só por sua extensão, mas por outras diversas características.

O gráfico abaixo (Gráfico 22) – que organiza a complexidade dos textos por meio das orientações de IF – demonstra que 95% dos textos podem ser considerados como fáceis, enquanto apenas 5% como médios; sem nenhuma ocorrência de textos complexos.

Gráfico 22 - Classificação percentual de complexidade de IF do livro *Elementary*

Fonte: Elaboração própria

Ainda que, conforme supracitado, o IF não possa ser utilizado como índice com total respaldo para estimar complexidade dos textos, ele serve como sinalizador de extensão. Logo, tendo em mente que adequações para fins de acessibilidade devem sempre ser orientadas pelo leitor para quem apresentaremos o texto, diante de indicações de alta complexidade pelo IF quando os textos são escritos para alunos de níveis básicos, por exemplo, podemos compreender que esses textos podem ter palavras muito extensas ou frases muito longas e, diante dessa corroboração, adotar estratégias que sejam coerentes com dimensões de léxico ou de estruturas frasais.

O texto mais simples por IF foi o E14 (93,44) e, ainda que não tenhamos resultados de textos considerados difíceis, o E12 apresenta o valor mais baixo de IF, 58,47 que, por meio das orientações da métrica, é sugestivo de complexidade média.

O quadro abaixo (Quadro 35) nos mostra alguns valores⁶⁵ relacionados à extensão dos textos com maior (E14) e menor (E12) resultados de IF. Além deles, apresentamos também essas informações sobre o segundo texto com maior resultado (E6, IF: 92,92) e o segundo texto com valor mais baixo (E5: 72,05) que, embora tenha essa classificação, ainda indica um valor relativo à alta acessibilidade:

⁶⁵ Por meio do *software* Word, podemos obter informações de médias de palavras, frases e caracteres para todos os textos.

Quadro 35 – Valores relacionados à extensão de textos do *Elementary* destacados por IF

	E14	E6	E12	E5
Palavras	242	439	276	358
Caracteres	1.040	1.890	1.385	1.684
Parágrafos	22	14	5	9
Frases	35	50	14	23
Média de frases por parágrafo	1,8	4,1	3,5	4,6
Média de palavras por frase	6,2	8,6	19,2	14,3
Média de caracteres por palavra	3,9	4,0	4,8	4,5

Fonte: Elaboração própria

Como é possível observar por meio dos números do quadro acima, existe uma grande distinção em extensão entre os textos considerados mais simples por IF e os dois mais complexos, principalmente na média de palavras por frase e na média de caracteres por palavra. Os textos mais simples têm médias de palavras por abaixo de nove e de caracteres por palavras inferiores ou iguais a quatro. Os textos mais complexos, por sua vez, apresentam médias superiores a 14 de palavras por frase e iguais e superiores a 4,5 de caracteres por palavras.

Ao observarmos a quantidade de caracteres e palavras, observamos que o segundo texto mais simples, E6, tem maiores quantidades do que o texto estimado como tendo complexidade média por IF. Entretanto, ainda que o E6 tenha mais palavras e caracteres de forma geral, o tamanho das palavras e a quantidade delas por frase é menor, fato que caracteriza esse material como sendo de baixa complexidade.

No quadro abaixo (Quadro 36), apresentamos um parágrafo de cada texto para que possamos ter um apoio visual das extensões mencionadas anteriormente:

Quadro 36 – Trechos de textos do *Elementary* destacados por IF

E14 (IF: 93,44)	E6 (IF: 92,92)	E12 (IF: 58,47)	E5 (IF: 72,05)
<i>'No. Jim is a man I met at a party last month. He's an actor, from New York. He says he's in love with me. It was his idea for me to come to Madame Yolanda.'</i>	<i>Hassan walks for five days through the mountains. The sun shines. At night, it's very cold. Then, one evening, he arrives at a rich and beautiful palace.</i>	<i>Father Brown was also a detective who was nearly as popular as Holmes; he was created by G.K. Chesterton (1874-1936). Father Brown only appeared in short stories, and Sherlock Holmes appeared in over 50 short stories and</i>	<i>One of Europe's biggest marathons takes place in spring every year, usually in April. About 40,000 people, including top runners and celebrities, run 26.2 miles through the city.</i>

		only four novels, including <i>The Hound of the Baskervilles</i> , where he solves the mystery of an enormous ghost dog.	
--	--	--	--

Fonte: Elaboração própria

Os parágrafos acima – que refletem com fidedignidade o restante de cada texto no que diz respeito às extensões – contribuiu para que possamos entender, na prática, os valores obtidos de IF. Os textos mais simples são formados por frases mais curtas ('No.', 'The sun shines.') e, também, por palavras menores ('man', 'met', 'he', 'says', 'love', 'one', 'cold'). Além disso, até mesmo os nomes próprios utilizados nas narrativas são curtos ('Jim' e 'Hassan').

Por outro lado, nos textos mais complexos, podemos ver não apenas maiores quantidades de palavras por frases (incluindo exemplos de coordenação e subordinação), mas também utilização de unidades lexicais mais extensas ('detective', 'appeared', 'including', 'enormous', 'marathons', 'celebrities'). Diferentemente dos textos E14 e E6, os textos E12 e E15 apresentam nomes próprios ou títulos mais longos e compostos por duas ou mais palavras ('*Sherlock Holmes*', '*The Hound of the Baskervilles*' '*Treetop Walkway*' – que não está no exemplo, mas no texto original). Evidentemente, a configuração ortográfica dos nomes próprios indica que tais unidades não são palavras do cotidiano do léxico em língua inglesa com sinônimos em português e, conseqüentemente, poderiam ser excluídas da análise relativa à complexidade por meio de extensão. Títulos de livros (como no texto E12) e nomes de locais (E5), por outro lado, são compostos por palavras do cotidiano e podem aumentar ou diminuir a acessibilidade às informações.

Para concluir a análise a respeito dos textos do livro *Elementary*, passamos para a observação comparativa entre as métricas de léxico do texto E e os valores dos textos LIS e LIC.

8.5.1.2.7 Análise contrastiva das métricas de sintaxe de E com os de LIS e LIC

Apresentamos agora, conforme fizemos com as métricas de léxico e semântica, os valores dos índices de sintaxe dos textos E comparados às médias das métricas

dos textos LIS e LIC. Lembrando que destacamos em verde os valores mais próximos aos do texto LIS (ou seja, potencialmente mais similares aos textos simples e, conseqüentemente, com maior acessibilidade) e em laranja os resultados mais próximos aos de LIC (menos acessíveis). O quadro abaixo (Quadro 37) apresenta e classifica as métricas de E em comparação às médias de LIS e de LIC:

Quadro 37 - Comparativo de métricas de sintaxe entre os textos do *corpus* comparativo e os textos do livro *Elementary*

	NT LIS	NT E	NT LIC	SS LIS	SS E	SS LIC	SIMS LIS	SIMS E	SIMS LIC
E1		23,27			89,62			122	
E2		84,61			88,3			205	
E3		93,45			31,21			114	
E4		98,21			55,17			131	
E5		39,36			72,57			81	
E6		89,44			87,49			155	
E7		64,06			65,91			77	
E8		42,47			22,06			116	
E9		94,52			57,93			108	
E10	Valor médio	93,7	Valor médio	Valor médio	91,47	Valor médio	Valor médio	133	Valor médio
E11	NT:	46,02	NT:	SS:	65,17	SS:	SIMS:	100	SIMS:
E12	72,82	24,2	33,43	75,59	33,36	46,10	134	79	79
E13		88,49			52,39			56	
E14		83,89			93,57			99	
E15		44,04			58,71			90	
E16		90,99			62,17			94	
E17		87,49			85,31			130	
E18		73,89			62,93			80	
E19		75,8			69,5			77	
E20		87,29			67			123	
Média		71,26			65,59			109	

	CR LIS	CR E	CR LIC	IF LIS	IF E	IF LIC
E1		8,08			81,07	
E2		24,2			87,31	
E3		60,64			84,14	
E4		87,08			84,1	
E5		10,38			72,05	
E6		28,43			92,92	
E7		33,36			72,54	
E8		68,79			72,3	
E9		64,43			87,54	
E10		30,15			91,37	
E11	Valor médio	49,2	Valor médio	Valor médio	76,37	Valor médio
E12	CR:	54,38	CR:	IF:	58,47	IF:
E13	51,98	35,2	26,68	85,17	86,84	44,17
E14		15,87			93,44	
E15		4,46			73,02	

E16		11,9			83,91	
E17		26,43			82,27	
E18		5,94			73,75	
E19		33			76,51	
E20		12,3			81,12	
Média		33,21			80,55	

Fonte: Elaboração própria

O Quadro 37 acima nos auxilia a visualizar se as métricas de sintaxe dos textos do livro Elementary se aproximam mais dos valores dos textos simples ou dos textos complexos do *corpus* comparativo. Visualmente, podemos perceber que o número de métricas marcadas em verde é superior às marcadas em laranja. Sendo assim, já podemos perceber que os índices de E estão mais próximos aos valores dos textos simples.

Contudo, ainda podemos perceber que, se compararmos o Quadro 37 ao Quadro 34 (que comparava os valores das métricas de léxico e semântica dos textos E com os textos LIS e LIC), os índices relativos à sintaxe parecem sugerir maior complexidade nesse aspecto textual do que a respeito de léxico e semântica.

A interpretação dos valores exibidos acima nos indica que:

a. 60% dos textos E apresentam valores de NT próximos ao texto simples (sendo que, ao analisarmos os valores de NT dos textos LIS podemos constatar que essa métrica é ainda mais alta para os textos E) e 30% aos resultados de LIC. A média de NT dos textos E também se aproxima dos valores de LIS, indicando que, por essa métrica, os textos podem ser considerados mais simples do que complexos;

b. 65% dos textos demonstram resultados de SS próximos à média de LIS, enquanto 35%, de LIC. Esses percentuais nos indicam que ainda que a estruturação sintática seja, em grande maioria, mais próxima aos textos simples, existem características que em determinados textos se assemelham mais à composição dos textos complexos – indício de que essa característica sintática pode requerer atenção quando tivermos por objetivo deixar o texto mais acessível, conforme mencionado na análise dessa métrica.

c. A métrica SIMS também serve de alerta a respeito do possível potencial de complexidade sintática dos textos E, visto que 50% dos valores dessa métrica são próximas à média de LIS e os outros 50%, de LIC. A média dos valores de SIMS dos textos E também se aproxima mais aos valores de LIS, mas a diferença entre os

valores não é muito grande. Além dos dados percentuais a respeito de SS (tópico b), os valores percentuais de SIMS também nos indicam necessidade de atenção aos aspectos sintáticos do texto, levando em consideração todas as observações apresentadas na interpretação dessa métrica;

d. Os valores de CR dos textos E indicam que a coesão referencial pode ser considerada o maior indício de potencial de complexidade desses textos, em razão de 70% dos valores dessa métrica serem mais próximos dos textos complexos, enquanto apenas 30% dos textos mais simples. Além disso, a média de CR dos textos E é muito mais próxima à média de LIC do que de LIS. Esses dados também nos alertam que, com o propósito de potencializar a acessibilidade dos textos, coesão é uma característica que precisa ser observada (conforme já discutido na análise da métrica);

e. Os resultados de IF, por outro lado, demonstram que em relação à extensão das palavras e frases dos textos, 95% dos textos E são similares aos textos mais simples. A média de IF dos textos E também é altamente similar à média dos textos LIS;

Em síntese, podemos observar e concluir que entre léxico, semântica e sintaxe, a última parece ser o elemento de maiores indícios de complexidade textual para os textos E.

Nossa análise nos demonstra que fatos como deslocamentos de elementos de suas posições sintáticas, distância pronominal e coesão referencial são algumas características que podem tornar os textos do livro *Elementary* potencialmente complexos para alunos desse nível. Além disso, é preciso destacar que os textos da coleção *English File* (não somente do nível elementar) tendem a apresentar constantes falhas em termos de pontuação – característica que merece revisão uma vez que, diante de palavras desconhecidas para os alunos, a pontuação pode ser o único indício de deslocamento de elementos e, dessa forma, contribuir (ou prejudicar) a compreensão da função que as palavras exercem.

Ao finalizarmos nossa análise de complexidade dos textos do livro *Elementary*, seguimos, na próxima seção, para o estudo das métricas do texto *Pre-intermediate*.

8.5.2 RESULTADOS MÉTRICOS DOS TEXTOS DO LIVRO DE NÍVEL PRE-INTERMEDIATE

Seguindo a mesma ordem da análise previamente realizada a respeito dos textos do livro *Elementary*, iniciaremos nosso estudo das métricas dos textos do livro *Pre-intermediate* pelos índices relativos ao léxico e à semântica.

8.5.2.1 Análise da complexidade lexical-semântica dos textos P

O quadro abaixo (Quadro 38) apresenta os valores das métricas de semântica e léxico dos textos P (*Pre-intermediate*) e as classificações de etiqueta de cores representativas de complexidade, conforme explicado no início do capítulo 8 desta tese:

Quadro 38 - Valores das métricas de semântica e léxico dos textos do livro *Pre-intermediate*

	SEMÂNTICA E LÉXICO				
	PS	IA	FPC	CLP	TT
P1	3,51	264,05	593,82	70,19	785
P2	4,71	248,83	590,67	42,86	706
P3	4,06	285,57	587,84	90,32	795
P4	5,09	255,13	592,56	33	730
P5	4,62	310,24	579,35	78,23	690
P6	4,3	300,31	588,93	75,17	705
P7	4,18	361,96	578,26	35,94	712
P8	4,28	296,91	584,38	46,81	823
P9	4,39	307,66	585,58	93,82	685
P10	4,39	302,85	591,47	37,45	765
P11	3,99	343,27	571,4	68,44	750
P12	4,17	337,44	587,03	28,43	764
P13	3,83	271,45	593,12	19,22	684
P14	3,8	313	586,05	5,05	643
P15	3,2	304,76	575,15	89,97	808
P16	3,85	295,47	584,05	94,74	684
P17	3,7	298,41	566,95	76,42	846
P18	4,06	307,44	584,5	42,07	739
P19	3,49	262,62	590,67	30,15	725
P20	4,26	268,33	583,53	89,97	637

P21	4,54	269,15	585,06	83,4	669
P22	4,23	300,27	571,78	81,86	774
MÉDIA	4,12	295,69	584,19	59,71	733

Fonte: Elaboração própria

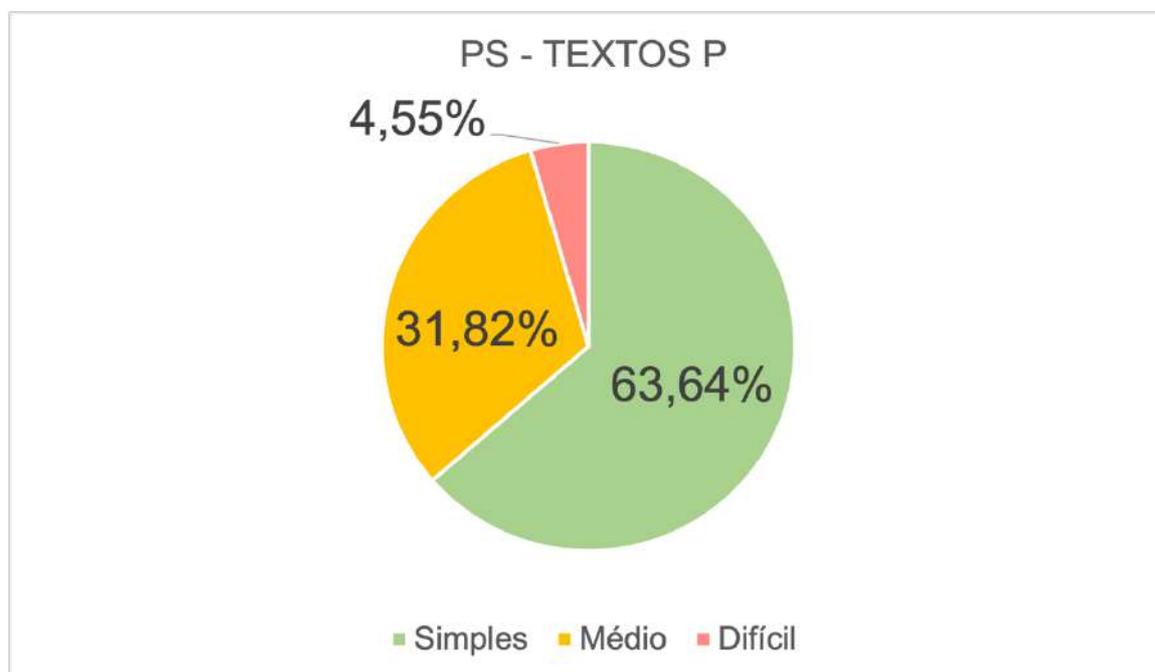
Legenda: PS: Polissemia | IA: Idade de aquisição | FPC: Familiaridade com palavras de conteúdo | CLP: Concretude Lexical | TT: *Type-Token*

Seguindo a ordenação de apresentação dos índices na tabela, iniciaremos nossa análise pela métrica PS.

8.5.2.1.1 Análise da métrica PS dos textos P

Por intermédio do gráfico abaixo (Gráfico 23), observamos que as orientações de interpretação de complexidade estimada pela PS sugerem que 63,64% dos textos são considerados simples, 31,82%, medianos e 4,55%, complexos.

Gráfico 23 - Classificação percentual de complexidade de PS do livro *Pre-intermediate*



Fonte: Elaboração própria

Curiosamente, se compararmos as informações do gráfico acima aos valores percentuais de classificação de complexidade da PS dos textos do livro *Elementary*, percebemos que a expectativa de resultados que indicassem maiores complexidades nos textos do livro *Pre-intermediate* (uma vez que é um nível acima do livro anterior)

não é alcançada. Essa métrica indica maior quantidade de textos mais simples no nível pré-intermediário do que nos materiais do nível elementar, como podemos observar pelo Quadro 39 a seguir:

Quadro 39 - Comparativo dos valores de PS entre os textos E e P

	TEXTOS E	TEXTOS P
Simple	35%	63,64%
Médio	45%	31,82%
Difícil	20%	4,55%

Fonte: Elaboração própria

Os números do quadro acima, quando comparados, revelam que há maiores frequências de palavras polissêmicas no livro *Elementary* do que no *Pre-intermediate*. Nossas observações sobre os textos classificados como mais simples e mais complexos dessa métrica podem auxiliar a compreender os motivos pelos quais a PS indicou maior complexidade nos textos E.

A tabela das métricas dos textos P destaca o texto P4 como tendo o melhor resultado de PS (PS: 5,09) e o P15 como pior (PS: 3,20) – textos que merecem nossa atenção pelas suas classificações.

Além dos textos supracitados (P4 e P15), também achamos válido observar a presença de polissemia em outros dois textos com posições de destaque: o texto com o segundo melhor resultado de PS, P2 (PS: 4,71) e o texto com segundo resultado mais baixo dessa métrica, P19 (PS: 3,49). Para isso, apresentamos no Quadro 40 abaixo exemplos de palavras polissêmicas nesses textos e os contextos nos quais essas unidades lexicais estão inseridas:

Quadro 40 – Palavras dos textos do livro *Pre-intermediate* que se destacam por PS

TEXTO	PALAVRAS POLISSÊMICAS	CONTEXTO
P4 (PS: 5,09)	<i>back</i>	<i>we were walking back to the hotel</i>
	<i>can</i>	<i>lots of people in the photo can be a good thing</i>
	<i>good</i>	<i>The light in the evening is also good</i>
	<i>look</i>	<i>look for little details that other photographers haven't seen</i>
	<i>perfect</i>	<i>So how do we take the perfect holiday photos?</i>

	<i>record</i>	<i>and we want to record our surroundings and happy memories</i>
	<i>right</i>	<i>and I think the picture on the right is much better</i>
P2 (PS: 4,71)	<i>back</i>	<i>and when we got back to the car</i>
	<i>distance</i>	<i>we could see the sea in the distance</i>
	<i>end</i>	<i>so in the end we all went up again</i>
	<i>leave</i>	<i>I didn't want to leave my phone up there</i>
	<i>own</i>	<i>'OK, but you can't go on your own'</i>
	<i>recording</i>	<i>Email us your stories or send us a recording</i>
	<i>top</i>	<i>at the top of the mountain</i>
P15 (PS: 3,2)	<i>age</i>	<i>At the age of 15</i>
	<i>books</i>	<i>She has written several bestselling children's books.</i>
	<i>course</i>	<i>and of course in her own life</i>
	<i>experiences</i>	<i>in the experiences of her children</i>
	<i>film</i>	<i>was in more than 50 films and many TV series</i>
	<i>like</i>	<i>Like mother, like daughter</i>
	<i>own</i>	<i>in her own life</i>
	<i>places</i>	<i>including the comedies 'Trading Places'</i>
<i>rest</i>	<i>and this marriage lasted for the rest of her life</i>	
P19 (PS: 3,49)	<i>Age</i>	<i>they're the same age and size</i>
	<i>colour</i>	<i>Their hair is the same colour and length</i>
	<i>couple</i>	<i>was adopted by a couple named Lewis in Lima</i>
	<i>like</i>	<i>they probably had a conversation something like this</i>
	<i>look</i>	<i>The two girls look identical</i>
	<i>project</i>	<i>that they set up an online project</i>

Fonte: Elaboração própria

Não só pelos resultados numéricos de PS, mas também pela análise contextual apresentada no quadro acima, podemos perceber que polissemia parece não ser um fator cujo potencial de complexidade seja alto para os textos do *livro Pre-intermediate*. O quadro acima nos auxilia a perceber que nos contextos nas quais as palavras polissêmicas estão inseridas, as pistas contextuais auxiliam o leitor a evocar o significado pretendido pelo texto.

Conforme já havíamos mencionado na análise da métrica PS para os textos *Elementary*, acreditamos que a polissemia não seja um fator que, em realidade, possa potencializar a complexidade dos textos para os níveis iniciais. Neste ponto da pesquisa, levantamos a hipótese de que a polissemia possa aumentar ou diminuir a acessibilidade dos textos ou em situações muito pontuais, ou em contextos em que mais de um significado possa ser ativado e, como já mencionamos, para que múltiplas significações possam ser evocadas, é preciso domínio de um vocabulário

relativamente extenso. Ter um léxico amplo está geralmente relacionado a maiores domínios da língua e, então, as pistas contextuais darão indícios, durante a leitura, de qual significado da palavra o texto se refere.

Ao mesmo tempo, é preciso lembrar que não podemos supor interpretações tomando o lugar do nosso leitor. É possível que algum estudante não consiga encontrar as pistas contextuais nas quais uma palavra polissêmica esteja inserida e relacione essa unidade ao significado não pretendido. Sendo assim, a interpretação global do texto pode ser prejudicada.

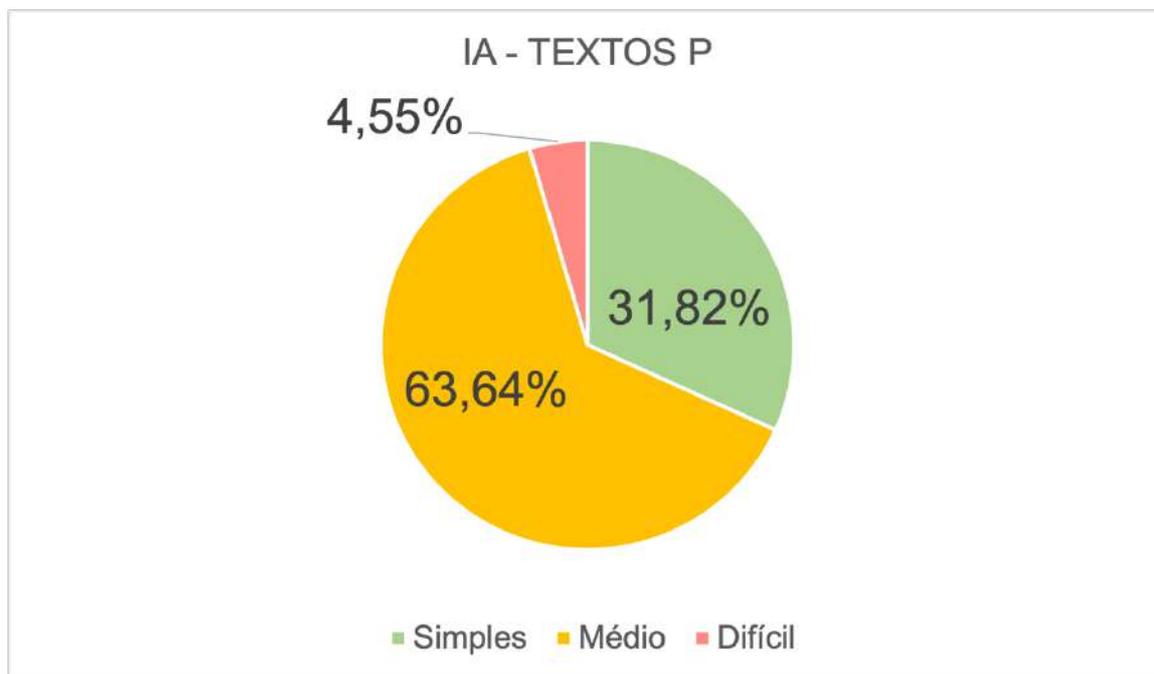
Além de tudo, especialmente no nosso caso, temos alunos nos níveis pré-intermediário cujo nível linguístico é, na realidade, básico. Esses estudantes geralmente optam por cursar os níveis pré-intermediários por exigências dos seus cursos de graduação. Uma vez que não existem requisitos nas disciplinas eletivas de língua, com grande frequência, alunos cursam níveis acima dos seus. Sendo assim, mesmo indicadores, como o caso da PS, que dificilmente se tornariam características que dificultariam a leitura (pelos motivos supracitados), precisamos, conforme já destacamos inúmeras vezes, manter em mente que sempre teremos diversos tipos de leitores e, conseqüentemente, de dificuldades de leitura.

Entretanto, ainda que não descartemos a validade da métrica PS como informativa para análises de complexidade textual, nossa observação dos textos destacados como tendo complexidade “média” e “alta” nos permitem afirmar que, precisamente para os textos do livro *Pre-intermediate*, são raras as instâncias que poderiam interferir na acessibilidade dos alunos.

Uma vez que a métrica PS não nos revela informações que nos permitam interpretar com fidedignidade o potencial de complexidade do léxico e da semântica, passamos para a análise dos valores de IA.

8.5.2.1.2 Análise da métrica IA dos textos P

O Gráfico 24 abaixo nos auxilia a iniciar nossa análise da métrica IA pela classificação percentual de complexidade dos textos P:

Gráfico 24 - Classificação percentual de complexidade de IA do livro *Pre-intermediate*

Fonte: Elaboração própria

Ao levarmos em consideração que os textos em foco são parte do livro *Pre-intermediate*, os percentuais de classificação de complexidade parecem ser coerentes com o nível: 63,64% são indicativos de complexidade média, enquanto 31,82% demarcados com complexidade baixa e 4,55%, alta.

Diante desses resultados, destacamos que o texto com melhor valor de IA é o P2 (IA: 238,83) e o texto com maior índice de complexidade por essa métrica é o P7 (IA: 361,96). Além do texto P2, outro cujo resultado de IA é positivo é o P4 (IA: 255,13). Curiosamente, tanto o texto P2 quanto o P4 foram os que resultaram a segunda e a primeira posição, respectivamente, nas classificações de acessibilidade de PS (métrica analisada anteriormente).

A leitura do texto P2 nos permite compreender o resultado positivo dessa métrica. O texto, que é um “*checklist*” do que é preciso verificar antes de viajar, é composto com vocabulários simples e concretos e que, geralmente, são aprendidos cedo na infância, como ‘*day*’, ‘*lunch*’, ‘*photos*’, ‘*sun*’, ‘*car*’ etc.

De forma similar, no texto P4 também há alta frequência de palavras comuns para crianças (mesmo *toddlers*), tais como ‘*place*’, ‘*photos*’, ‘*camera*’, ‘*sky*’, ‘*light*’ etc. Ainda que nele existam outras palavras que, naturalmente, são aprendidas em idades mais avançadas (como ‘*crowded*’, ‘*enthusiastic*’ e ‘*angles*’, por exemplo) a proporção de palavras mais simples para palavras mais complexas é de aproximadamente 7:1;

ou seja, proporcionalmente estão presentes mais unidades lexicais simples do que complexas. Além desses dois textos, apresentamos no quadro abaixo (Quadro 41) outros com valores de IA indicativos de alta acessibilidade e alguns trechos que ilustrem o vocabulário presente neles:

Quadro 41 – Trechos de textos do livro *Pre-intermediate* que se destacam positivamente por IA

TEXTO	TRECHOS
P19 (IA: 262,62)	<i>The two girls look identical. Their hair is the same colour and length, they're the same age and size, and when you see them together, in the photo they put on social media, you would think that they were identical twins.</i>
P1 (IA: 264,05)	<i>My dad is the perfect man,' she says. 'He's warm, he's generous, he knows how to look after a woman, how to treat a woman.</i>
P20 (IA: 268,33)	<i>He stopped at a petrol station, filled up the car with petrol, and went to the toilet.</i>
P21 (IA: 269,15)	<i>Every day he used to sit with his friends and talk about other people. 'Can you believe she did that?'</i>
P13 (IA: 271,45)	<i>Her hair was red gold, and her eyes were sea-blue. She was wearing a white top and a dark skirt.</i>

Fonte: Elaboração própria

Ao lembramos que a métrica IA tem como base para cálculo a relação entre aprendizado e idade e que ela destaca textos com grandes frequências de palavras que são aprendidas na primeira infância (fator indicativo de acessibilidade), podemos observar pelos textos apresentados no Quadro 41 acima muitas palavras que são comuns a crianças (mesmo *toddlers*) e, portanto, tendem a potencializar a acessibilidade textual.

Destacamos que o texto P1 (por ser o primeiro texto do livro *Pre-intermediate*) embora com resultados positivos de IA, apresenta palavras e expressões que, ao mesmo tempo, podem ser consideradas complexas para alunos brasileiros (como *'spark', 'clever', 'end up', 'dates', 'look after'*). Embora essas unidades não sejam, em realidade, tão complexas, podemos destacá-las como possíveis fatores de complexidade. Nesse ponto, não vemos a necessidade de adoção de estratégias para simplificar os textos antes da leitura dos alunos, mas a verificação do docente para essas palavras para garantir a compreensão dos alunos, especialmente porque *phrasal verbs* (*'end up'*) é um conteúdo ensinado apenas na metade final do livro e a palavra *'date'* é um falso cognato para o português (embora, atualmente, ela esteja sendo frequentemente utilizada na língua portuguesa denotando seu sentido verdadeiro, “encontro”).

Já nos textos cuja valoração os classifica como os mais complexos pela IA (P7, com 361,96 e P11, com 343,27), podemos perceber que a frequência de palavras mais complexas (e que geralmente não são aprendidas nas primeiras infâncias, diretriz da métrica IA) é muito mais alta.

No texto P7, que fala sobre criação e publicação de novas palavras em língua inglesa, constam palavras como ‘*noun*’, ‘*networking*’, ‘*barista*’, ‘*latte*’, ‘*gadgets*’, ‘*ringtone*’, ‘*phenomenon*’ etc., que também estão presentes no texto P11, como, por exemplo ‘*civil*’, ‘*hostile*’, ‘*advisers*’, ‘*speech*’, ‘*preachers*’, etc. Ainda que algumas dessas palavras sejam cognatas verdadeiras da língua portuguesa, elas são apenas exemplos da complexidade do vocabulário desses dois textos, sendo o P11, inclusive, classificado como o segundo texto mais complexo pela métrica FPC, sobre a qual falaremos posteriormente.

Além dos textos supracitados, no quadro abaixo (Quadro 42) apresentamos alguns textos e trechos com vocabulários que podem ser considerados complexos e que não fazem parte do léxico comum de uma criança (indicativo de potencial de complexidade). Esses destaques são provenientes não somente das indicações da métrica IA, mas também da nossa experiência em sala de aula. No quadro, destacamos em negrito palavras que consideramos potencialmente complexas:

Quadro 42 – Trechos de textos do livro *Pre-intermediate* com valores médios de IA

TEXTO	TRECHOS
P14 (IA: 313)	... because you're busy unloading your trolley Most people are right-handed join the queue with the fewest people...
P6 (IA: 300,31)	The survey by a travel website, shows that people invent stories to make their lives appear more interesting than they really are. ... and 3% of people even put on fake tan Social media is becoming increasingly important in our lives...
P22 (IA: 300,27)	An Italian couple ran out onto the runway ... The engines were running and the stairs had been removed, but the couple tried signalling to the pilots... ... spokesman Greg Parashak said on Monday...

Fonte: Elaboração própria

Como podemos observar, algumas palavras tendem a ser complexas para nossos alunos de nível pré-intermediário. Visto que as palavras destacadas não são os focos da unidade e que existem estratégias de substituição para algumas delas, acreditamos que seria possível otimizar o vocabulário em determinadas instâncias para melhorar a acessibilidade. Voltamos a destacar que é preciso ter bom senso ao

adotar estratégias e fazê-las de acordo com a necessidade e limitações dos alunos e que cada edição com finalidades de simplificação deve levar em conta o aspecto da manutenção linguística. Como os textos são voltados para o ensino, é preciso sempre ter um equilíbrio entre otimizar a acessibilidade e promover insumos suficientes para o aprendizado.

É importante destacar que apenas um texto foi classificado como tendo alta complexidade pela IA (P7). O texto P11, assim como os outros 63,64% (14 dos 21 textos) foram considerados como tendo complexidade média. É essencial reafirmar, conforme mencionamos no início dessa seção, que os valores observados por essa métrica não são fatores de alta preocupação, visto que os textos estão presentes no livro *Pre-intermediate* e precisam apresentar complexidades um pouco mais altas que os textos do livro *Elementary* para a manutenção linguística dos estudantes.

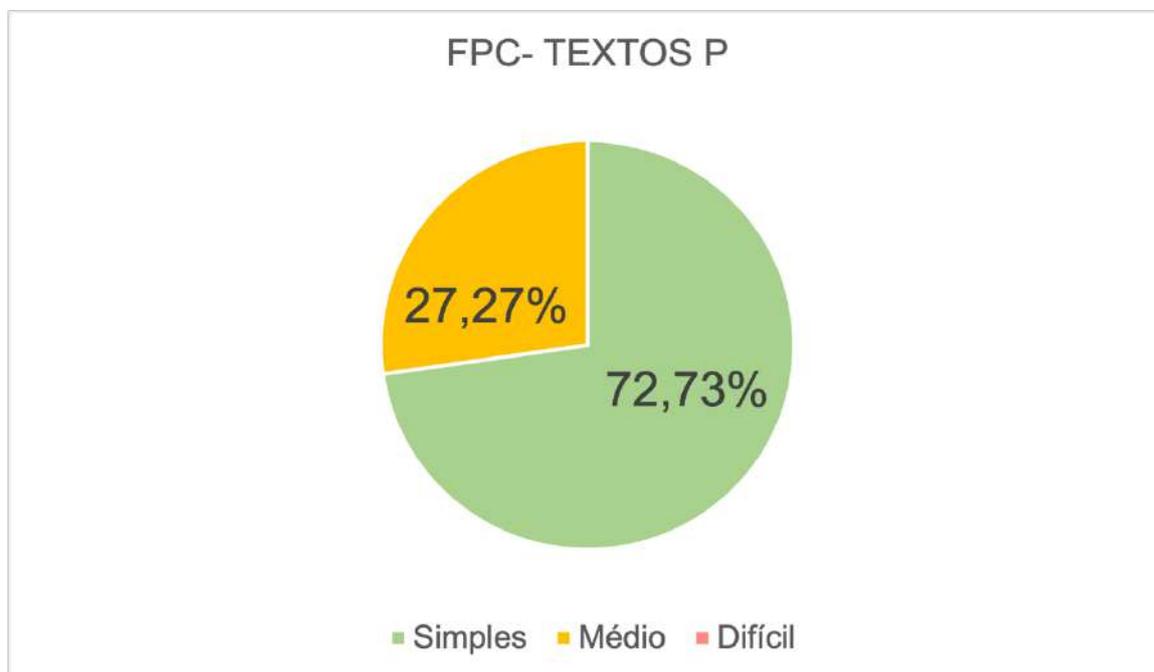
Ainda assim, lembramos que em casos particulares (como alunos neuroatípicos ou estudantes que estão cursando uma disciplina de nível pré-intermediário quando seu nível linguístico é, em realidade, básico), os textos destacados em amarelo (indicativo de complexidade média) no quadro de métricas devem ter seu léxico analisado para que estratégias de simplificação sejam adotadas quando se julgar necessário.

Para complementar a análise de complexidade lexical e semântica dos textos P, passamos para o estudo da métrica FPC.

8.5.2.1.3 Análise da métrica FPC dos textos P

Conforme já mencionamos, a métrica FPC tem forte relação com a IA – embora ela estime a familiaridade com palavras de conteúdo e a IA observe idade de aquisição. Sendo assim, as duas perspectivas nos permitem uma observação mais global do comportamento lexical dos textos. Para que possamos analisar esse índice para os textos E, iniciamos as observações pela classificação percentual conforme o gráfico abaixo (Gráfico 25):

Gráfico 25 - Classificação percentual de complexidade de FPC do livro *Pre-intermediate*



Fonte: Elaboração própria

O gráfico acima (Gráfico 25) exibe que 72,73% dos textos do livro *Pre-intermediate* podem ser considerados simples pela métrica FPC, enquanto 27,27%, como médios. Nenhum texto é classificado como tendo alta complexidade.

Como procedemos até o momento, destacamos o texto apontado como o mais acessível e o menos acessível por cada métrica. No caso da FPC, esses textos são, respectivamente, o P1 (FPC: 593,82) e o P17 (FPC: 566,95), sobre os quais falaremos a seguir.

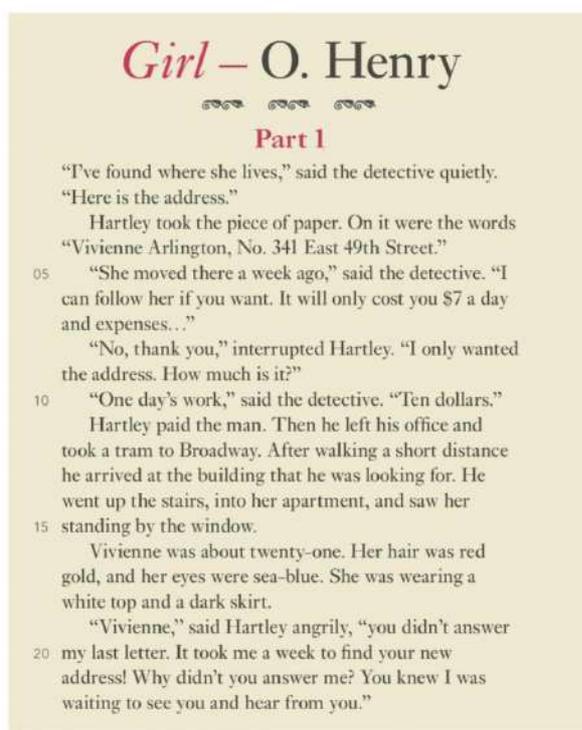
A indicação de que o texto P1 é classificado como tendo mais alta acessibilidade pela FPC é o cenário que seria ideal ao falarmos de progressão de complexidade. Se buscássemos aumentar a complexidade de forma progressiva, seria interessante termos textos mais acessíveis nas primeiras unidades de cada livro e, com o avanço das unidades, os valores métricos nos demonstrarem aumento de complexidade (de forma equilibrada e coerente). Posteriormente, analisaremos se a progressão que é esperada ocorre dentro de cada nível e entre eles.

No texto P1, podemos observar uma grande frequência de substantivos comuns (como, por exemplo, '*parent*', '*child*', '*dad*', '*man*', '*woman*') e de adjetivos que

são usualmente ensinados em livros e cursos de língua em níveis básicos (como, por exemplo, *'perfect'*, *'warm'*, *'fun'*, *'young'*, *'difficult'* etc.). Por trabalhar com palavras de conteúdo, a presença de substantivos e adjetivos simples (como os exemplos apresentados) potencializa o valor de FPC.

Outro texto com valores positivos de FPC é o texto P13 (FPC: 593,12). Esse texto é uma narrativa que – com a finalidade de “alimentar” o imaginário do leitor para que ele possa construir a história em sua mente – apresenta uma grande quantidade de palavras de conteúdo simples (como *'paper'*, *'week'*, *'address'*, *'building'*, *'window'* e *'stairs'*, por exemplo). Conforme mencionado no parágrafo anterior, a indicação de baixa complexidade pela FPC é consequência da proporção favorável às palavras simples do texto, uma vez que, ao verificarmos a lista de palavras desse material, podemos perceber que são poucas as unidades lexicais de conteúdo que poderiam ser consideradas menos familiares (como *'suburbs'* e *'masterfully'*, por exemplo). Na imagem abaixo (Figura 30) podemos ler um trecho do texto P13 para ilustrar seu resultado positivo de FPC:

Figura 30 – Texto do livro *Pre-intermediate* com valor positivo de FPC



Fonte: Latham-Koenig, Oxenden e Lambert (2019, p. 66)

Seguindo um caminho inverso aos textos P1 e P13, o texto P17 (considerado o mais complexo pela FPC) é rico em palavras de conteúdo menos familiares. Nesse texto, encontramos palavras que podem ser complexas até mesmo para um falante nativo com instrução formal limitada. Alguns exemplos dessas palavras são: ‘*sportsperson*’, ‘*blaming*’, ‘*renewed*’, ‘*disqualified*’, ‘*dehydrated*’, ‘*referee*’, ‘*banned*’ e ‘*delighted*’, entre outras palavras que podemos observar pelo trecho do texto na Figura 31 abaixo:

Figura 31 – Texto do livro *Pre-intermediate* com valor positivo indicativo de maior complexidade pela FPC



Fonte: Latham-Koenig, Oxenden e Lambert (2019, p. 85)

A presença de palavras menos familiares também ocorre no texto P11, cujo valor de FPC o indica como o segundo mais complexo e que também foi classificado nessa posição pela IA. Alguns exemplos de palavras de conteúdo que não são tão familiares incluem: ‘*legend*’, ‘*hostile*’, ‘*civil*’, ‘*audience*’, ‘*racism*’, ‘*content*’ etc. Além do fato de serem palavras classificadas com valores baixos na lista-base da FPC, algumas delas podem ser ainda mais complexas para falantes do português por serem falsos cognatos, como ‘*legend*’, ‘*audience*’ e ‘*content*’ que podem ser confundidas com “legenda”, “audiência” (de tribunal) e “contente”, respectivamente.

Apresentamos no quadro abaixo (Quadro 43) outros textos⁶⁶ cujos valores de FPC os classificam como tendo complexidade média e destacamos alguns trechos com palavras menos acessíveis:

Quadro 43 – Trechos de textos do livro *Pre-intermediate* com valores médios de FPC

TEXTO	TRECHOS
P7 (FPC: 578,26)	<i>We can make a new word by combining... ... like gastropub... ... like barista or latte... ... the names of brands or companies... ... describe new technology or new gadgets...</i>
P5 (FPC: 579,35)	<i>The aim of the game... The presenter gives the contestant...</i>

Fonte: Elaboração própria

Os exemplos apresentados no quadro acima são unidades lexicais que interferem diretamente nos valores da métrica FPC. Entretanto, ainda que algumas delas possam ser consideradas não tão familiares, o contexto social de um aluno brasileiro auxiliaria na interpretação do significado (como nos exemplos de ‘*barista*’ e ‘*latte*’), visto que essas palavras são utilizadas também no vernáculo da língua portuguesa. Sendo assim, destacamos a importância de observar os indícios da métrica não como respostas sumárias de complexidade, mas sempre diante do contexto social do leitor, conforme explicamos no capítulo teórico sobre AT.

Ainda a respeito dos exemplos apresentados acima, é preciso atentar para o caso da palavra “*gastropub*”. Embora complexa, sua contextualização no texto explica sua presença. Essa unidade é apresentada como forma de ilustrar a criação de novas palavras ao combinar duas já existentes. Dessa forma, essa palavra interferiu no cálculo realizado pelo Coh-Metrix para o FPC, mas deve ser desconsiderada como exemplo de alta complexidade por ser propositalmente apresentada e delimitada.

Uma vez que, 72,73% dos textos foram classificados como simples pela métrica FPC, talvez nossa preocupação a respeito do que isso indica não seja a respeito da falta de acessibilidade por alunos com maiores limitações, mas, talvez, de alta acessibilidade para alunos que estão nos níveis pré-intermediário; o que nos leva a questionar: será que alguns desses textos não poderiam conter palavras um pouco

⁶⁶ Outro texto que poderíamos destacar é o P22, mas apresentamos exemplos sobre ele a respeito da métrica IA que também se aplicam para a FPC.

mais complexas para a manutenção de vocabulário com a finalidade de aprender inglês como língua adicional?

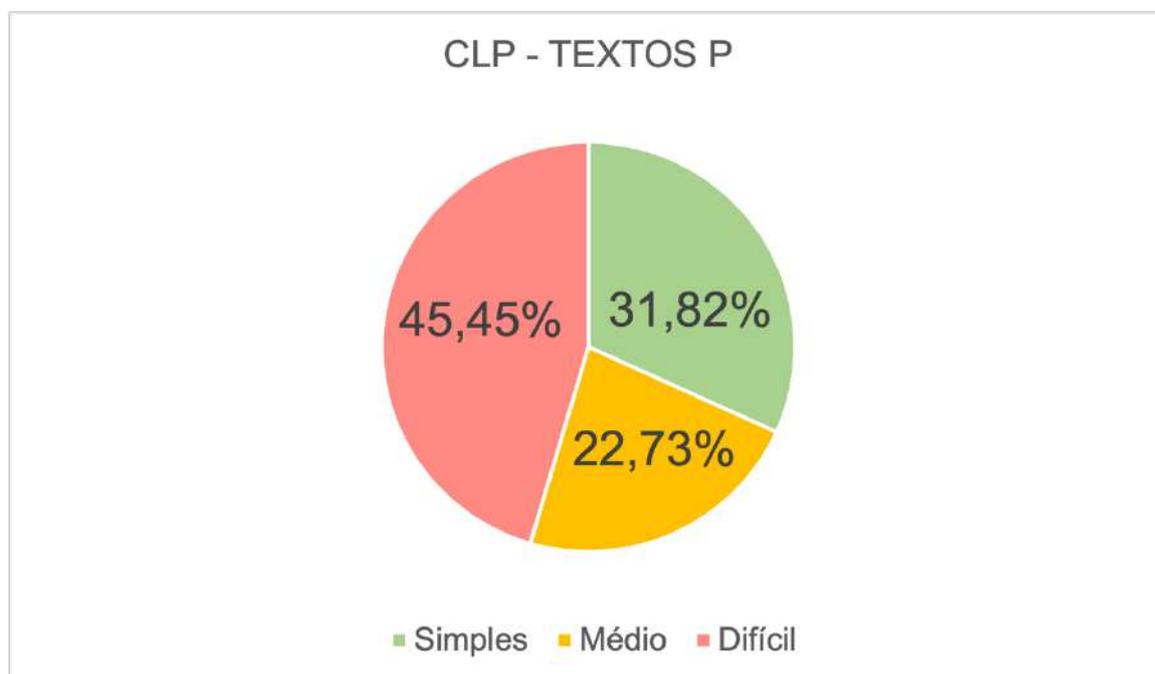
Se pegarmos o texto P12, por exemplo, cujo valor de FPC indica baixa complexidade e que está posicionado entre um texto considerado como o segundo mais simples por essa métrica (P13) e o segundo mais complexo (P11), podemos perceber que, com a finalidade de progredir a complexidade para a manutenção lexical, existem palavras que poderiam ser facilmente substituídas por outras um pouco mais complexas ('*sometimes*' por '*occasionally*', '*stop*' por '*cease*', '*continue*' por '*pursue*', por exemplo).

Obviamente, a estratégia de substituir palavras por sinônimos mais complexos deve ser adotada apenas quando o professor tiver por finalidade proporcionar uma interpretação de texto mais desafiadora e apresentar vocabulários que podem ser novos para alunos e turmas que tenham a necessidade ou o potencial de desenvolver conhecimentos linguísticos por meio de textos mais complexos (sendo válido lembrar que a customização, como mencionado no capítulo sobre Educação 3.0, 4.0 e 5.0 não é prática exclusiva para alunos com limitações, mas que também é aconselhável para aumentar os insumos de aprendizado para alunos com maiores capacidades).

Prosseguindo com a análise das métricas de léxico e semântica dos textos P, passamos para as observações sobre a métrica CLP.

8.5.2.1.4 Análise da métrica CLP dos textos P

Ainda no contexto da análise de léxico e semântica, o gráfico abaixo (Gráfico 26) exhibe a classificação percentual de complexidade dos textos do livro *Pre-intermediate* pela métrica CLP, que classifica 31,82% dos textos como simples, 22,73% como tendo complexidade média e 45,45% como difíceis.

Gráfico 26 - Classificação percentual de complexidade de CLP do livro *Pre-intermediate*

Fonte: Elaboração própria

Dado que a métrica CLP resulta em percentuais de concretude lexical, precisamos observar a proporção de palavras concretas pelas abstratas nos textos porque, conforme indica as diretrizes da métrica, maiores presenças de unidades lexicais concretas tendem a potencializar a acessibilidade. Diante de 45,45% dos textos tendo a classificação de CLP como indicativas de alta complexidade, precisamos observar essa característica léxico-semântica com cuidado.

No texto classificado como mais simples por essa métrica, P16 (CLP: 94,74%), encontramos uma grande quantidade de palavras concretas cujo significante evoca uma imagem mental de algo concreto no mundo, como, por exemplo *'tennis'*, *'football'*, *'basketball'*, *'winter'*, *'summer'* e outras palavras relacionadas a esportes (tema do texto). Já palavras abstratas são poucas e, em escalas de abstração, não demandam muito esforço cognitivo para ativação de significado, tais como *'popular'*, *'interesting'* e *'strength'*, por exemplo.

O segundo texto mais acessível pela CLP (P9, CLP: 93,82%) também é rico em palavras concretas, tais como *'coffee'*, *'juice'*, *'water'*, *'milk'*, *'tea'* e outras unidades lexicais referentes à temática "bebidas saudáveis". Em comparação, são poucas palavras que podem ser consideradas abstratas, como *'stress'*, *'high'* e *'risk'*, por exemplo.

Por outro lado, nos textos considerados mais complexos pela CLP (P14, com 5,05% e P13, com 19,22%), a presença de palavras concretas é menor, não só em comparação às palavras abstratas, mas principalmente pela constante repetição das palavras concretas que são difíceis de retomar por meio de sinônimos.

Nos textos P14 e P13, as palavras concretas tendem a se repetir (como *'queue'*, *'supermarket'* e *'girl'*), diminuindo a proporção entre palavras concretas e abstratas. Ao mesmo tempo, a frequência de adjetivos, advérbios e substantivos abstratos, (como *'fast'*, *'simple'*, *'slowly'* e *'home'*) é alta.

Diante da sinalização das métricas indicando baixos valores de concretude lexical (e tendo consciência de que essa indicação é apontada na literatura de AT como indício de alta complexidade textual), voltamos nossa atenção para os textos com valores baixos de CLP para compreender os motivos pelos quais a métrica apresentou resultados tão baixos. Contudo, durante nosso estudo, relatamos que a seleção de palavras abstratas (de forma manual) que pudessem ser consideradas verdadeiras indicações de alta complexidade foi uma tarefa difícil e, ao mesmo tempo, ao encontrarmos palavras possivelmente abstratas, percebemos que a maior parte delas não são unidades lexicais complexas. Diante disso, questionamos nossa compreensão sobre a diferença entre palavras abstratas e concretas e, então, decidimos aprofundar nossos conhecimentos sobre esses conceitos. Apresentamos, então, uma conceitualização do que significa “abstrato” que julgamos compreensiva e verossímil.

Para César (2010), a definição de palavra abstrata como “tudo aquilo que não pode ser observado por um dos cinco sentidos humanos” é muito reducionista. Para ele, os autores modernos - ao compreender o conceito de ‘palavras abstratas’ como uma questão não só gramatical, mas também filosófica – têm apresentado definições mais eficientes para esse conceito, como explica no parágrafo abaixo:

Substantivo concreto é aquele que designa ou pode designar o ser de existência real em si mesmo, independentemente de outro sujeito. (Note-se: é necessário um pouco de abstração filosófica para entender-se esta doutrina.) — Atente-se contudo desde já o leitor cauteloso que um mesmo substantivo pode ser abstrato ou concreto conforme o sentido e o valor funcional que leva dentro da frase. [...] "Fogo", v. g., será concreto quando designar o fogo real, existente em si só, independente de qualquer outro sujeito, mas será abstrato, quando sua existência for fictícia ou imaginária, significando mais uma semelhança com o nosso fogo comum e real (por isso concreto) ; assim, falando-se do fogo do amor, da caridade, o substantivo "fogo" é tomado no sentido abstrato (ou pelo menos, como fictício, pseudoconcreto), porque abstratos são os substantivos de existência

imaginária, que não subsistem por si sós, mas se devem apoiar em outro sujeito (substantivo). São, em geral, meras qualidades de seres concretos, p. ex., amor, justiça, beleza. Daí se explica a força de expressões como “a bondade em pessoa”, “a justiça personificada”, onde se procura concretizar certos substantivos que, de natureza, são abstratos. Quem, pois, partir do princípio absoluto de que “concreto” é substantivo que tem substância (tal é a afirmação do distinto consulente) cairá em erro e se verá em dificuldades para classificar certos vocábulos que podem assumir acepções várias. O substantivo concreto, sobre ser substância, deve vir acompanhado de existência real, independente de qualquer outro sujeito. O abstrato, ao contrário, anda sempre encostado noutro sujeito, tendo existência independente somente no intelecto ou na imaginação. (CÉSAR, 2010, p. 26-27).

Ora, se unidades lexicais podem ser consideradas concretas ou abstratas não necessariamente por si só, mas na ativação de seu significado no contexto, então a distinção entre concreto e abstrato quando realizada por uma máquina pode ser altamente imprecisa.

Além das observações acima, mesmo levando em consideração os distintos aspectos que podem classificar uma palavra como concreta ou abstrata, nossas análises dos textos nos permitiram concluir que, talvez, palavras que designem conceitos abstratos não sejam fatores de complexidade pelo fato de designarem conceitos abstratos. O que pode ser visto como potencial de complexidade é o [des]conhecimento do léxico em si, ou a dificuldade de compreender determinadas palavras, mas – raramente – a relação entre significante e significado. Sendo assim, compreendemos que embora a métrica possa auxiliar na estimativa de complexidade, os conceitos expressos por palavras abstratas não são estranhos aos nossos leitores, uma vez que são estudantes adultos de níveis superiores. Além do mais, são escassas as ocorrências de palavras nos textos analisados designem conceitos que são desconhecidos para nosso perfil de leitor.

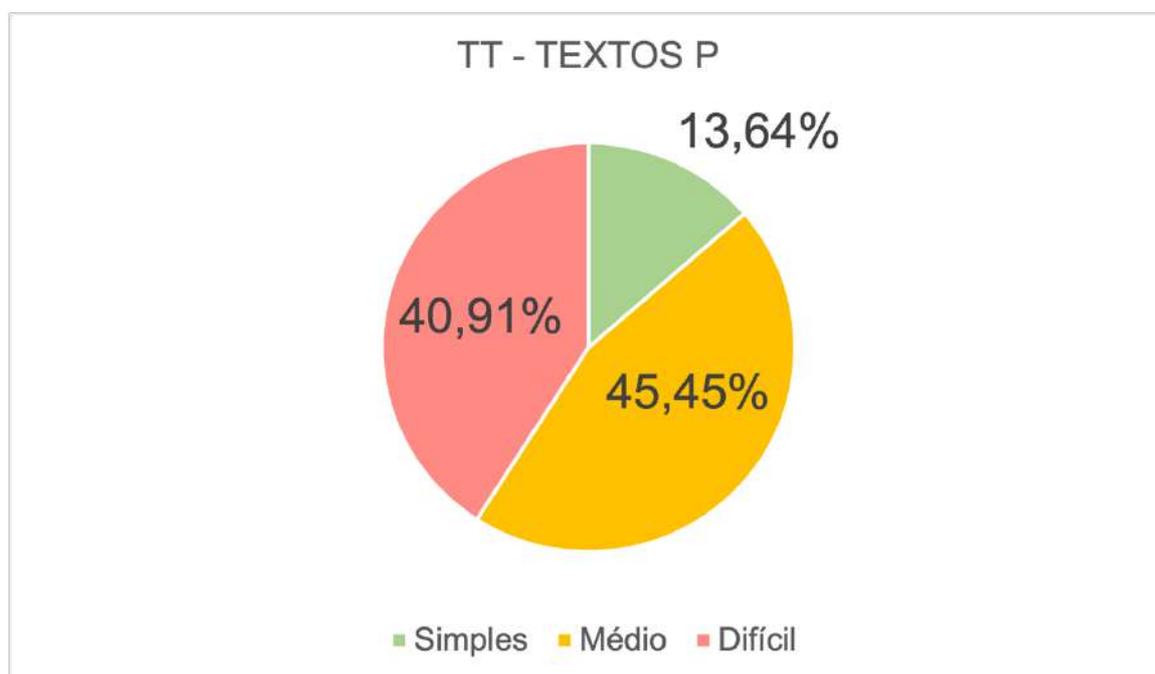
Embora pudéssemos optar por remover essa métrica da nossa pesquisa, acreditamos que nossas conclusões a respeito dela podem ser válidas para futuros pesquisadores que considerem utilizar a CLP como índice de complexidade. Obviamente, esse índice pode auxiliar em determinados casos a, principalmente, destacar textos que possam apresentar palavras abstratas que tornem difícil para um leitor relacionar o significante ao significado para que um docente, diante dessa sinalização, possa observar o texto com detalhes e, ao julgar necessário, adequá-lo.

Para que possamos finalizar nosso estudo das métricas de léxico e semântica dos textos P, prosseguimos para a análise de TT.

8.5.2.1.5 Análise da métrica TT dos textos P

O gráfico abaixo (Gráfico 27) demonstra o percentual de classificação de complexidade de cada texto pelos indicativos de TT. Como podemos observar, 13,64% dos textos apontam que essa métrica indica complexidade baixa, 45,45%, complexidade média e 40,91%, alta.

Gráfico 27 - Classificação percentual de complexidade de TT do livro *Pre-intermediate*



Fonte: Elaboração própria

Uma vez que os textos em foco fazem parte do livro *Pre-intermediate*, era esperado (e idealizado) que os valores das métricas indicassem complexidade um pouco mais alta do que os textos do nível *Elementary* – fato que pôde ser corroborado pela TT. Entretanto, o percentual de textos considerados como difíceis pode ser um sinal de alerta para a complexidade relacionada à variedade lexical, especialmente porque a maior parte dos textos destacados com complexidade alta apresentam valores distantes do limite entre a classificação média e alta.

Ao mesmo tempo, é importante lembrar que textos narrativos (ou com sequência similar à narrativa) tendem a ter altas variações lexicais, conforme apontou Silva (2018) – uma vez que para o desenvolvimento lógico das histórias, novas palavras são necessárias para descrever personagens, cenários, objetos etc. Ainda

assim, havíamos observado valores com maiores indícios de baixa complexidade nos textos *Elementary*, fato que contribui para que possamos observar os textos do nível pré-intermediário com atenção.

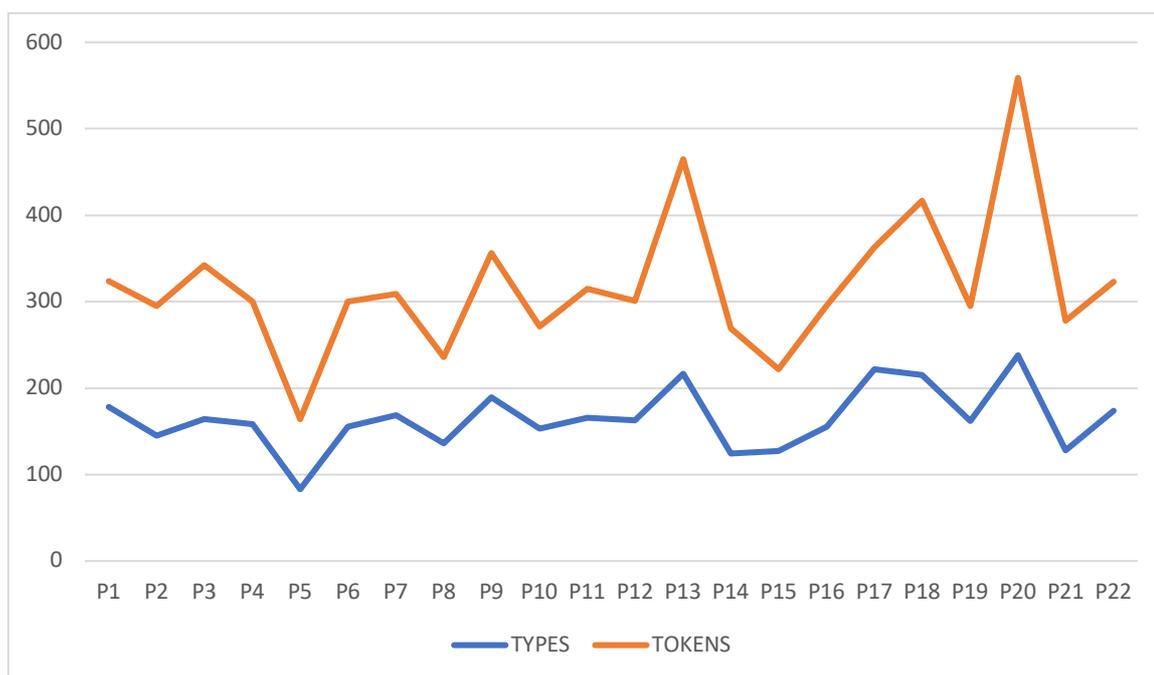
O quadro abaixo (Quadro 44) apresenta a quantidade de *types* e *tokens* de cada um dos textos P:

Quadro 44 – Quantidade de *types* e de *tokens* dos textos do livro *Pre-Intermediate*

TEXTO	TYPES	TOKENS
P1	178	324
P2	145	295
P3	164	342
P4	158	300
P5	83	164
P6	155	300
P7	169	309
P8	136	236
P9	189	356
P10	153	271
P11	166	315
P12	163	301
P13	217	465
P14	124	269
P15	127	222
P16	155	296
P17	222	363
P18	215	417
P19	162	295
P20	238	559
P21	128	278
P22	174	323

Fonte: Elaboração própria

Conforme realizamos a respeito dos textos do livro *Elementary*, o quadro acima (Quadro 44) e o gráfico abaixo (Gráfico 28) nos auxiliam a verificar se a extensão de cada texto tem sido progressiva (demonstrando maior potencial de complexidade, uma vez que extensão é apontadas por pesquisadores de AT - como Flesch (1970) e DuBay (2004) - como uma das características textuais de maior potencial de complexidade).

Gráfico 28 – *Types e tokens dos textos do livro Pre-intermediate em gráfico-linha*

Fonte: Elaboração própria

Pela leitura do quadro e do gráfico podemos constatar que:

- I. O primeiro e o último texto têm a mesma extensão;
- II. Existem mais picos de aumento de extensão do que de redução;
- III. A progressão da quantidade de palavras apresentadas por texto não é constante;
- IV. Os textos menos extensos (e potencialmente menos complexos por essa característica) estão presentes na parte inicial do livro;
- V. Os textos mais extensos (que pode representar maior complexidade) estão na parte final do livro.

Ao aplicarmos nosso procedimento metodológico com o apoio do Excel para destacar as palavras novas em cada texto, obtemos os seguintes resultados (exibidos no Quadro 45 abaixo):

Quadro 45 – Quantidade de palavras novas/únicas em cada texto do livro *Pre-intermediate*

TEXTO	QUANTIDADE DE PALAVRAS NOVAS/ÚNICAS
P1	42
P2	25
P3	44
P4	28

P5	23
P6	33
P7	61
P8	39
P9	84
P10	45
P11	60
P12	37
P13	55
P14	37
P15	48
P16	41
P17	90
P18	64
P19	46
P20	92
P21	30
P22	57

Fonte: Elaboração própria

Conforme os números nos indicam, a quantidade de palavras novas apresentadas por texto varia entre 23 e 92. Se imaginarmos que essas palavras podem estar sendo apresentadas pelos alunos pela primeira vez, seria uma prática didática inviável e ineficiente esperar que estudantes de ILA aprendam essas quantidades de palavras por texto, especialmente se não fizerem parte do vocabulário controlado da unidade. Contudo, é importante apresentarmos as seguintes observações:

I. Não é possível controlar ou estimar se, de fato, essas palavras são realmente novas para os alunos;

II. Ao levarmos em consideração a interpretação das métricas de léxico e semântica apresentadas nesta tese e destacarmos as palavras que são apresentadas pela primeira vez em cada unidade, podemos verificar quais delas poderiam ser substituídas caso não sejam fundamentais para a manutenção linguística;

III. Os dados a respeito das palavras novas servem apenas como mais um indício que pode auxiliar a destacar textos que podem ter altos potenciais de complexidade por léxico e semântica, mas esses números não podem ser os únicos (ou os mais importantes) indicativos de complexidade textual referente ao léxico. Constatando, por exemplo, que no texto P20, 92 palavras são únicas e apresentadas somente neles, podemos voltar nossa atenção para esse texto para ver se essa

quantidade não é demasiada e se existem unidades que podem potencializar a complexidade para os nossos alunos.

Pela interpretação dos valores de TT e dos percentuais de textos que foram classificados como tendo complexidade média (45,45%) e alta (40,91%), conforme já mencionado, precisamos atentar aos textos com valores altos de TT.

O texto com maior índice de complexidade pela TT é o P17 (TT: 846) e com valor sugestivo de maior acessibilidade, o P20 (TT: 637).

Ao analisarmos o texto P17 (seja pela leitura do texto ou pela lista de palavras), podemos ver que a variação lexical é muito alta, de fato. Ainda que o texto seja rico em nomes próprios (tanto de atletas, quanto de locais), ele também apresenta léxico bastante diversificado, como pode ser ilustrado pelo trecho a seguir:

“South Korean footballer Ahn Jung-Hwan scored the goal that sent Italy out of the 2002 World Cup when they beat them 2-1. But Jung-Hwan also played for the Italian football club Perugia. After the match, the president of the club, Luciano Gucci, announced that the player's contract would not be renewed. 'That gentleman will never set foot in Perugia again, Gucci said. 'I have no intention of paying a salary to somebody who has ruined Italian football.' Gucci later apologized, but Ahn Jung-Hwan left the club and never went back to an Italian club.”

Além da alta variedade lexical, grande parte das palavras utilizadas no texto P17 não são simples e podem dificultar a compreensão de um aluno no nível pré-intermediário – fato já sinalizado pelas demais métricas, uma vez que todas as anteriores destacaram esse texto como tendo resultados indicativos de complexidade média ou alta (sendo que, os índices com valores classificados como “complexidade média” estão, geralmente, no limite entre os valores de classificação “média” e “alta”).

Além do texto P17, outros resultaram valores de TT como indicativos de alta complexidade. Para ilustrar esses resultados, o quadro abaixo (Quadro) apresenta trechos de textos que foram classificados por TT como potencialmente complexos:

Quadro 46 – Trechos de textos do livro *Pre-intermediate* com valores indicativos de alta complexidade por TT

TEXTO	TRECHOS
P8 (TT: 823)	<i>Fast food restaurants have changed the way we eat. The growth of mobile phone apps like Tinder has changed the way we meet new people. We don't need to</i>

	<i>wait a week to see the next episode of a TV series - we can download it on Netflix. We are even walking 10% faster than 20 years ago, and talking more quickly. Everything is getting faster, but is it getting better? That depends on how fast you like to live.</i>
P15 (TT: 808)	<i>She has written several bestselling children's books. She says she finds the inspiration for her writing all around her - in the experiences of her children, her godchildren (one is actor Jake Gyllenhaal), her friends and of course in her own life.</i>
P3 (TT: 795)	<i>In 1972, I was living in London. I was in my early twenties, and I was working for an advertising agency. That summer, I went on a camping holiday with my parents in the south of France. One night, I went to a nightclub on the beach and I met a young Frenchman, and we fell in love. When I got home, I immediately started looking for a job in Paris. He was at university there - he was studying medicine. In the end, I found a job as a secretary with UNESCO, and I went to live there.</i>
P22 (TT: 774)	<i>Matteo Clementi, 26, and Enrica Apollonio, 23, got stuck in terrible traffic on their way to the airport on Wednesday. When they arrived, the gate was closed and they were not allowed to board their flight back to Italy. They went to the next gate, forced open a security door, and ran towards the plane. The engines were running and the stairs had been removed, but the couple tried signalling to the pilots to let them get on. However, they were not allowed to board and were arrested by security staff.</i>
P10 (TT: 765)	<i>Then they look at each possible problem and plan how to avoid it. So for the exam situation, they go to bed early and have a good night's sleep; they find out in advance exactly where the exam is; they eat a good breakfast, and take lots of pens and pencils, and a bottle of water; and they leave home early.</i>
P12 (TT: 754)	<i>Sometimes they try to say a couple of phrases in the local language, but they stop making an effort as soon as they discover that the waiter knows a little English. Some British people who live abroad often spend all their time with other Brits, and never learn the language at all.</i>
P11 (TT: 750)	<i>She was Aretha Franklin's mentor, and in 1961 she was the first gospel singer to win a Grammy Award. She was also an important member of the Civil Rights Movement in the USA, and she was a close friend of Martin Luther King. Mahalia often went with King on civil rights demonstrations and marches, including into the most hostile parts of the South, and she sang at the events. As a presenter from National Public Radio said, 'her voice became the soundtrack of the Civil Rights Movement.</i>

Fonte: Elaboração própria

Embora os trechos destacados pelo quadro acima sejam apenas fragmentos do texto, por intermédio deles podemos ter uma ideia de suas variedades lexicais. Em alguns casos, como no texto P12, por exemplo, percebemos a utilização de outras palavras para retomar conceitos já mencionados ('*British people*' – '*Brits*'), característica que aumenta a variedade lexical e, conseqüentemente, o potencial de complexidade (visto que exige maior domínio de vocabulário).

Já os textos cujos resultados de TT indicam baixo potencial de complexidade, como o P20 (TT: 637), P14 (TT: 643) e P21 (TT: 669) apresentam pouca variação lexical especialmente pela repetição de palavras. No quadro abaixo, apresentamos

fragmentos desses textos destacando, em negrito, algumas repetições (fato que reduz a variedade lexical) com a finalidade de ilustrar os valores positivos de TT:

Quadro 47 – Trechos de textos do livro *Pre-intermediate* com valores indicativos de baixa complexidade por TT

TEXTO	TRECHOS
P20 (TT: 637)	<p>... police in Sydney, Australia, received several phone calls about shouting and loud noises that were coming from an apartment in a suburb of the city...</p> <p>The callers had heard a woman screaming, a man shouting...</p> <p>A police car went to the apartment immediately...</p> <p>The officer told the man that his neighbours had heard shouting and screaming.</p> <p>'What about the woman who was screaming?'</p> <p>The police looked around the apartment and confirmed that nobody was hurt.</p>
P14 (TT: 643)	<p>Do you know why the queues at the other checkouts in the supermarket always seem to move faster than yours?</p> <p>... because you only notice how fast the other queues are moving when yours is moving slowly...</p> <p>If your queue moves fast, then you won't notice the slower queues at all...</p> <p>If there are three queues in the supermarket and you join the middle one, there is a two in three chance that one of the other queues will be the fastest...</p>
P21 (TT: 669)	<p>Every day he used to sit with his friends and talk about other people.</p> <p>And his friends told their wives, who told their neighbours, who told their friends.</p> <p>... she gave him a feather pillow...</p> <p>'Take the pillow to the top of the hill, then cut it open and release all the feathers...</p> <p>He cut open the pillow. All the feathers flew out...</p> <p>... go back up the hill with the empty pillow and refill it with the feathers.</p> <p>'The feathers have blown everywhere.'...</p> <p>'Your gossip is like the feathers,'...</p>

Fonte: Elaboração própria

Como podemos perceber pelos trechos acima, nos textos com melhores resultados de TT a repetição de palavras é frequente. Ainda que em determinados momentos fosse possível substituir as palavras destacadas por sinônimos ou pronomes, os autores optaram pela repetição dos substantivos. Conforme Silva (2018) demonstrou, uma das estratégias de acessibilidade mais eficientes é a recorrência de palavras quando um sinônimo ou pronome estiver muito distante da referência original.

Entretanto, é preciso destacar que o texto P14, por exemplo, foi classificado por outras métricas como tendo complexidades médias e altas. Em partes anteriores da nossa análise já havíamos destacado algumas características desse texto que podem ser compreendidas como potencialmente complexas. Sendo assim, enfatizamos a importância (conforme já repetimos inúmeras vezes) de utilizar

múltiplas métricas para análise de CT. Caso tomássemos a métrica TT como única indicativa de complexidade, poderíamos acreditar que o texto P14 é acessível, uma vez que o resultado dessa métrica o classificou como tendo o segundo melhor resultado relativos à relação entre complexidade e variação lexical.

Para concluir a análise de potencial de complexidade de léxico e de semântica por meio de uma perspectiva mais global, comparamos os valores dos índices dos textos do livro *Pre-intermediate* com as médias das métricas dos textos simples e complexos do *corpus* comparativo, como apresentaremos na próxima seção.

8.5.2.1.6 Análise contrastiva das métricas de léxico e semântica de P com os de LIS e LIC

O quadro abaixo (Quadro M) nos auxilia a visualizar se os valores métricos dos textos P se aproximam mais da média desses índices dos textos simples ou dos textos complexos do nosso *corpus* comparativo. Essa comparação (quando somada à interpretação individual das métricas), complementa a análise dos potenciais de complexidade do nosso *corpus* de estudo.

Quadro 48 - Comparativo de métricas de léxico e semântica entre os textos do *corpus* comparativo e os textos do livro *Pre-intermediate*

	PS LIS	PS P	PS LIC	IA LIS	IA P	IA LIC	FPC LIS	FPC P	FPC LIC
P1	Valor médio PS: 4,07	3,51	Valor médio PS: 3,55	Valor médio IA: 267,18	264,05	Valor médio IA: 367,08	Valor médio FPC: 578,87	593,82	Valor médio FPC: 563,43
P2		4,71			248,83			590,67	
P3		4,06			285,57			587,84	
P4		5,09			255,13			592,56	
P5		4,62			310,24			579,35	
P6		4,3			300,31			588,93	
P7		4,18			361,96			578,26	
P8		4,28			296,91			584,38	
P9		4,39			307,66			585,58	
P10		4,39			302,85			591,47	
P11		3,99			343,27			571,4	
P12		4,17			337,44			587,03	
P13		3,83			271,45			593,12	
P14		3,8			313			586,05	
P15		3,2			304,76			575,15	
P16		3,85			295,47			584,05	
P17		3,7			298,41			566,95	
P18		4,06			307,44			584,5	
P19		3,49			262,62			590,67	

P20		4,26			268,33			583,53	
P21		4,54			269,15			585,06	
P22		4,23			300,27			571,78	
Média		4,12			295,69			584,19	

	CLP LIS	CLP P	CLP LIC	TT LIS	TT P	TT LIC
P1		70,19			785	
P2		42,86			706	
P3		90,32			795	
P4		33			730	
P5		78,23			690	
P6		75,17			705	
P7		35,94			712	
P8		46,81			823	
P9		93,82			685	
P10		37,45			765	
P11	Valor médio CLP: 77,49	68,44	Valor médio CLP: 49,22	Valor médio TT: 661	750	Valor médio TT: 734
P12		28,43			764	
P13		19,22			684	
P14		5,05			643	
P15		89,97			808	
P16		94,74			684	
P17		76,42			846	
P18		42,07			739	
P19		30,15			725	
P20		89,97			637	
P21		83,4			669	
P22		81,86			774	
Média		59,71			733	

Fonte: Elaboração própria

Em um primeiro momento enxergamos que, visualmente, as três primeiras métricas de léxico e semântica (PS, IA e FPC) se aproximam mais dos valores dos textos simples (células destacadas em verde), enquanto as duas últimas (CLP e TT), dos textos complexos.

A interpretação do quadro nos permite observar que:

a. 77,2% dos textos P têm valores de PS mais próximos à média dos textos simples, enquanto 22,7%, dos textos complexos. Além da quantidade de textos, a média de PS dos textos P (4,12) também é mais próxima à média dos textos LIS (4,07) do que dos textos LIC (3,55). Esses números, com a interpretação já realizada sobre a métrica em questão, auxiliam a concluir que o comportamento dos textos P em relação à presença de palavras polissêmicas é semelhante ao que ocorre nos textos simples e, portanto, o potencial de acessibilidade é alto;

b. A respeito da IA, 86,3% dos textos resultam valores próximos aos textos LIS e apenas 13,6% aos textos LIC. A média de IA para os textos P (295,69) também é mais próxima à média de LIS (267,18) do que de LIC (367,08), sugerindo, de forma similar ao índice PS, que o léxico dos textos do livro *Pre-intermediate* também pode ser interpretado como potencialmente acessível;

c. Dentre as métricas de léxico e semântica, a FPC é a que mais demonstra aproximação à média de LIS, visto que 95,4% dos textos resultam em valores próximos aos de LIS e apenas 4,5% aos de LIC – ou seja, dentre os 22 textos do livro *Pre-intermediate*, 21 são similares à composição dos textos simples no que diz respeito à presença de palavras de conteúdo familiares, enquanto apenas 1 se difere (embora, se observarmos os números, possamos perceber que a diferença é mínima). Além da quantidade de textos, a média de FPC para os textos P (584,19) também é mais próxima à média de LIS (578,87) do que de LIC (563,43);

d. Embora 54,5% dos textos P tenham valores de CLP mais próximos à média dos textos LIS, a média de CLP dos textos P (59,71) tem valor mais próximo à média dos textos LIC para essa métrica (bem como 45,5% dos textos). Ainda que esses dados sinalizem que algumas instâncias podem ser potencialmente complexas, é preciso lembrar que os textos em análise estão no livro de um nível pré-intermediário e, dessa forma, a acessibilidade deve ser um pouco menor do que os textos presentes no nível elementar a fim de promover manutenção linguística. Além disso, conforme explicamos na análise de CLP apresentada anteriormente, nossos estudos a respeito dessa métrica nos permitem concluir que ela não deve ser tomada como referência autoritária para estimar complexidade lexical e semântica;

e. A média de TT⁶⁷ dos textos P (733) se aproxima muito mais da média dos textos LIC (734) do que dos textos LIS (661). Além disso, 68,1% dos textos têm valores mais próximos aos dos textos complexos, enquanto 31,9% dos textos simples. Sendo assim, a variedade lexical dos textos do livro *Pre-intermediate* pode sugerir que essa característica seja compreendida como fator de potencial de complexidade – embora seja necessário atentar para a interpretação da métrica apresentada anteriormente.

⁶⁷ Lembrando que valores mais altos de TT indicam maior complexidade.

Em suma, a respeito do léxico e da semântica nos textos P, podemos observar que houve progresso de complexidade em algumas características, mas que os indícios de baixa acessibilidade se referem a instâncias particulares.

No que diz respeito ao tipo de palavra apresentada, seja por polissemia, por relação à idade de aprendizado ou por familiaridade, os textos P são - em grande maioria - redigidos com léxico adequado para o nível pré-intermediário. Por outro lado, essa adequação pode ser um sinalizador para que estratégias de acessibilidade sejam adotadas para tornar o vocabulário dos textos mais complexos para alunos mais avançados.

As características que alertam para indícios de potencial de complexidade são relacionadas à concretude e à variação lexical, mas nossa análise nos permitiu concluir que embora a variação possa ser um fator complicador, os resultados e a leitura dos textos não indicam altíssima complexidade. Para casos específicos onde estratégias de customização sejam necessárias para aumentar a acessibilidade dos textos (como para alunos com limitações linguísticas e/ou estudantes neuroatípicos), é possível optar por estratégias que reduzam a variedade lexical (como a repetição de palavras, sobre a qual falaremos posteriormente no capítulo de orientações). A respeito da concretude, voltamos a enfatizar que, conforme explicamos na análise da métrica, os resultados não devem ser tomados como referências máximas para interpretação de complexidade.

Na próxima seção, continuamos nossa análise dos textos do livro *Pre-intermediate* com a interpretação das métricas de sintaxe.

8.5.2.2 Análise da complexidade sintática dos textos P

No primeiro contato com o Quadro 49 abaixo, podemos perceber, ainda que visualmente, uma clara diferença entre os valores das métricas sintáticas dos textos P para os textos E. Para que possamos corroborar que os textos do nível pré-intermediário apresentam, de fato, maior complexidade sintática do que os textos do nível elementar, precisamos analisar cada um dos índices individualmente (como temos feito até então).

Quadro 49 - Valores das métricas de sintaxe dos textos do livro *Pre-intermediate*

	SINTAXE				
	NT	SS	SIMS	CR	IF
P1	87,08	70,54	124	70,19	81,32
P2	99,52	33,36	74	42,86	85,13
P3	87,29	48,01	148	90,32	78,1
P4	80,23	44,83	88	33	76,37
P5	68,79	35,57	101	78,23	75,25
P6	78,23	22,96	82	75,17	70,57
P7	46,81	63,68	72	35,94	71,02
P8	67,36	59,87	82	46,81	76,37
P9	41,68	81,86	89	93,82	78,15
P10	64,43	51,6	61	37,45	66,46
P11	60,26	42,86	95	68,44	72,3
P12	81,33	31,92	73	28,43	69,57
P13	97,67	81,59	113	19,22	88,2
P14	61,41	47,61	83	5,05	83,55
P15	47,21	51,99	121	89,97	68,66
P16	61,03	64,06	101	94,74	73,86
P17	44,43	63,31	97	76,42	64,32
P18	93,45	50	85	42,07	77,83
P19	51,6	62,93	111	30,15	65,68
P20	85,31	62,17	95	89,97	81,03
P21	95,73	58,32	98	83,4	93,89
P22	76,42	37,83	76	81,86	71,39
MÉDIA	71,69	53,04	94	35,63	75,86

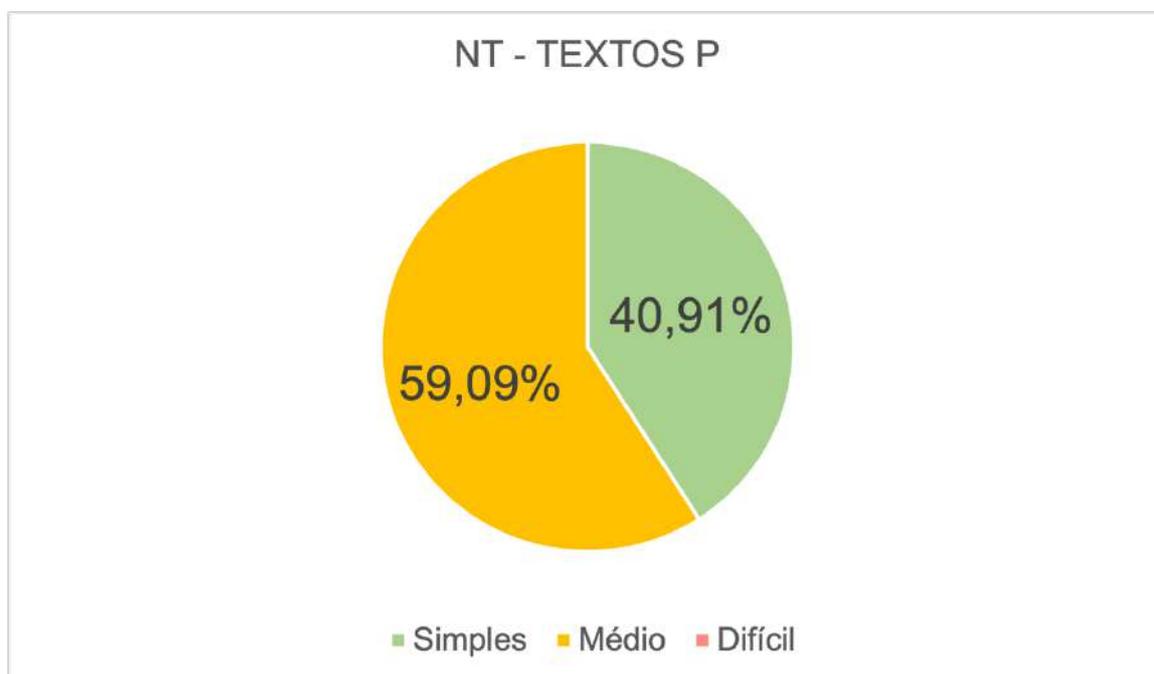
Fonte: Elaboração própria

Legenda: NT: Narratividade | SS: Simplicidade sintática | SIMS: Similaridade Sintática | CR: Coesão Referencial | IF: Índice Flesch

Seguindo a ordenação que temos adotado até o momento, iniciaremos a análise da sintaxe dos textos P por meio da interpretação da métrica NT.

8.5.2.2.1 Análise da métrica NT dos textos P

Por intermédio do gráfico abaixo (Gráfico 29), podemos relatar que a NT classifica 40,91% dos textos como simples e 59,09% como médios, sem nenhuma ocorrência de indicação de alta complexidade por essa métrica.

Gráfico 29 - Classificação percentual de complexidade de NT do livro *Pre-intermediate*

Fonte: Elaboração própria

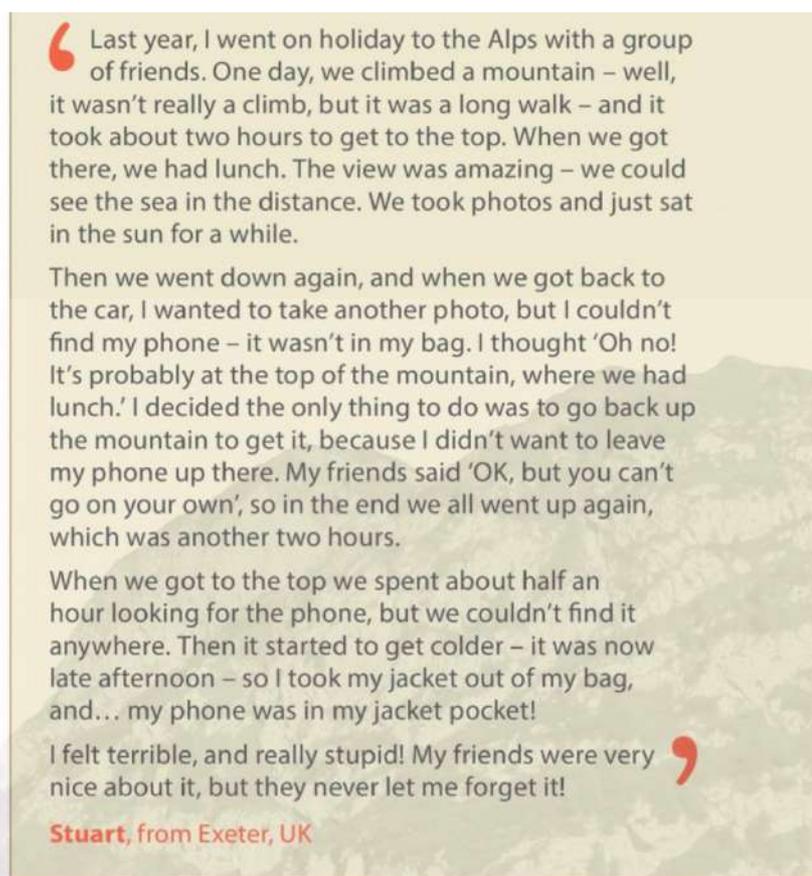
Ao lembrarmos que os textos P fazem parte de um livro didático voltado para um nível pré-intermediário, os valores indicados pelo gráfico acima podem ser interpretados como positivos, uma vez que, pelo nível, a complexidade não poderia ser exclusivamente baixa (caso contrário, a manutenção linguística poderia ser prejudicada).

O texto P indicado como mais simples pela NT é o P2 (NT: 99,52%), enquanto o mais complexo, o P9 (NT: 41,68%) – sobre os quais falaremos a seguir.

O primeiro fato importante a ser mencionado em relação ao texto P2 é sua posição no livro. Conforme explicamos anteriormente, uma vez que os materiais têm uma finalidade pedagógica, o ideal seria termos progressão constante da complexidade de todos os traços (fato que, até então, percebemos que não ocorre com tanta regularidade). Sendo assim, o segundo texto do livro ser indicado como o mais acessível pela NT é um sinal positivo. Para que possamos compreender a indicação da métrica, é necessário que o analisemos holisticamente.

O texto P2 é, de fato, uma narrativa. Esse texto costuma ser uma leitura fácil para os alunos em sala de aula. A demarcação de personagens é clara e o desenvolvimento lógico-temporal é fácil de ser acompanhado, como podemos observar pelos trechos do texto exibido pela imagem abaixo (Figura 32):

Figura 32 – Texto do livro *Pre-intermediate* com valor de alta acessibilidade por NT



Fonte: Latham-Koenig, Oxenden e Lambert (2019, p. 14)

A leitura do texto nos permite observar que o desencadeamento lógico é simples e que existem marcadores de temporalidade e de localização claros ('*Last year*', '*then*', '*so*' etc.). Sendo assim, o valor alto de NT é um reflexo verossímil dessa característica do texto P2.

Além do P2, outros textos resultaram valores altos de NT. No quadro abaixo (Quadro 50), destacamos alguns trechos desses textos com a finalidade de ilustrar suas estruturas narrativas (que servem como indicativos de alta acessibilidade):

Quadro 50 - Trechos de textos do livro *Pre-intermediate* com valores altos de NT

TEXTO	TRECHOS
P13 (NT: 97,67%)	<p>"I've found where she lives," said the detective quietly.</p> <p>"Here is the address." Hartley took the piece of paper. On it were the words "Vivienne Arlington, No. 341 East 49th Street?"</p> <p>"She moved there a week ago," said the detective. "I can follow her if you want. It will only cost you \$7 a day and expenses."</p>
P21 (NT: 95,73%)	<p>One day a wise old woman in the town asked to speak to the man. When he arrived, she gave him a feather pillow. 'Take the pillow to the top of the hill, then cut it open and release all the feathers, she said. But why?' he</p>

	<i>asked. 'Just do as I say,' she answered, 'and come back tomorrow.' So the man went to the top of the hill. He cut open the pillow. All the feathers flew out, and the wind carried them in all directions.</i>
P18 (NT: 93,45%)	<i>When he was 14 he broke his arm in my PE lesson. He was in hospital for a week, and we collected money and bought him a CD, so he was probably beginning to get interested in music.</i>
P3 (NT: 87,29%)	<i>That summer, I went on a camping holiday with my parents in the south of France. One night, I went to a nightclub on the beach and I met a young Frenchman, and we fell in love. When I got home, I immediately started looking for a job in Paris.</i>
P1 (NT: 87,07%)	<i>Charlotte and Clint are very close. They often go out together in the evening, they're planning a trip to India for later in the year, and they're both Bruce Springsteen fans. But Charlotte thinks that Clint needs a new partner, and he agrees.</i>
P20 (NT: 85,31%)	<i>An Argentinian family was driving when the husband, Walter, made an unfortunate mistake. He stopped at a petrol station, filled up the car with petrol, and went to the toilet. But when he drove off, he didn't notice that his wife Claudia wasn't in the car.</i>
P12 (NT: 81,33%)	<i>When his course finished he went to Madrid for the weekend to do his tests. A teacher called Paula met him there and gave him a mark out of ten for each test and then a final mark for everything.</i>

Fonte: Elaboração própria

Como o quadro acima nos auxilia a perceber, diversos recursos linguísticos marcam os textos como narrativas, tais como marcadores lógico-temporais, alta presença de verbos e diálogos. Logo, os valores de NT refletem que, de fato, esse aspecto textual pode ser considerado como indicativo de acessibilidade para esses textos.

Já o texto com valor mais baixo de NT, P9 (NT: 41,68%), não é uma narrativa. Ao analisarmos os textos, percebemos que sua divisão em tópicos não aproxima a leitura do texto a uma história, conforme podemos observar por meio da imagem a seguir (Figura 33) que exhibe um trecho do texto P9:

Figura 33 – Trecho de texto do livro *Pre-intermediate* com valor de baixa acessibilidade por NT

How do we know what to believe when doctors give us advice about what to drink and what not to drink? Every week, it seems, a scientific study contradicts last week's research. **Jeremy Laurance** reviews the latest medical research into drinks.

ALCOHOL

GOOD: For the **heart**, in small amounts, according to some studies. Two small glasses of red wine a day can cut the risk of heart attack by 30%, especially for middle-aged men.

BAD: For the **liver**. Liver disease is increasing. Some studies also say that one drink a day increases a woman's risk of all types of cancer by 6%. Other recent studies suggest that no alcohol at all is the ideal for both men and women.



Fonte: Latham-Koenig, Oxenden e Lambert (2019, p. 43)

Como o trecho acima nos permite observar, a estrutura do texto não é similar a uma narrativa. Além da formatação em tópicos, os conselhos sobre os benefícios e desvantagens de cada bebida (tema do texto) inicia com uma frase com zeugmas, omitindo o sujeito e o verbo das frases – característica que impactou nos valores de NT.

Ainda sobre o texto P9, observamos que além dos valores de NT indicarem potencial de complexidade, o vocabulário desse texto pode ser indicativo de complexidade, especialmente porque apresenta diversas unidades lexicais relacionadas aos órgãos do corpo humano e a problemas de saúde. Entretanto, é válido apontar que muitas dessas palavras fazem parte do vocabulário de controle da unidade na qual o texto é apresentado.

Além do texto P9, destacamos no quadro abaixo (Quadro 51) outros textos (e alguns trechos de exemplo) cujos valores de NT são indícios de complexidade média:

Quadro 51 – Trechos de textos do livro *Pre-intermediate* com valores médios de NT

TEXTO	TRECHOS
P17 (NT: 44,43%)	<i>In 2016, at the Olympics in Rio de Janeiro, the German men's football team lost to the host nation in the final. Brazil won their first ever Olympic gold medal in the event and the local fans were delighted.</i>
P7 (NT: 46,81%)	<i>New words are created in many different ways. We can make a new word by combining two words, like gastropub (gastronomy + pub) or vlog (video + blog).</i>

P15 (NT: 47,21%)	<i>Janet Leigh (1927-2004) was one of film director Alfred Hitchcock's favourite actresses, and was in more than 50 films and many TV series. Her daughter Jamie Lee Curtis (1958-) is also a successful actress.</i>
P19 (NT: 51,6%)	<i>It seems that it is not uncommon for people who are unrelated to look almost identical. Niamh Geaney, from Dublin, and two friends were so interested in trying to find their 'twins' that they set up an online project called Twin Strangers.</i>
P5 (NT: 68,79%)	<i>In this game there is a wheel with 25 letters of the alphabet (not including 'X'). The aim of the game is to complete the alphabet wheel by saying the correct word for each letter in five minutes. The presenter gives the contestant definitions for words starting with the letters, for example, 'A - the place where you catch a plane' = 'Airport'.</i>

Fonte: Elaboração própria

Os trechos destacados no quadro acima auxiliam a compreender os resultados de NT. Com exceção do texto P17, os demais textos apresentados no quadro apresentam características mais descritivas do que narrativas. A estruturação das frases (especialmente pela presença de voz passiva e pelo fato de os sujeitos serem, em muitos casos, frases verbais) diferencia a composição dos textos de narrativas com personagens claros e com marcações lógico-temporais evidentes. Além disso, em muitos desses textos há a presença do que chamamos de “*dummy subject*” (quando há ausência de um antecedente explícito para pronomes), característica que não é frequente em narrativas.

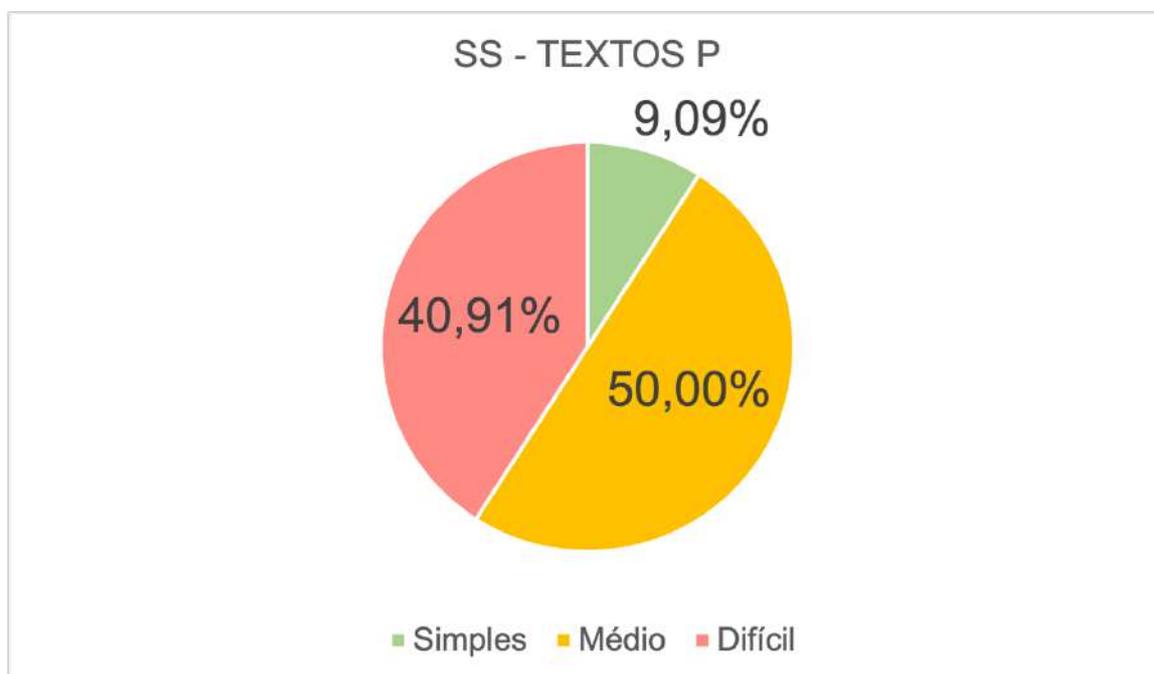
Conforme destacamos no início desta seção, a indicação de que 59,09% dos textos apresentam NT consideradas como indícios de complexidade média não é um fator de alerta tão grande. Uma vez que os textos fazem parte de um nível intermediário, esse valor é coerente com a progressão linguística. Obviamente, os valores de NT podem servir de marcadores para adoção de estratégias de simplificação quando professores de ILA julgarem necessário. Ao mesmo tempo, também enfatizamos que, na nossa experiência como docentes, é extremamente importante oportunizar o contato dos alunos com outras tipologias e gêneros textuais que não, necessariamente, narrativas – especialmente alunos universitários.

Na próxima seção, analisaremos a métrica SS.

8.5.2.2.2 Análise da métrica SS dos textos P

Os indicativos da SS nos alertam que grande parte dos textos podem ser considerados complexos por essa métrica (40,91%), enquanto 50% e 9,09% são classificados como indícios de complexidade média e baixa, respectivamente, conforme indica o Gráfico 30 abaixo:

Gráfico 30 - Classificação percentual de complexidade de SS do livro *Pre-intermediate*



Fonte: Elaboração própria

Diante do fato de que os textos são voltados para estudantes de nível pré-intermediário, 50% dos textos serem considerados médios é uma informação pertinente ao nível. Entretanto, o ideal seria que a outra metade estivesse dividida de forma equilibrada entre as classificações de complexidade média e alta.

Visto que, neste ponto da pesquisa, já compreendemos que os valores de SS demonstram maior acessibilidade quando as frases do texto seguem formações sintáticas simples, não consideramos necessário direcionar nossa atenção para os textos que indicam baixa acessibilidade. Entretanto, tendo um percentual tão alto de textos que podem ser classificados como complexos pela SS, achamos importante observá-los para entender tais resultados.

Para essa observação, destacamos, no quadro abaixo (Quadro 52), textos cujos resultados de SS sejam indicativos de alta complexidade e trechos desses

materiais que ilustrem as formações sintáticas que podem ser consideradas complexas. Apresentaremos os textos partindo do mais complexo segundo a SS, para o segundo mais complexo e assim por diante:

Quadro 52 – Trechos de textos do livro *Pre-intermediate* com valores de alta complexidade por SS

TEXTO	TRECHOS
P6 (SS: 22,96)	<i>Psychologist Judi James, one of the organizers of the survey, said, 'When some people read their friends' posts and see their photos on Facebook, Twitter, and Instagram, they begin to feel jealous of them. They think that their friends are having a much more exciting life. So they invent details about their own weekend.'</i>
P12 (SS: 31,92)	<i>Many British people think 'I don't have to learn a foreign language because everyone speaks English nowadays'. This is partly true. In many multinational companies, for example, employees have to speak English as it is the company's official language of communication.</i>
P2 (SS: 33,36)	<i>When we got to the top we spent about half an hour looking for the phone, but we couldn't find it anywhere. Then it started to get colder - it was now late afternoon - so I took my jacket out of my bag, and... my phone was in my jacket pocket!</i>
P5 (SS: 35,57)	<i>To start the game the presenter reads a definition for a word which begins with A and the contestant has to quickly say the word. If the contestant can't think of the word, he or she says 'Pass'. When a contestant passes or gets a word wrong, the presenter goes on to the next letter.</i>
P22 (SS: 37,83)	<i>An Italian couple ran out onto the runway of Malta's international airport to stop a Ryanair jet from leaving for Italy without them, a Maltese court heard on Thursday. Matteo Clementi, 26, and Enrica Apollonio, 23, got stuck in terrible traffic on their way to the airport on Wednesday. When they arrived, the gate was closed and they were not allowed to board their flight back to Italy.</i>
P11 (SS: 42,86)	<i>On 28th August 1963, there was a famous march in Washington against racism. Mahalia was with King, who was going to give a five-minute speech. Before the day itself, he and his advisers decided what he was going to say.</i>
P4 (SS: 44,83)	<i>The light in the evening is also good, but there are a lot more people around. There are times, in the local market, for example, when having lots of people in the photo can be a good thing, but most of the time I prefer my photos to be less crowded.</i>
P14 (SS: 47,61)	<i>4. If you can, choose a checkout which is 'cash only'. Using cash is usually quicker than paying by card. 5. Don't think too much! Sometimes it's best just to join the queue with the fewest people.</i>
P3 (SS: 48,01)	<i>In 1972, I was living in London. I was in my early twenties, and I was working for an advertising agency. That summer, I went on a camping holiday with my parents in the south of France. One night, I went to a nightclub on the beach and I met a young Frenchman, and we fell in love. When I got home, I immediately started looking for a job in Paris. He was at university there - he was studying medicine. In the end, I found a job as a secretary with UNESCO, and I went to live there.</i>

Fonte: Elaboração própria

Por meio do quadro acima e da leitura dos textos, é fácil de identificar os motivos pelos quais a métrica SS destacou um grande percentual como potencialmente complexos. A maior parte dos textos P apresenta frases e parágrafos longos, deslocamentos de elementos de suas posições canônicas, uso de imperativo (que interfere nos valores de SS uma vez que o sujeito está omissos), presença de apostos e alta frequência de coordenação e subordinação que, também frequentemente, posicionam a frase coordenada ou subordinada antes da principal.

Pelo trecho do texto P6, podemos observar apostos, orações subordinadas alocadas antes das principais, presença de discurso indireto e orações coordenadas – características que modificam as estruturas sintáticas e, ao mesmo tempo, potencializam a complexidade dos textos.

Os traços supracitados também podem ser observados nos demais textos destacados pelo quadro acima. Em muitos deles, além de tais características, também podemos perceber deslocamento de adjuntos, presença de frases imperativas e utilização de gerúndios na posição de sujeitos – exemplos de recursos linguísticos que modificam a formação sintática canônica. O último caso mencionado (gerúndios como sujeito) tende a ser complexo para alunos brasileiros, visto que, em língua portuguesa, ao utilizarmos um verbo como sujeito de uma oração, esse verbo deve estar na forma infinitiva.

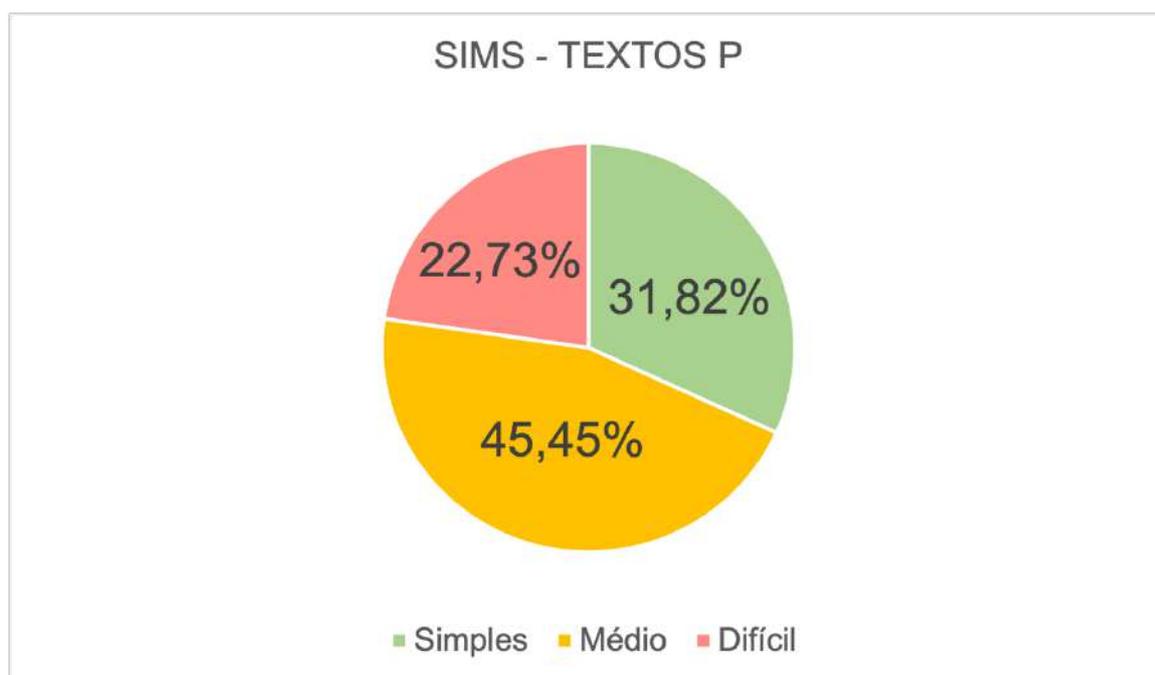
Todas as informações supracitadas nos permitem concluir que, de fato, os textos presentes no livro *Pre-intermediate* podem ser sintaticamente complexos para os nossos alunos em foco. Evidentemente, é importante que estudantes de ILA tenham contato com diversas estruturas textuais que configuram distintos gêneros textuais. Entretanto, acreditamos que a apresentação dessas diferentes formas deve ser gradual e equilibrada. Se observarmos a posição no livro dos textos destacados como sintaticamente complexos, podemos observar que muitos deles ocupam posições iniciais (como no caso dos textos P2, P3 e P4); sendo assim, é possível que a SS evidencie uma característica textual que seja verdadeiramente indício de complexidade. Diante dessa verificação, acreditamos que estratégias de simplificação possam ser adotadas, conforme sugeriremos, posteriormente, no capítulo de orientações desta tese.

Visto que os resultados de SS nos alertam para alto potencial de complexidade, a análise de SIMS poderá corroborar ou refutar tais conclusões, conforme veremos na próxima seção.

8.5.2.2.3 Análise da métrica SIMS dos textos P

Pelo gráfico abaixo (Gráfico 31) podemos observar que somado ao fato de a métrica SS ter destacado alto potencial de complexidade, os valores de SIMS apontam que a similaridade semântica entre os textos também deve ser observada com atenção, uma vez que 22,73% e 45,45% dos textos P apresentam valores de SS que indicam complexidade alta e média, respectivamente (e 31,82% como tendo complexidade baixa). Por outro lado, conforme comentamos no início da análise de SS, um percentual que aponte que metade dos textos possuem complexidade média e que a outra metade seja dividida entre complexidades baixas e altas pode ser um resultado coerente com o nível pré-intermediário.

Gráfico 31 - Classificação percentual de complexidade de SIMS do livro *Pre-intermediate*



Fonte: Elaboração própria

Visto que já compreendemos que a métrica SIMS analisa o quão similar ou diferente são as formações sintáticas de todas as frases de um texto, não consideramos necessária a observação dos textos destacados como simples; sabemos que esses textos possuem, então, formações sintáticas que se repetem ou que sejam altamente similares. Sendo assim, reservamos essa análise mais detalhada para os textos com resultados indicativos de alta complexidade.

Considerando que já havíamos destacado textos com valores baixos de SS na análise da métrica e que esses traços também são, possivelmente, os responsáveis pelas diferenças sintáticas nos textos, nosso primeiro procedimento de análise é observar quais dos textos que são destacados por SIMS como potencialmente complexos também são por SS. Pela leitura do quadro das métricas de sintaxe dos textos P (Quadro 49 apresentado no início do capítulo), constatamos que os textos sinalizados por SS e por SIMS como complexos são: P2, P12 e P22 – curiosamente um livro no início, um no meio e outro no final do livro.

Para ilustrar a baixa similaridade sintática nesses textos, selecionamos trechos de cada um deles, conforme o quadro abaixo (Quadro 53):

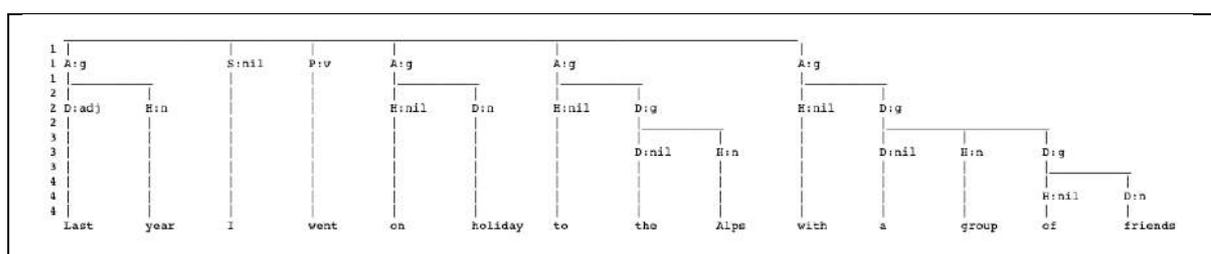
Quadro 53 – Trechos de textos do livro *Pre-intermediate* com valores de alta complexidade por SIMS

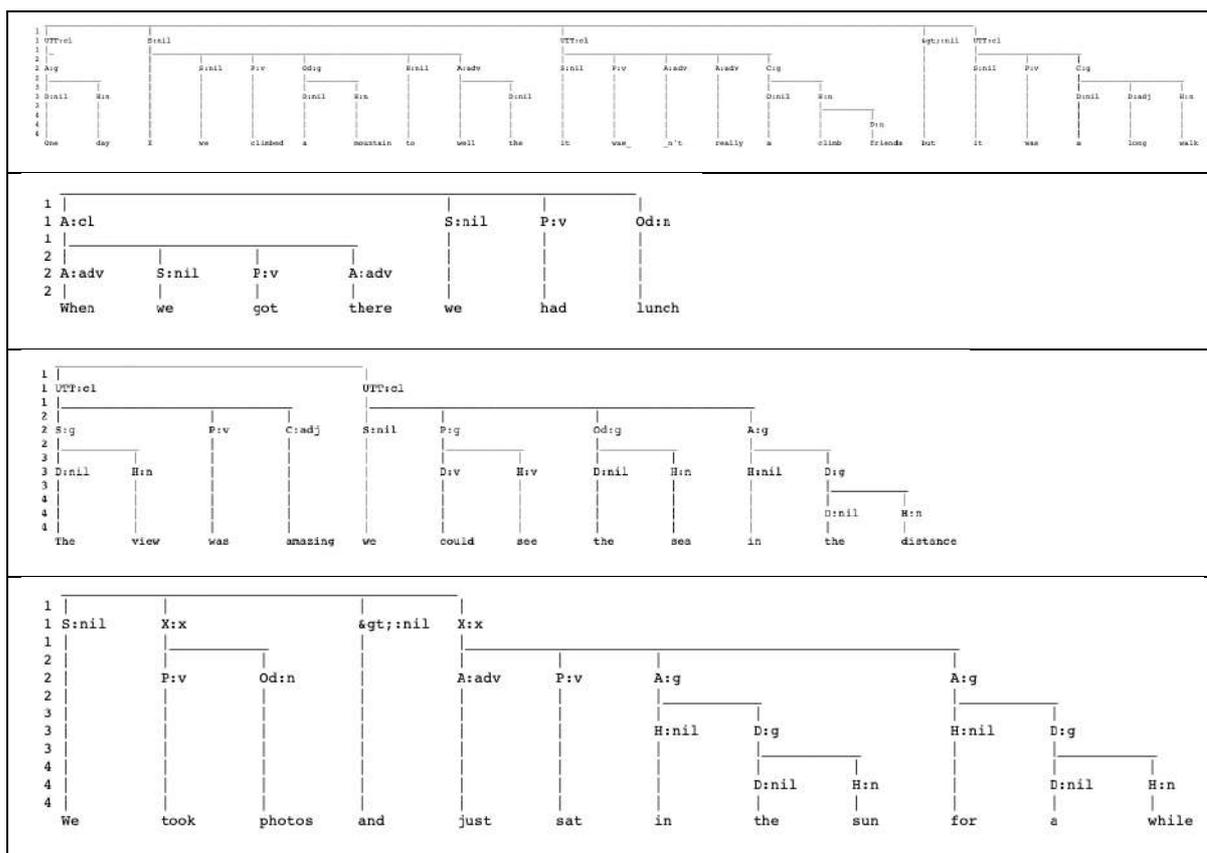
TEXTO	TRECHOS
P12 (SIMS: 73)	<i>A newspaper decided to find out by sending Max, one of its journalists, on an intensive language course. He then had to go to the country and do some 'tests' to see if he could 'survive' in different situations.</i>
P2 (SIMS: 74)	<i>Last year, I went on holiday to the Alps with a group of friends. One day, we climbed a mountain - well, it wasn't really a climb, but it was a long walk. [...] When we got there, we had lunch. The view was amazing - we could see the sea in the distance. We took photos and just sat in the sun for a while.</i>
P22 (SIMS 76)	<i>Tang, who works as a barman at the airport, had not played the lottery for the last ten years. He said that he planned to continue working and had not decided how to spend the money. But he said that he might leave \$20 notes in different places so that other people could find them and be lucky like him.</i>

Fonte: Elaboração própria

Ao realizarmos a análise sintática dos trechos acima, percebemos que mesmo em trechos tão pequenos, a formação das frases é distinta. Se observarmos os textos em suas totalidades, notamos que a variação de estruturação sintática é ainda maior. Na imagem abaixo (Figura 34) apresentamos as árvores sintáticas do trecho destacado do texto P2:

Figura 34 – Árvores de análise sintática do texto P2





Fonte: Elaboração própria por meio da plataforma visl.sdu.dk

No texto P2, já podemos observar uma grande diversidade de formações sintáticas. A primeira frase inicia com um adjunto deslocado, seguido por sujeito, verbo e três outros adjuntos preposicionados. Logo em seguida, podemos ver uma frase formada por um adjunto deslocado, sujeito, verbo e objeto direto. Posteriormente, há outro tipo de formação: adjunto adverbial, sujeito, verbo e predicativo do sujeito, seguido por uma frase subordinada posicionada antes da frase principal cuja estrutura é: adjunto adverbial, [sujeito, verbo, adjunto], [sujeito, verbo, objeto direto]. A frase a seguir tem a forma “sujeito, verbo de ligação, predicativo do sujeito”, seguida por outra cuja estrutura é “sujeito, verbo (modal e principal), objeto direto, adjunto”. Por fim, no exemplo apresentado no trecho, temos uma frase que é estruturada em: [sujeito, verbo, objeto direto], conjunção, advérbio, [verbo, adjunto preposicionado, adjunto preposicionado].

Se colocarmos as análises sintáticas das frases de forma resumida, podemos constatar sete diferentes formações, conforme podemos ver abaixo:

1. a, s + v + a + a + a;
2. a, s + v + od;
3. a, s + v + ps;

4. a, [s + v + a], [s, v, od];
5. s + v + ps;
6. s + v + od + a;
7. [s + v + od], conj., adv., [v + a + a].

Se em um parágrafo encontramos sete distintas formações sintáticas, podemos supor que a formação do texto tenda a ser altamente diferenciada no que diz respeito à sintaxe – fato que podemos corroborar por meio da leitura.

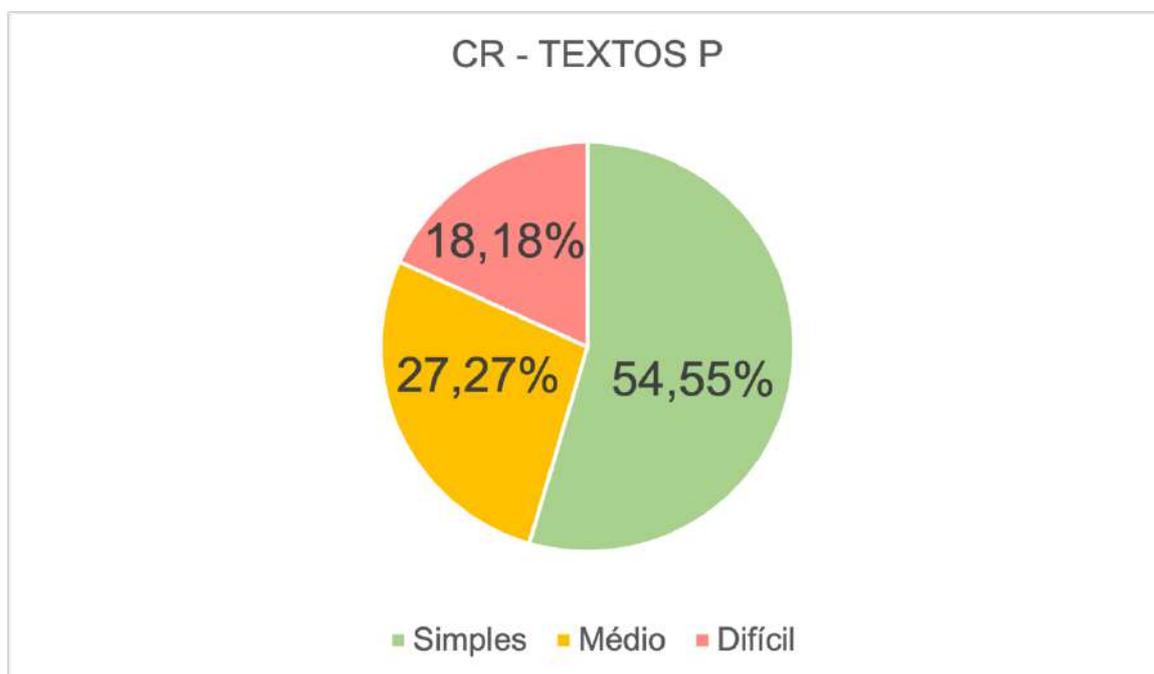
De forma similar, uma breve análise dos trechos apresentados no Quadro 53 acima nos mostra frases com formações sintáticas únicas e diferentes das demais – característica que se repete não só nos textos e trechos destacados no quadro, mas nos demais textos do livro *Pre-intermediate* (embora com menores variações nos textos que foram indicados como médios em complexidade).

Por fim, a interpretação dos valores de SIMS dos textos P somada às conclusões apresentadas a respeito da métrica SS nos permite afirmar que existem textos nos quais a sintaxe pode ser entendida como um fator de alto potencial de complexidade. Diante disso, é válido que professores atentem para os textos classificados como complexos e adotem estratégias de simplificação quando tiverem por finalidade adequar os textos para determinados estudantes (estratégias sobre as quais falaremos posteriormente).

Continuando nossa análise dos índices de sintaxe dos textos do livro *Pre-intermediate*, passamos para o estudo de CR.

8.5.2.2.4 Análise da métrica CR dos textos P

Por intermédio do gráfico abaixo (Gráfico 32), observamos que 54,55% dos textos são destacados como tendo valores de CR que os enquadram como simples, 27,27%, como médios e 18,18% como complexos:

Gráfico 32 - Classificação percentual de complexidade de CR do livro *Pre-intermediate*

Fonte: Elaboração própria

Se compararmos esses percentuais com o comportamento da CR nos textos do livro *Elementary*, podemos observar indícios de redução de complexidade, uma vez que, para os textos E, 55% eram destacados como complexos e 20% como médios. Em relação aos valores de E, os percentuais de P representariam inadequação ao levarmos em consideração progressão de complexidade (posto que os textos do nível pré-intermediário deveriam ser um pouco mais complexos do que os do nível elementar).

Contudo, ao analisarmos os valores de CR dos textos P sem comparação com o nível antecedente, podemos observar que os resultados são, até certo ponto, coerentes com o nível linguístico. Conforme já mencionamos para outras métricas, por serem textos de um nível pré-intermediário, espera-se que a complexidade não seja totalmente baixa, mas nem tão alta. Logo, a divisão dos textos conforme podemos observar pelo gráfico acima parece indicar adequação de nível.

Visto que já compreendemos como funciona a CR, entendemos que a leitura dos textos marcados como “simples” facilita a construção de relação entre pronomes e substantivos, verbos e sujeitos e qualquer outro tipo de referência realizada por recursos linguísticos. É possível, também, que esses textos utilizem estratégias como repetição de substantivos para evitar que a substituição de um substantivo por um pronome possa causar ambiguidade de referência e/ou que evitem grandes

distâncias entre pronomes e os substantivos aos quais se referem. No trecho do texto P16 abaixo – considerado o mais acessível pela CR (com valor de 94,74) – podemos observar algumas dessas estratégias. Destacamos em cores palavras que se relacionam (ou seja, todas as palavras em verde “apontam” para o mesmo referencial, todas as palavras em laranja se referem ao mesmo referente e assim por diante). Esses destaques nos auxiliam a observar a coesão textual:

Trecho do texto P16:

Men usually hit the ball so hard, especially when they serve, that their opponent can't return it.

In men's football, the players fall over all the time and act like babies. The women don't do that. I've read research that says that women hardly ever pretend to be hurt. And when they are hurt, they get up again 30 seconds faster than men.

Além da fácil relação entre sujeito e verbo, entre pronome e substantivo e entre pronome demonstrativo e frase antecedente, no texto P16 também encontramos várias instâncias do pronome pessoal de primeira pessoa – característica que facilita a coesão referencial, visto que esses pronomes são dêiticos.

Logo, ao compreendermos bem o que os resultados positivos de CR nos indicam, acreditamos que – para as finalidades da nossa análise – é mais válido que nossos focos de observação sejam os textos com resultados indicativos de complexidade alta por essa média. Para isso, apresentamos no quadro abaixo (Quadro 54) os textos indicados como potencialmente complexos e alguns trechos sobre os quais discutiremos posteriormente:

Quadro 54 – Trechos de textos do livro *Pre-intermediate* com valores de alta complexidade por CR

TEXTO	TRECHOS
P14 (CR: 5,05%)	<i>Lots of people with a few items can be slower than a few people with lots of items. Sometimes it's best just to join the queue with the fewest people.</i>
P13 (CR: 19,22%)	<i>Hartley paid the man. Then he left his office and took a tram to Broadway. After walking a short distance he arrived at the building that he was looking for. He went up the stairs, into her apartment, and saw her standing by the window.</i>
P12 (CR: 28,43%)	<i>That's been true for a long time. In any city around the world you can hear British tourists asking for the restaurant menu in English.</i>

	<p><i>Many British people think 'I don't have to learn a foreign language because everyone speaks English nowadays'.</i></p> <p><i>Children only have to learn a language until they are 14. After that, they don't have to continue if they don't want to. Thirty per cent of young people say that they don't want to carry on with a foreign language because 'it's too difficult.'</i></p>
<p>P19 (CR: 30,15%)</p>	<p><i>It seems that it is not uncommon for people who are unrelated to look almost identical. Niamh Geaney, from Dublin, and two friends were so interested in trying to find their 'twins' that they set up an online project called Twin Strangers. Very quickly, Niamh found a remarkably similar-looking stranger who lived just a few miles away.</i></p> <p><i>Jim Springer's parents told him that he had an identical twin brother, but that he was dead.</i></p>

Fonte: Elaboração própria

No trecho do texto P12 apresentado no quadro acima, a primeira frase (“*That's been true for a long time.*”) é, de fato, a primeira frase do texto. O pronome demonstrativo ‘*that*’ se refere à pergunta feita no título do texto (“*Are the British really so bad at learning languages?*”). Ainda que, para nós - leitores proficientes - a conexão entre o pronome demonstrativo e o título seja clara, não podemos supor que também será para um aluno do nível pré-intermediário. Sendo assim, o fato de o texto iniciar com o pronome ‘*that*’ pode ser uma característica potencialmente complexa.

Ainda sobre o texto P12, podemos observar alguns traços que impactam, negativamente, na coesão referencial, como destacamos a seguir:

- Na frase “*you can hear British tourists asking for the restaurant menu in English*”, o uso do presente particípio como objeto direto pode ser potencialmente complexo visto que, alunos do nível pré-intermediário, ainda não estudaram esse tipo de construção frasal e, geralmente, esperam encontrar um verbo ‘*to be*’ ligando um sujeito a um verbo no gerúndio;
- O uso do pronome ‘*I*’ na frase “*Many British people think 'I don't have to learn a foreign language because everyone speaks English nowadays'.*” pode causar ambiguidade, uma vez que os alunos podem compreender que ele se refere ao narrador do texto (em vez de compreender, pelo uso das aspas, que essa frase é um pensamento de cada cidadão britânico);
- No parágrafo “*Children only have to learn a language until they are 14. After that, they don't have to continue if they don't want to. Thirty per cent of young people say that they don't want to carry on with a foreign*

language because 'it's too difficult.', o pronome *'they'* pode ser interpretado erroneamente. Caso o estudante não consiga associar que *'young people'* é uma expressão utilizada para retomar *'children'*, é possível que compreendam que outros jovens digam que as crianças não querem continuar os estudos de uma língua estrangeira por ser muito difícil (em vez de compreender que esse *'they'* se refere aos 30% das crianças/jovens entrevistados).

Características semelhantes são observadas nos trechos dos textos apresentados no Quadro 54 acima. No texto P13, por exemplo, o pronome *'her'* retoma um substantivo que é mencionado três parágrafos acima. Já a referência do pronome *'he'* pode ser tanto *'Hartley'* quanto *'the man'*.

No texto P14, também podemos perceber alguns traços que podem causar dificuldades em coesão referencial. Na frase *"Lots of people with a few items can be slower than a few people with lots of items."*, por exemplo, a extensão da frase que ocupa a posição de sujeito é grande. É possível, então, que estudantes com limitações linguísticas associem o verbo *'can'* à palavra *'items'*, em vez de associar à frase *'lots of people with few items.'* Além desse exemplo, também existem ocorrências de *dummy subjects* no texto P14 que, por terem referências que não são explicitamente mencionadas nos textos, dificultam o processo de coesão referencial.

Já no texto P19, além da presença de *dummy subjects*, também podemos perceber que existem muitas relações entre pronomes (pessoais e relativos) e sujeitos (*people – who / Niamh – two friends – were – their – they / stranger – who*), característica que impacta na coesão referencial. Ainda que pouco provavelmente ocorra, o pronome *'he'* na frase *"Jim Springer's parents told him that he had an identical twin brother, but that he was dead."* pode ser relacionado ao Jim, em vez do seu irmão gêmeo, alterando totalmente, o sentido do texto.

Ainda que tais os apontamentos sobre os textos acima possam parecer *far fetched*⁶⁸, precisamos, como professores de ILA, antever todos os cenários possíveis de baixa acessibilidade textual para, então, adotar estratégias que sejam adequadas para aprimorar a compreensão (e, ao mesmo tempo, a manutenção linguística) dos nossos alunos. Silva, Moll e Perna (2021) enfatizam a importância de “nos colocarmos

⁶⁸ Expressão da língua inglesa com significado parecido à expressão “forçação de barra” ou algo implausível.

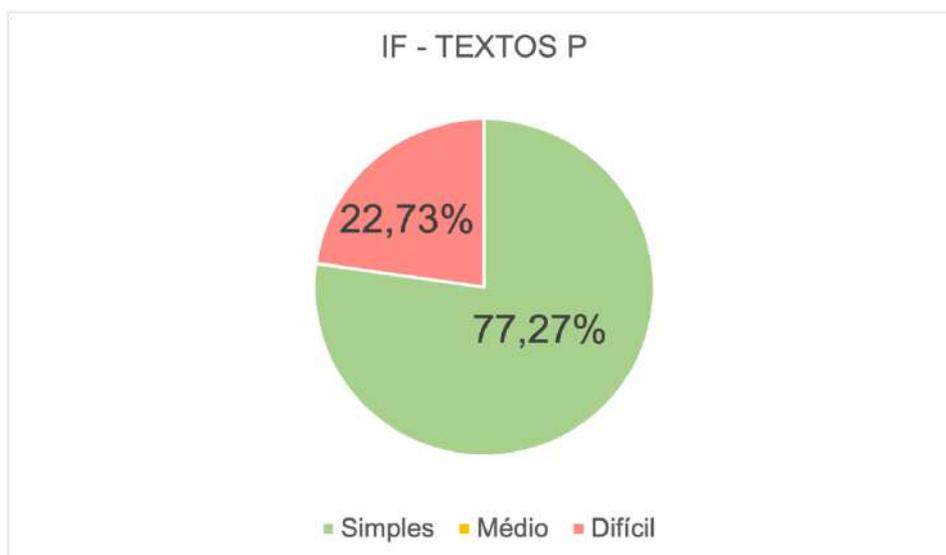
no lugar” dos nossos leitores (nossa audiência social) para que possamos prever dificuldades e adotar estratégias quando temos por objetivo potencializar acessibilidade textual.

Por fim, ainda que a maior parte dos textos tenham sido classificados como tendo valores de CR com baixo potencial de complexidade, podemos observar que, em algumas instâncias, a coesão referencial pode ser um indicativo de baixa acessibilidade. Evidentemente, professores de ILA podem utilizar as ocorrências de alta complexidade de CR (como no parágrafo apresentado como ilustração do texto P19, por exemplo) e os casos de frases e parágrafos com muitos pronomes (principalmente) para desenvolver o raciocínio lógico e referencial de seus alunos. Contudo, em cenários com alunos que tenham limitações linguísticas e/ou com alunos neuroatípicos, as múltiplas possibilidades de referência dentro de um mesmo texto (seja de forma global ou dentro de um mesmo parágrafo) ou a grande distância entre um pronome e seu substantivo podem ser características que dificultam a interpretação de um texto. Além disso, por limitações temporais (visto que, ao menos no contexto para o qual estamos realizando nossa pesquisa, os professores têm, geralmente, uma hora e trinta minutos para trabalhar com cada unidade), desviar a atenção do conteúdo-foco da unidade pode ser improdutivo e incoerente com as propostas.

Por fim, para que possamos concluir a análise das métricas relativas à sintaxe dos textos do livro *Pre-intermediate*, prosseguimos para as observações a respeito do IF.

8.5.2.2.5 Análise da métrica IF dos textos P

Concluindo a análise individual das métricas, o gráfico abaixo (Gráfico 33) mostra que 77,27% dos textos P são classificados como sendo de baixa complexidade pelo IF, enquanto 22,73%, como de alta complexidade:

Gráfico 33 - Classificação percentual de complexidade de IF do livro *Pre-intermediate*

Fonte: Elaboração própria

Nos exemplos apresentados para analisar as demais métricas, até então, já conseguimos observar que a extensão de alguns textos P é, em determinados momentos, longa (como no trecho utilizado para ilustrar a métrica CR do texto P19, por exemplo).

É interessante observar por meio do quadro de métricas dos textos P que o primeiro caso de IF indicativo de alta complexidade é o texto P10 e os quatro próximos, do texto P12 em diante. Isso nos indica que, em termos de extensão, a complexidade aumenta somente da metade do livro em diante – característica que pode ser percebida como positiva, uma vez que, pelo interesse didático, a complexidade deveria ser progressiva. Outro fato importante de destacar é que embora cinco textos possam ser considerados menos acessíveis pela IF, seus valores estão mais próximos daqueles que indicam complexidade média do que dos números que representam altíssimas complexidades.

No quadro abaixo (Quadro 55), apresentamos os textos considerados complexos por IF e trechos que ilustram os motivos pelos resultados dessa métrica:

Quadro 55 – Trechos de textos do livro *Pre-intermediate* com valores de alta complexidade por IF

TEXTO	TRECHOS
P17 (IF: 64,32)	<i>In 2016, at the Olympics in Rio de Janeiro, the German men's football team lost to the host nation in the final. Brazil won their first ever Olympic gold medal in the event and the local fans were delighted. As they celebrated, one very disappointed German player, Robert Bauer, decided to show 7 fingers to the fans, to remind them of the time Germany beat Brazil 7-1 in the 2014 World Cup semi-final.</i>

P19 (IF: 65,68)	<p><i>Two women recently met by chance at Bremen University, in Germany. One was English, and one was Irish, and they were both on Erasmus scholarships.</i></p> <p><i>Nothing unusual there, except for one thing. The two girls look identical. Their hair is the same colour and length, they're the same age and size, and when you see them together, in the photo they put on social media, you would think that they were identical twins. In fact, Cordelia Robert and Ciara Murphy are unrelated.</i></p>
P10 (IF: 66,46)	<p><i>Then they look at each possible problem and plan how to avoid it. So for the exam situation, they go to bed early and have a good night's sleep; they find out in advance exactly where the exam is; they eat a good breakfast, and take lots of pens and pencils, and a bottle of water; and they leave home early. That puts them in control, and it means that the exam will be better than for an optimist, who just thinks 'Oh, everything will be fine!' Because sometimes everything goes wrong, and it's good to be prepared.</i></p>
P15 (IF: 68,66)	<p><i>She has been in many different kinds of films, including the comedies <i>Trading Places</i>, <i>Fish Called Wanda</i>, and <i>True Lies</i>, for which she won a Golden Globe Award for Best Actress in a Musical or Comedy. She has also starred in the comedy-horror TV series <i>Scream Queens</i>. In one episode she recreated the famous scream from <i>Psycho</i>.</i></p>
P12 (IF: 69,57)	<p><i>Many British people think 'I don't have to learn a foreign language because everyone speaks English nowadays'. This is partly true. In many multinational companies, for example, employees have to speak English as it is the company's official language of communication.</i></p>

Fonte: Elaboração própria

Pela nossa experiência com os textos P, sabíamos, de antemão, que os resultados de IF não seriam indicativos de altíssima complexidade. Acreditamos, inclusive, que sobre extensão, alguns textos do livro *Pre-intermediate* poderiam ser um pouco mais extensos para oportunizar maior contato com exemplos gramaticais e de vocabulários que sejam os pontos-foco de cada unidade.

Mesmo analisando os textos cujos valores dessa métrica indicam menor acessibilidade, é possível observar que extensão não é a característica de maior dificuldade nesses textos (embora, em alguns momentos pontuais, seja possível adotar estratégias para reduzir ainda mais esses indícios de complexidade, como falaremos posteriormente).

No texto com valor de IF que indica maior complexidade textual (P17), podemos perceber que o fator que influenciou no resultado dessa métrica foi a extensão das palavras, uma vez que ele adota palavras relativamente extensas (como '*delighted*', '*celebrated*', '*disappointed*'). Embora o IF não seja motivo de alta preocupação a respeito da acessibilidade, é preciso lembrar que o P17 foi destacado por todas as métricas de léxico e semântica e por quatro (dos cinco) índices de sintaxe como

potencialmente complexo; sendo assim, sua extensão se soma às demais características que reduzem sua acessibilidade.

Ainda sobre o P17, é preciso destacar que muitas das palavras apresentadas no texto são nomes próprios (nomes de atletas ou de países) e que são, em grande maioria, extensos. Logo, essas unidades contribuem para que os valores de IF sejam mais baixos, embora a interpretação dessas palavras não seja necessariamente complexa, já que são facilmente identificadas como substantivos próprios e/ou parte do vocabulário dos alunos (visto que nomes de países são ensinados no nível *Elementary* e nas unidades iniciais).

No trecho do texto P10, podemos observar um exemplo de extensão que pode, de fato, potencializar a complexidade. Somente no texto destacado encontramos orações complexas (formadas por coordenação e subordinação) e usos de ponto-e-vírgula que contribuem para demarcar o quão grande as frases são. Além disso, também podemos perceber que a extensão das palavras é grande (embora essas unidades não sejam complexas).

Nos textos P12, P15 e P19 também percebemos características similares aos citados anteriormente. Ainda que nem todos tenham frases extensas, todos apresentam palavras relativamente longas, embora algumas sejam substantivos próprios.

Em síntese, acreditamos que a extensão (representada pelo IF) dos textos P é adequada na maior parte do *corpus* e que, inclusive, textos um pouco mais longos pudessem ser apresentados nesse nível – como no texto da unidade 3C, por exemplo, que poderia apresentar um texto mais longo e com mais exemplos de *relative clauses* (foco gramatical da unidade) para otimizar o insumo de aprendizado para os alunos.

Nos textos destacados como mais complexos, precisamos analisar se a extensão se soma a outras indicações de complexidade e, caso isso não ocorra, se os valores de IF não são consequência de substantivos próprios ou outras palavras que, embora longas, sejam simples. Em casos como no texto P17, por exemplo, algumas estratégias de simplificação alunos (como a divisão de orações em duas ou mais, sobre a qual falaremos posteriormente) podem ser adotadas quando tivermos por objetivo otimizar a acessibilidade dos textos para determinados.

Para concluir nossa análise dos textos P, seguimos na próxima seção para a comparação das métricas desses textos com os valores das médias de LIS e de LIC.

8.5.2.2.6 Análise contrastiva das métricas de sintaxe de P com os de LIS e LIC

No quadro abaixo (Quadro 56), conforme apresentamos anteriormente para as métricas de léxico e semântica, destacamos em verde os valores das métricas dos textos P que se aproximam da média de cada índice dos textos LIS (textos simples do *corpus* comparativo) e, em laranja, os resultados que se aproximam mais das médias de LIC (textos complexos do *corpus* comparativo):

Quadro 56 - Comparativo de métricas de sintaxe entre os textos do *corpus* comparativo e os textos do livro *Pre-intermediate*

	NT LIS	NT P	NT LIC	SS LIS	SS P	SS LIC	SIMS LIS	SIMS P	SIMS LIC
P1		87,08			70,54			124	
P2		99,52			33,36			74	
P3		87,29			48,01			148	
P4		80,23			44,83			88	
P5		68,79			35,57			101	
P6		78,23			22,96			82	
P7		46,81			63,68			72	
P8		67,36			59,87			82	
P9		41,68			81,86			89	
P10	Valor médio	64,43	Valor médio	Valor médio	51,6	Valor médio	Valor médio	61	Valor médio
P11	NT:	60,26	NT:	SS:	42,86	SS:	SIMS:	95	SIMS:
P12	72,82	81,33	33,43	75,59	31,92	46,10	134	73	79
P13		97,67			81,59			113	
P14		61,41			47,61			83	
P15		47,21			51,99			121	
P16		61,03			64,06			101	
P17		44,43			63,31			97	
P18		93,45			50			85	
P19		51,6			62,93			111	
P20		85,31			62,17			95	
P21		95,73			58,32			98	
P22		76,42			37,83			76	
Média		71,69			53,04			94	

	CR LIS	CR P	CR LIC	IF LIS	IF P	IF LIC
P1		70,19			81,32	
P2		42,86			85,13	
P3		90,32			78,1	
P4		33			76,37	
P5		78,23			75,25	
P6		75,17			70,57	
P7		35,94			71,02	
P8		46,81			76,37	
P9		93,82			78,15	
P10		37,45			66,46	

P11	Valor médio CR: 51,98	68,44	Valor médio CR: 26,68	Valor médio IF: 85,17	72,3	Valor médio IF: 44,17
P12		28,43			69,57	
P13		19,22			88,2	
P14		5,05			83,55	
P15		89,97			68,66	
P16		94,74			73,86	
P17		76,42			64,32	
P18		42,07			77,83	
P19		30,15			65,68	
P20		89,97			81,03	
P21		83,4			93,89	
P22		81,86			71,39	
Média		35,63			75,86	

Fonte: Elaboração própria

A leitura dos dados do Quadro R nos permitem concluir que:

a. Além da média de NT dos textos P ser mais próxima à média de LIS para esse índice, 77,2% dos textos também indicam maior proximidade aos valores indicativos de baixa complexidade e somente 22,8% dos resultados de maiores complexidades. Tais percentuais indicam que a maior parte dos textos apresenta estruturas e componentes similares a narrativas e, conseqüentemente, potencializam a acessibilidade e interpretação de informações;

b. No que diz respeito à SS, 63,6% dos textos P apresentam valores mais próximos à média dos textos complexos e 36,4% dos textos simples. A média para esse índice também sugere maior proximidade ao resultado médio dos textos complexos. Esses dados indicam que as formações sintáticas das frases dos textos P podem ser compreendidas como fatores de alta complexidade (visto que os textos LIC são acadêmicos e científicos nos quais podemos observar construções sintáticas distintas e complexas);

c. A análise comparativa dos valores de SIMS corroboram as indicações de complexidade sintática dos textos P, visto que além da média dessa métrica ser mais próxima da média dos textos LIC, 77,2% dos textos também apresentam valores muito mais próximos à média dos textos complexos. Somente 22,8% resultam em números similares à média de LIS. Dessa forma, tanto a SS quanto a SIMS indicam que os textos P podem ter traços de alta complexidade relacionados à sintaxe;

d. No quesito coesão referencial, embora 68,1% dos textos apresentem valores de CR mais próximos à média dos textos simples, a média dos textos P para essa métrica é muito mais próxima à média dos textos complexos. Então, ainda que apenas 31,9% dos textos sejam destacados como indicativos de alta complexidade pela CR,

seus valores são tão baixos que aproximam classificam a média como indício de baixa acessibilidade. Logo, ainda que na maior parte dos textos a coesão referencial não seja uma característica de alta complexidade, nos textos indicados por essa métrica como potencialmente complexos, esse traço textual tende a ser um dos maiores fatores de complexidade para esse grupo de textos e, para esses casos, estratégias de simplificação devem ser adotadas (sobre as quais falaremos posteriormente);

e. O IF, conforme já havíamos concluído na sua análise individual, indica que no quesito extensão textual, tanto a média da métrica dos textos P quanto 95,4% dos textos tem valores mais próximos à média dos textos LIS. Logo, a extensão não é percebida como uma característica potencializadora de complexidade.

Em síntese, nossa análise nos permitiu concluir que existem algumas características sintáticas que podem tornar os textos potencialmente complexos (especialmente para alunos com limitações linguísticas e/ou estudantes neuroatípicos).

Por outro lado, uma vez que os textos fazem parte de um livro voltado para um nível pré-intermediário, compreendemos que esses traços de complexidade devam ser levados em consideração para cenários específicos, uma vez que não parecem ser características que impedem totalmente a compreensão dos textos e, simultaneamente, contribuem para a manutenção linguística.

É preciso destacar, conforme mencionamos antes, que sobre a extensão textual, acreditamos que alguns textos poderiam oportunizar mais insumos de aprendizado; não somente por meio de textos mais longos, mas com mais exemplos dos pontos gramaticais e de vocabulários que são trabalhados em cada unidade.

Por intermédio da tabela que exhibe os valores métricos (tanto de sintaxe, quanto de léxico e semântica) dos textos P, é possível localizar quais textos merecem atenção e quais são as características desses textos que podem ser observadas e modificadas quando professores de ILA buscarem aumentar a acessibilidade. Para isso, é mais interessante que os docentes observem não somente uma métrica dos textos (embora seja possível fazê-lo, especialmente para orientar para estratégias de simplificação pontuais), mas que, ao analisar complexidade, observe as indicações de todas as métricas. Alguns textos (como o P7, P11, P15 e P17) resultam em mais métricas indicativas de alto potencial de complexidade do que índices que os classifique como acessíveis. Para esses textos, é preciso cuidado ao aplicar

estratégias de simplificação uma vez que a complexidade não é relatada apenas por um aspecto textual, mas por múltiplos. A menos, obviamente, que se deseje reduzir apenas um fator de complexidade, os textos devem ser observados na sua totalidade.

Na próxima seção, concluiremos nossa análise métrica por meio dos dados do último livro proposto para nossa pesquisa. Para isso, iniciaremos pelos índices de léxico e semântica dos textos do nível *Intermediate* voltando nossa atenção, principalmente, para os textos indicativos de alta complexidade.

8.5.3 RESULTADOS MÉTRICOS DOS TEXTOS DO LIVRO DE NÍVEL INTERMEDIATE

Nesta seção, faremos uma análise prévia e global das métricas dos textos presentes no livro *Intermediate*. Posto que os textos em questão são voltados para um grupo de alunos, supostamente, em nível intermediário, é aceitável que tenhamos maiores frequências de indicações de complexidade média e alta e – conforme previamente mencionado – traços de alta complexidade não tendem a impedir a leitura ou desservir a manutenção linguística para alunos desse nível.

Entretanto, ainda precisamos lembrar que embora os textos sejam primeiramente voltados para esse nível, é possível que tenhamos alunos que não tenham conhecimentos linguísticos que os classificariam como “intermediários” e que, ainda assim, estejam cursando disciplinas que adotem esse livro (fato que ocorre com grande frequência). Além disso, também é possível que estudantes neuroatípicos acessem os materiais e, portanto, estimar complexidade e observar traços que podem ser complicadores ainda é válido, mesmo que, em tese, estudantes de ILA que são verdadeiramente intermediários já tenham um repertório linguístico que os ajudaria a “contornar” as características textuais complexas com mais facilidade e, ao mesmo tempo, utilizar esses traços como insumo de aprendizagem.

Portanto, ainda que não tenhamos o objetivo de analisar com os mesmos detalhes as métricas e os textos do livro *Intermediate* como fizemos com os níveis *Elementary* e *Pre-intermediate*, acreditamos que essa análise prévia nos permitirá observar se, futuramente, é válido investigar os textos desses livros de forma mais minuciosa, especialmente para estudantes neuroatípicos.

Iniciaremos nossa análise global pelas métricas indicativas de complexidade lexical-semântica.

8.5.3.1 Análise prévia da complexidade lexical-semântica dos textos I

O quadro abaixo (Quadro 57) organiza e exhibe os valores das métricas relacionadas ao léxico e à semântica dos textos do livro *Intermediate*. Lembramos que as métricas estão destacadas em cores representativas de complexidade de acordo com a interpretação de cada um dos índices (verde: simples, amarelo: médio, vermelho: difícil) e que, além das cores, suas gradações também demonstram níveis de complexidade. Logo, verde escuro indica altíssima acessibilidade, enquanto verdes mais claros, menores acessibilidades em relação ao verde escuro (mas ainda indicativos de alta acessibilidade). Da mesma forma, amarelo escuro indica que o índice é considerado médio, mas mais próximo aos valores indicativos de 'simples', enquanto amarelos mais claros, valores que se aproximam mais das classificações de alta complexidade. Já para o vermelho, quanto mais escura for a cor, mais complexo o texto é pela interpretação de cada métrica.

Quadro 57 - Valores das métricas de semântica e léxico dos textos do livro *Intermediate*

	SEMÂNTICA E LÉXICO				
	PS	IA	FPC	CLP	TT
I1	4,4	313,36	587,74	72,91	679
I2	3,39	306,69	582,4	89,62	772
I3	4	297,96	585,47	48,01	620
I4	4,59	293,69	584,81	43,25	669
I5	4,45	362,94	587,56	73,89	664
I6	4,06	351,53	572,2	53,19	736
I7	3,86	339,53	576,02	38,21	700
I8	3,74	288,43	585,41	25,14	746
I9	4,56	336,38	575,49	58,71	818
I10	4,22	273,7	580,23	94,29	792
I11	4,1	273,84	588,54	67	754
I12	4,28	330,45	580,88	34,09	680
I13	3,43	298,2	583,37	50	693
I14	4,29	325,36	591,24	49,2	833
I15	4,18	365,94	586,97	57,53	756
I16	4,14	389,09	581,99	6,3	777
I17	3,63	343,17	567,87	48,01	734

I18	3,47	279,33	582,56	22,66	689
I19	3,96	263,32	593,69	40,13	745
I20	3,5	303,36	576,59	89,8	749
I21	3,78	272,08	589,4	67,36	565
MÉDIA	4,00	314,68	582,88	53,78	722

Fonte: Elaboração própria

Legenda: PS: Polissemia | IA: Idade de aquisição | FPC: Familiaridade com palavras de conteúdo | CLP: Concretude Lexical | TT: *Type-Token*

Embora as cores possam auxiliar a perceber as indicações de complexidade, é sempre necessário a interpretação humana para confirmar se as indicações são coerentes e quais componentes textuais tiveram impacto no valor de cada índice. Sendo assim, é importante destacarmos algumas observações sobre as métricas, iniciando pela PS.

8.5.3.1.1 Análise prévia da PS dos textos I

A métrica PS classifica 57,14% dos textos do livro *Intermediate* como simples, 33,33%, como tendo complexidade média e 14,29%. Isso significa que apenas três entre os vinte e um textos podem indicar alta complexidade referente à polissemia. Conforme já havíamos advertido a respeito da PS nas análises dos textos dos livros *Elementary* e *Pre-intermediate*, embora os valores desses três textos indiquem alto potencial de complexidade, nossa observação indica que, na verdade, as palavras polissêmicas apresentam baixo potencial de complexidade quando as observamos dentro do contexto. Visto que os textos em questão sejam utilizados com alunos de um nível, presumidamente, intermediário, seus conhecimentos linguísticos possivelmente auxiliam na correta “ativação” dos significados das palavras.

Para exemplificar, o trecho abaixo (com palavras polissêmicas em negrito) é parte do texto destacado como mais complexo pela PS (texto I2, PS: 3,39):

*But she learned the **hard** way that you can't really socialize if you don't want to spend money. The evening after her year-long challenge ended, she bought all her friends a **drink** in the pub. But she has not gone **back** to her past spending **habits**.*

Para prosseguirmos com a análise de léxico e semântica, passamos agora para comentários sobre os valores de IA.

8.5.3.1.2 Análise prévia da IA dos textos I

O percentual de classificação de complexidade por IA dos textos do livro *Intermediate* indica uma divisão que seria [quase] ideal para o nível, visto que 61,90% dos textos são considerados como tendo complexidade média e 19,05% como alta e baixa. Se os textos apresentassem progressão constante de complexidade dessa métrica, ou seja, se nos textos iniciais a métrica indicasse maior presença de palavras simples e a complexidade fosse progredindo com o passar das unidades, teríamos o cenário ideal para manutenção linguística e que fosse coerente para o nível proposto. Entretanto, ao destacarmos os textos considerados complexos pela IA (e pela análise de progressão que será realizada ao final desse capítulo), percebemos que isso não ocorre.

Tendo em vista que os valores da métrica IA refletem no tipo de palavras presentes no texto (sendo que valores mais baixos indicam alta presença de palavras que são aprendidas mais cedo na infância, ou seja, que são, teoricamente, mais simples), textos destacados como complexos alertam para a possível frequência de unidades lexicais potencialmente difíceis. Ainda que esses textos sejam utilizados com alunos de um nível intermediário, palavras que são consideradas complexas até para falantes nativos também serão, conseqüentemente, para estudantes de ILA.

O texto I16 foi considerado o mais complexo pela IA (389,09). Abaixo, um fragmento desse texto para ilustrar a complexidade do seu vocabulário:

*“It depends what you want to do. Some **degrees** are worth it, like medicine or **dentistry**.*

*There are so many better alternatives out there, in my opinion. I got a place at uni to do **accountancy**, but I chose to do an **apprenticeship**. All my friends are now at uni and in **debt**. I'm 20 and I'm **earning** money and learning on the job.*

*I'm a software **engineer** at a global tech **company**.*

*A **degree** is preferred, but not essential. The **recruitment** team always say if they have two people, and one is **self-taught** and has experience, and the other has just finished uni with no experience, they'll choose the first.”*

Em suma, em alguns textos do livro *Intermediate*, podemos encontrar vocabulários que podem ser altamente complexos para um estudante brasileiro de ILA e que poderiam ser alterados sem prejudicar a manutenção linguística. Ainda que

tenhamos destacado somente os textos sinalizados como tendo complexidade alta pela IA, enfatizamos que é sempre importante um olhar cuidadoso sobre o vocabulário apresentado em textos para alunos de ILA visto que, conforme já explicado, pesquisadores de complexidade textual (como DuBay, 2004 e Silva, 2018) alertam que o léxico é, na maior parte das vezes, a principal característica de alta complexidade textual.

A seguir, uma métrica que pode complementar os indicativos da IA para léxico e semântica, a FPC.

8.5.3.1.3 Análise prévia da FPC dos textos I

Nenhum texto foi classificado como tendo complexidade alta pela FPC. 76,19% dos textos são considerados simples, enquanto 23,81%, médios em complexidade.

De forma similar às observações que apresentamos para a métrica IA, algumas palavras de conteúdo (ou seja, substantivos, verbos, adjetivos e advérbios) podem ser potencialmente complexas para nosso perfil de leitor. Algumas dessas unidades lexicais, conforme podemos perceber nos destaques dos trechos apresentados, são complexas pela extensão, pela ortografia, pelo significado, ou por todas essas características simultaneamente.

Outra observação importante que pode ser considerada pelas orientações da FPC é que em casos de *clusters*⁶⁹ (como na parte ‘*British Architects Gold Medal*’ do texto I20), embora as palavras não sejam consideradas complexas de forma isolada, sua posição antes do substantivo e sua funcionalidade como adjetivos podem ser uma característica com alto potencial de complexidade, especialmente porque não utilizamos essa forma de adjetivação em língua portuguesa. Ainda que tenhamos apresentado essa observação com base no exemplo I20, ela é válida para todo e qualquer texto, especialmente quando os *clusters* são longos.

Em síntese, ainda que a métrica FPC não tenha destacado nenhum texto como tendo complexidade potencialmente alta, análises “manuais” dos textos nos permitem encontrar palavras de conteúdo que podem ser complexas. Para casos específicos, é válido adotar estratégias de simplificação como apresentaremos nos capítulos 9

⁶⁹ Sequências de palavras que ocupam a mesma função sintática (como quando posicionadas em frente a um substantivo, exercem a função de adjetivo, por exemplo).

(‘Sugestões de Estratégias de Acessibilidade para os textos do *Corpus* de Estudo’) e 10 (‘Orientações para Acessibilidade Textual Linguodidática’) desta tese.

Na próxima subseção, discutiremos os indicativos da CLP.

8.5.3.1.4 *Análise prévia da CLP dos textos I*

Para a CLP, 14,29% dos textos têm complexidade baixa, 28,57%, média e 57,14%, alta. Contudo, mesmo diante desse alto percentual de textos complexos, a leitura dos textos nos permite que concluir que, de forma similar aos textos dos demais livros, as palavras não-concretas não evocam significados desconhecidos para os alunos, especialmente por serem adjetivos e advérbios com equivalências em português.

Evidentemente, como já destacamos antes, essas classes de palavras (adjetivos e advérbios) são apontadas pela literatura como indicativos de complexidade. Entretanto, se tomarmos o texto I6 como exemplo (considerado o mais complexo pela CLP, com valor da métrica de 6,3), podemos perceber que as palavras “abstratas” não estariam em um degrau alto de abstração (conforme a escada de abstração de Hayakawa, 1978, apresentada anteriormente.). Abaixo, um trecho desse texto com palavras abstratas destacadas que auxilia a perceber que seu potencial de complexidade não é tão alto para estudantes de um nível intermediário:

*“A degree is **preferred**, but not **essential**. The recruitment team **always** say if they have two people, and one is self-taught and has **experience**, and the other has **just finished uni with no experience**, they'll choose the first.”*

Para concluir a análise prévia do léxico e da semântica dos textos do livro *Intermediate*, prosseguimos para comentários sobre os valores de TT.

8.5.3.1.5 *Análise prévia da TT dos textos I*

A TT classifica a complexidade dos textos do livro *Intermediate* da seguinte forma: 28,57% tendo complexidade baixa, 38,10%, média e 33,33% alta.

Se os textos apresentassem progressão constante (ou seja, textos mais acessíveis em termos de variedade lexical – característica estimada pela TT - nas

unidades iniciais, com acessibilidade média no meio e com vocabulário com maiores variações no final), teríamos o cenário ideal para o nível intermediário, visto que a progressão auxiliaria na manutenção linguística.

Ao analisarmos holisticamente os textos destacados pela TT como tendo alta complexidade, percebemos que a maior parte deles apresenta substantivos próprios (fator que interfere na variedade lexical, uma vez que aumenta a quantidade de *types*). O trecho abaixo foi retirado do texto I14 (considerado mais complexo pela TT, com valor de 833 para essa métrica), no qual podemos ver não só a presença de muitos nomes próprios em apenas um parágrafo, mas também a variedade lexical:

“In the 21st century, charisma is the quality that people in all fields of life, from business to politics, would most like to have. But can you learn it? The man I have just met thinks so. His name is Danish Sheikh, and he is a charisma coach. He has worked with Microsoft, Yahoo, and the BBC. and he thinks he can turn anyone into George Clooney. He charges £150 an hour, and plenty of people are paying. And for two days, I'm going to be his student.”

Evidentemente, somente variedade lexical não é o suficiente para estimar complexidade relacionada ao léxico, por esse motivo, é essencial levar em consideração as métricas previamente mencionadas e realizar, preferencialmente, um cruzamento de dados para destacar quais textos têm muitas palavras diferentes e, simultaneamente, unidades com altas densidades semânticas.

Na seção a seguir, iniciaremos a análise prévia relacionada às métricas de sintaxe dos textos do livro *Intermediate*.

8.5.3.2 Análise prévia da complexidade sintática dos textos I

No quadro a seguir, (Quadro 58) organizamos as métricas relativas à sintaxe dos textos do livro *Intermediate*. Os valores estão classificados pelo sistema de cores indicativas de complexidade:

Quadro 58 - Valores das métricas de sintaxe dos textos do livro *Intermediate*

	SINTAXE				
	NT	SS	SIMS	CR	IF
I1	59,48	39,36	59	14,01	72,14
I2	97,13	39,74	100	66,28	75,9
I3	87,9	58,32	108	43,25	81,94
I4	80,51	71,23	92	38,21	84,1
I5	58,71	20,61	69	50	59,75
I6	56,36	74,22	96	9,18	70
I7	49,6	63,68	83	17,62	65,32
I8	90,49	50	87	26,11	76,32
I9	41,68	61,03	92	12,71	62,77
I10	86,86	36,32	73	48,4	67,82
I11	97,88	42,07	99	46,81	80,93
I12	59,48	61,79	84	21,77	75,55
I13	78,81	34,09	78	62,55	79,16
I14	93,82	61,79	71	13,35	81,07
I15	23,89	30,85	55	40,52	64,75
I16	82,38	64,43	112	12,71	82,15
I17	53,98	25,14	65	42,07	65,69
I18	95,45	56,36	89	40,9	90
I19	99,27	62,55	134	68,08	91,8
I20	48,01	31,92	97	8,53	65,73
I21	94,84	68,08	115	52,39	82,3
MÉDIA	73,17	50,17	88	35,02	75,01

Fonte: Elaboração própria

Legenda: NT: Narratividade SS: Simplicidade sintática SIMS: Similaridade Sintática CR: Coesão Referencial IF: Índice Flesch
--

Conforme realizado nas análises anteriores (*Elementary* e *Pre-intermediate*), iniciamos nossas interpretações (ainda que prévias) das métricas pelos valores de NT (na próxima subseção).

8.5.3.2.1 Análise prévia da NT dos textos I

Pela métrica NT, 52,38% dos textos são classificados como simples, 47,62% como tendo complexidade média e apenas 4,76%, complexidade alta. Visto que os textos desse livro são voltados para alunos de um nível intermediário, maiores distâncias às estruturas de narrativas não tendem a ser, geralmente, um fator de alto potencial de complexidade, especialmente para alunos universitários.

Conforme explicamos na descrição da métrica, a NT indica a semelhança ou diferença de um texto a uma estrutura de narrativa. Em termos de acessibilidade, materiais com valores altos de NT tendem a ser considerados mais simples, uma vez que as narrativas costumam ser os primeiros tipos de texto com os quais temos contato desde a primeira infância e, também, pelo seu desencadeamento ser lógico e, usualmente, claro.

Dito isso, é preciso lembrar que os textos do nosso *corpus* de estudo são utilizados com alunos universitários que, em consequência, têm contato com outros gêneros textuais e estão (em geral) habituados a diferenciar e compreender traços de textos acadêmicos, descritivos etc. Para níveis iniciais, textos com valores altos de NT são de grande auxílio (pelas razões já citadas), mas acreditamos que essa métrica tende a “perder a força” com o avanço dos níveis linguísticos dos alunos.

Obviamente, os resultados dos nossos estudos também podem ser válidos para adaptação de textos para casos de turmas com alunos com altas limitações linguísticas ou neuroatípicos. Para esses alunos, a análise de complexidade e a adoção de estratégias de simplificação devem ser práticas comuns para seus professores de ILA.

Em suma, ainda que a NT indique alta complexidade para apenas um texto (I15), é válido lembrar que os textos destacados pela métrica como tendo complexidade média (marcados em amarelo na tabela de métricas) também podem ser potencialmente complexos para estudantes com limitações linguísticas e/ou neuroatípicos. Entretanto, nossa experiência em sala de aula com esses textos nos alerta que os traços de complexidade são geralmente relacionados ao léxico (conforme já discutido) ou, em termos de sintaxe, às diferentes configurações, aos múltiplos deslocamentos de elementos e às pontuações – características que analisaremos nas próximas subseções pelas demais métricas de sintaxe.

Visto que a NT é uma métrica com baixo potencial de estimativa de complexidade para o nível intermediário, precisamos observar os indicativos de outros índices. A seguir, analisaremos, previamente, os valores de SS.

8.5.3.2.2 Análise prévia da SS dos textos I

Para a métrica SS, nenhum texto do livro *Intermediate* pode ser classificado como tendo baixa complexidade. 57,14% são destacados como tendo complexidade média, enquanto 42,86%, alta.

Conforme mencionamos na nossa descrição do livro *Intermediate* (seção 8.4.3 desta tese), os textos desse nível tendem a apresentar configurações sintáticas muito distintas, com uma grande quantidade de deslocamentos, de casos de coordenação e subordinação e de elementos de coesão referencial.

Como podemos perceber pelo trecho abaixo (do texto I5, classificado como o mais complexo pela SS, com valor da métrica de 20,61), a sintaxe é, de fato, uma característica que pode ser potencialmente complexa:

“Researchers from the University of Maryland separated baby rats from their mothers and discovered that baby rats with more Foxp2, a kind of protein necessary for the development of speech and language, made more noise and called to their mothers more often. Moving on to humans, scientists then discovered that women have 30% more Foxp2 than men.”

No trecho acima, é possível constatar algumas características que influenciam na complexidade sintática do texto, tais como dois verbos para um mesmo sujeito, presença de um aposto longo (que separa um sujeito do seu predicado), deslocamento de elementos de suas posições canônicas e uso de *Present Participle* (característica que tende a ser complexa para estudantes brasileiros quando relacionam o sufixo ‘-ing’ somente como o ‘-ndo’ do português.).

Em resumo, as características de sintaxe medidas pela SS dos textos do livro *Intermediate* pode ser compreendida como um fator de alto potencial de complexidade mesmo para o nível intermediário. Para expandir ainda mais a análise, prosseguimos para observações sobre os valores de SIMS.

8.5.3.2.3 Análise prévia da SIMS dos textos I

A métrica SIMS aponta que 19,05% dos textos podem ser considerados simples, 47,62%, com complexidade média e 33,33%, alta.

Ainda que tenhamos percentuais classificados como “simples”, é preciso ter coerência na análise. A métrica SIMS, como já mencionado, observa a similaridade sintática entre as frases do texto. Quanto mais similares forem, menos conhecimentos sintáticos são exigidos e, em consequência, mais acessíveis tendem a ser. Entretanto, ao levarmos em consideração os dados relacionados à métrica SS (descrita anteriormente), é preciso interpretar que embora muitos textos tenham frases que sejam parecidas sintaticamente, elas ainda são complexas. Como analogia, se analisarmos manuais de aviação pela SIMS, teremos indicações de muita semelhança sintática, visto que as frases nesses textos tendem a se configurar de formas parecidas; entretanto, elas ainda são sintaticamente complexas.

O texto I2, por exemplo, foi classificado pela SIMS como tendo complexidade média (com valor de SIMS de 100), mas pela SS, alta (SS: 39,74). Ao destacarmos o primeiro parágrafo deste texto, podemos perceber que a sintaxe pode ser compreendida como potencialmente complexa:

“Michelle McGagh, a 34-year-old journalist, had a mortgage of £230,000 which she wanted to reduce. Last November, she looked at her finances and she discovered that every year she spent a lot of money unnecessarily - for example, £1,570 in the pub, £1,110 in restaurants, and £400 on coffees. So she made a radical plan - to stop spending money for a year. She continued to pay her broadband, gas, and electricity bills, and allowed herself £30 a week to buy food and household essentials. But she couldn't spend money on anything else - no clothes, no buses or flights, no meals out, no drinks, and no cosmetics.”

Como é possível perceber pelo trecho destacado acima, existem diversas características que potencializam a complexidade sintática dos textos. Ainda que no I2 existam outras frases que sejam altamente similares (ou até mesmo, completamente) entre si, as configurações são – de forma geral – complexas.

Além da estruturação sintática, outro elemento importante de observar é a coesão referencial, métrica discutida a seguir.

8.5.3.2.4 Análise prévia da CR dos textos I

Para a CR, a classificação de complexidade é a seguinte: 23,81%, simples; 33,33%, médios e; 42,86%, complexos.

Ao lermos os textos destacados como complexos pela métrica, compreendemos que existem muitas referências dêiticas ou fora do texto (como por meio de generalizações). Além disso, alguns pronomes retomam palavras mencionadas em outros parágrafos (e não no mesmo onde o pronome é apresentado).

No trecho do texto I20 apresentado abaixo (texto considerado como o mais complexo pela CR, com valor de 8,53 para essa métrica), podemos ver que a palavra *'people'* é repetida diversas vezes e é utilizada para generalização (apontando para um referencial que está fora do texto):

“Many people think that 2016 was one of the worst years in history for deaths of influential people. They were people who entertained us and educated us, people we loved and people we hated, but all of them were icons...”

Já no texto I6 (segundo mais complexo, CR: 9,18), além da generalização feita pela palavra *'people'*, o pronome *'him'* se refere a um sujeito mencionado em outro parágrafo – característica que impacta nos valores de CR – conforme podemos observar pelo trecho abaixo:

“Good for him. It's time people protested against this kind of behaviour. If people can't live without their phones for two hours, they shouldn't go to concerts.”

Em suma, embora os textos sejam destinados a alunos de um nível intermediário, acreditamos que existem instâncias nas quais a coesão pode ser compreendida como um fator de alto potencial de complexidade.

Por fim, prosseguimos para os resultados de IF.

8.5.3.2.5 Análise prévia da IF dos textos I

Para o IF, apenas 4,76% dos textos são considerados complexos – o que significa, apenas um texto. 66,67% são classificados como simples, enquanto 28,57% como médios em complexidade.

Visto que esse índice observa extensão de palavras e de frases, seus valores indicam que mais da metade dos textos tem unidades lexicais e/ou sentenças com tamanhos adequados.

O texto destacado como sendo mais complexo por essa métrica (IF: 59,75) foi o I5 e, ao realizarmos sua leitura, é possível notar que existem muitas palavras extensas, não só substantivos comuns, mas também próprios, conforme mostra o trecho desse texto abaixo:

“Psychologists at the University of Herefordshire gave 50 male and 50 female students eight minutes to perform three tasks at the same time: maths problems, finding restaurants on a map, and drawing a plan for finding a lost key in a field.”

O IF, então, sinaliza que em extensão, o potencial de complexidade é baixo. Entretanto, é preciso sempre lembrar que embora essa seja a única métrica dentre as utilizadas neste trabalho que foi desenvolvida especificamente para estimar complexidade, sua utilização sem outras métricas não é aconselhada (especialmente ao percebermos que textos como o I1 e o I12, por exemplo, são considerados simples por IF, mas como tendo complexidades médias ou altas pelas demais métricas.

Na próxima seção, para concluir as análises objetivadas neste estudo, apresentamos gráficos para interpretação de progressão de complexidade e nossas observações sobre seus indicativos.

8.6 ANÁLISE DE PROGRESSÃO DE COMPLEXIDADE

Com a finalidade de verificar se a hipótese de insumo de Krashen (1987) está sendo atendida – hipótese que define que para manutenção e evolução linguística de aprendizes de línguas adicionais devem ser apresentados *inputs* que sejam um pouco

além das suas competências linguísticas atuais (c+1)⁷⁰ – inserimos os valores das métricas textuais de cada um dos livros em gráficos linhas. Sendo assim, com o objetivo de averiguar se os textos do nosso *corpus* de estudo apresentam progressão de conteúdos lexicais e gramaticais de forma gradual e constante, os gráficos-linha de progressão nos auxiliarão a visualizar se a gradação adequada de complexidade está ocorrendo.

Nesta seção, temos a finalidade de observar dois fenômenos: 1. Progressão de complexidade intranívelar, ou seja, se existe progressão de complexidade dentro de cada nível (se há um avanço constante, gradual e “natural” de complexidade entre todos os textos de um mesmo livro) e; 2. Progressão de complexidade internívelar, ou seja, se ocorrem progressões de um nível para o outro (se os textos demonstram maiores complexidades para cada uma das métricas não só entre unidades do mesmo livro, mas se a gradação também ocorre quando níveis são avançados).

Evidentemente, o fato de os textos apresentarem progressões de complexidade perfeitamente constantes é um cenário altamente improvável e irreal - uma vez que seria praticamente impossível controlar todos os níveis de cada um dos textos para que as métricas apresentassem sempre um crescimento de progressão de forma constante. Visto que textos são escritos com uma língua viva e dinâmica, seria utópico imaginar comportamentos métricos que demonstrassem que todos os níveis textuais crescem em complexidade com o avanço das unidades. Entretanto, ainda que saibamos que a progressão não será impecavelmente constante, esperamos que os gráficos demonstrem a maior coerência possível no avanço de complexidade para potencializar a manutenção linguística. Sendo assim, em cenários ideais perceberíamos: I. Indicativos de que a complexidade dos textos avança de forma gradual e constante (mesmo com poucos intervalos em que alguma métrica indique redução da complexidade em vez de aumento) e; II. Valores das métricas que classifiquem os textos do livro *Intermediate* como sendo mais complexos que os textos do *Pre-intermediate*, e que os materiais do nível pré-intermediário indiquem maior complexidade que os do *Elementary*.

⁷⁰ (c+1) é uma fórmula na qual c representa a quantidade de conteúdos e competências do aluno em determinado momento, enquanto o +1 indica o input que é necessário para progressão linguística.

Tendo a finalidade de tornar a progressão mais visual e acessível⁷¹, conduziremos a interpretação de progressão por meio de quadros nos quais apresentaremos e comentaremos as indicações de progressão (ou a falta dela) com o suporte de um gráfico-linha para cada métrica. Os gráficos são organizados da seguinte forma:

1. Cada nível será representado por uma linha (laranja: *Elementary*, amarelo: *Pre-intermediate* e verde: *Intermediate*);
2. O eixo horizontal de cada gráfico representa a progressão ordenada dos textos de cada livro. Uma vez que cada nível tem quantidades minimamente distintas de textos, utilizaremos a nomenclatura 'Texto 1', 'Texto 2' etc. no lugar das etiquetas específicas de cada livro (ainda assim, a etiqueta Texto 1 posiciona no gráfico o texto E1 na linha laranja, o texto P1 na linha amarela e o texto I1 na linha verde - e assim progressivamente);
3. O eixo vertical dos gráficos representa a progressão de complexidade. É importante lembrar que o comportamento de cada métrica é distinto (característica que alterará a exibição das linhas). Para facilitar a interpretação dos gráficos, considere o Quadro 59 abaixo que resume a relação métrica-complexidade e explica o movimento de cada linha para cada índice:

Quadro 59 - Resumo dos valores e indicativos das linhas na relação métrica-complexidade

MÉTRICAS	RESUMO DOS VALORES	INDICATIVOS DAS LINHAS
NT, SS, SIMS, CR, IF, PS, FPC e CLP	Valores e acessibilidade em relação diretamente proporcional. Ou seja, valores mais baixos representam indícios de maior complexidade, enquanto valores mais altos, indícios de menor complexidade.	Pontos mais altos no eixo vertical indicam maior acessibilidade e linhas em ascensão retratam crescimento de acessibilidade entre textos.
IA e TT	Valores e acessibilidade em relação inversa, ou seja,	Pontos mais baixos no eixo vertical indicam maior

⁷¹ Nas versões anteriores desta tese, estávamos conduzindo a análise de progressão pela apresentação do gráfico seguida por uma interpretação em forma de texto. Entretanto, observamos que muitas frases acabavam se tornando repetitivas, apenas alterando os dados apresentados. Se tivéssemos optado por continuar a análise daquela forma, nosso texto se tornaria ainda mais extenso e o potencial de acessibilidade seria prejudicado. Logo, para não contradizermos as orientações que nós mesmos defendemos por meio desta tese, optamos por apresentar a análise de progressão de forma mais simplificada por meio de quadros.

	valores mais altos representam indícios de maior complexidade, enquanto valores mais baixos, indícios de menor complexidade.	acessibilidade e linhas em declínio representam redução de complexidade entre textos.
--	--	---

Fonte: Elaboração própria

Destacamos que, ainda que possamos perceber aumento ou diminuição de complexidade entre um texto e outro (por exemplo, a diminuição do valor de PS entre o texto E1 e E2, indicando maior progressão de complexidade), para nós, é interessante considerarmos intervalos com três ou mais textos para verificar progressão intranívelar, visto que diferenças de valores perceptíveis apenas entre dois textos não configura um intervalo por meio do qual seja possível observar constância de progressão. Dessa forma, destacaremos intervalos com três ou mais textos que demonstrem aumento ou diminuição de complexidade para cada uma das métricas.

No quadro abaixo (Quadro 60), ilustramos e explicamos os quadros que serão utilizados para a análise de progressão:

Quadro 60 - Explicação dos quadros de análise de progressão

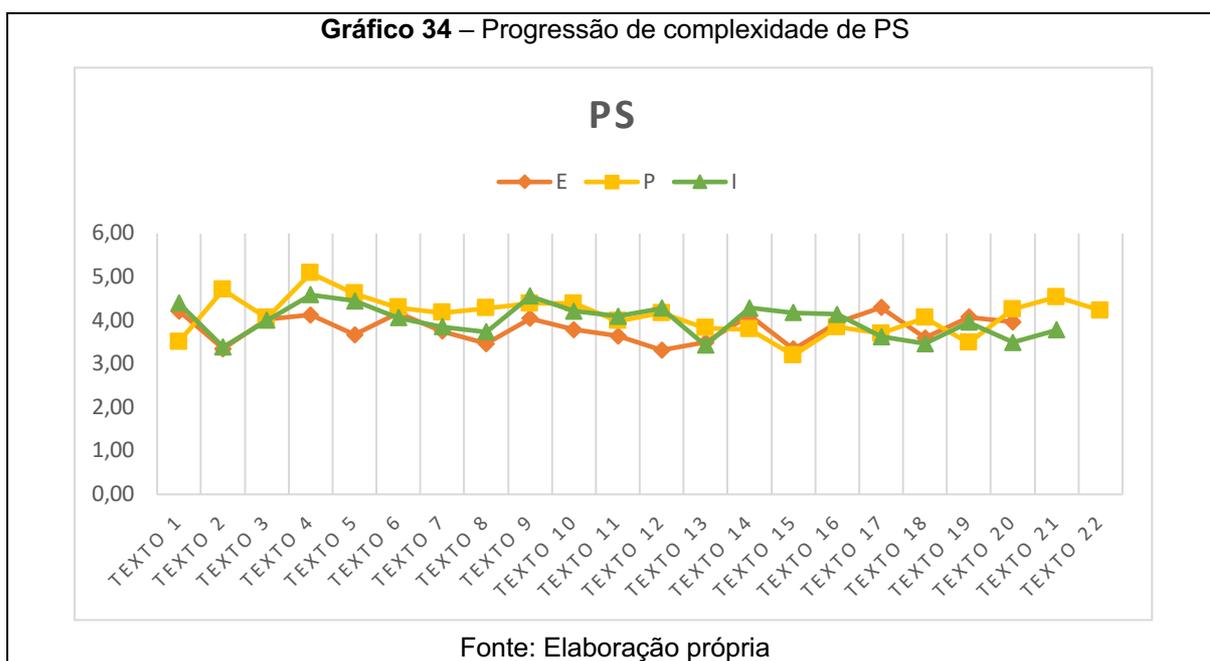
Neste campo, apresentaremos os gráficos em linha de cada uma das métricas.	
PROGRESSÃO INTRANIVELAR	
Progressão de complexidade intranívelar constante?	Para este campo, responderemos 'sim' caso as linhas indiquem que a progressão de complexidade foi constante dentro de cada nível e 'não' caso a constância esperada para manutenção linguística não tenha ocorrido.
Indicações de aumento ou redução de complexidade para cada nível:	Neste campo, apresentaremos a quantidade de vezes que a complexidade aumentou e a quantidade de vezes que diminuiu para cada um dos níveis.
Maior(es) intervalos de progressão de complexidade para cada nível:	Aqui, destacaremos intervalos com três ou mais textos para cada um dos níveis que tenha demonstrado progressão constante de complexidade.
Observações:	Neste campo, apresentaremos observações que julgarmos interessantes para a análise de progressão de complexidade intranívelar.

PROGRESSÃO INTERNIVELAR	
Progressão de complexidade internivelar constante?	Para este campo, responderemos 'sim' caso as linhas indiquem que a progressão de complexidade foi percebida entre cada nível (ou seja, se as métricas indicam complexidades maiores para os textos do livro <i>Intermediate</i> em relação ao <i>Pre-intermediate</i> etc.) e 'não' caso a constância esperada para manutenção linguística não tenha ocorrido.
Nível com o texto mais complexo pela métrica:	Aqui, apresentaremos qual nível tem o texto indicado como o mais complexo para a métrica.
Nível com o texto mais acessível pela métrica:	Nesta célula, apresentaremos qual nível tem o texto indicado como o mais acessível para a métrica.
Observações:	Neste campo, apresentaremos observações que julgarmos interessantes para a análise de progressão de complexidade internivelar.

Fonte: Elaboração própria

Seguindo a ordem das análises realizadas até o momento, iniciaremos pela observação das métricas relacionadas ao léxico e à semântica. A seguir, apresentamos a análise de progressão pela métrica PS conforme o quadro abaixo (Quadro 61):

Quadro 61 – Análise de progressão de complexidade da métrica PS



PROGRESSÃO INTRANIVELAR			
Progressão de complexidade intranivelar constante?	Não		
Indicações de aumento ou redução de complexidade para cada nível:		Aumento	Redução
	<i>Elementary:</i>	10	10
	<i>Pre-intermediate:</i>	11	11
	<i>Intermediate:</i>	13	8
Maior(es) intervalos de progressão de complexidade para cada nível:	<i>Elementary:</i>	E6 – E8 e E9 – E12	
	<i>Pre-intermediate:</i>	P4 – P7 e P10 – P15	
	<i>Intermediate:</i>	I4 - I8 e I14 - I18	
Observações:	Livro <i>Elementary</i> : valores que indicam maior acessibilidade pela PS estão na metade final do livro; fenômeno contrário do que seria ideal.		
PROGRESSÃO INTERNIVELAR			
Progressão de complexidade internivelar constante?	Não		
Nível com o texto mais complexo pela métrica (ponto mais baixo no gráfico):	<i>Pre-intermediate</i>		
Nível com o texto mais acessível pela métrica (ponto mais alto no gráfico):	<i>Pre-intermediate</i>		
Observações:	<ul style="list-style-type: none"> ▪ O livro <i>Pre-intermediate</i> apresenta o valor mais alto e o mais baixo de PS, indicando que dentro desse nível temos o texto com maior e o texto com menor complexidade entre todos os textos de todos os níveis; ▪ No intervalo entre os textos E3 – E12 do livro <i>Elementary</i>, percebemos maior constância entre os picos mais baixos de PS. Se compararmos o intervalo similar dos outros níveis, vemos que os valores dessas métricas são mais altos, indicando maior acessibilidade (pela interpretação desse índice) nos níveis pré-intermediário e intermediário; contrário do que seria o ideal 		

Fonte: Elaboração própria

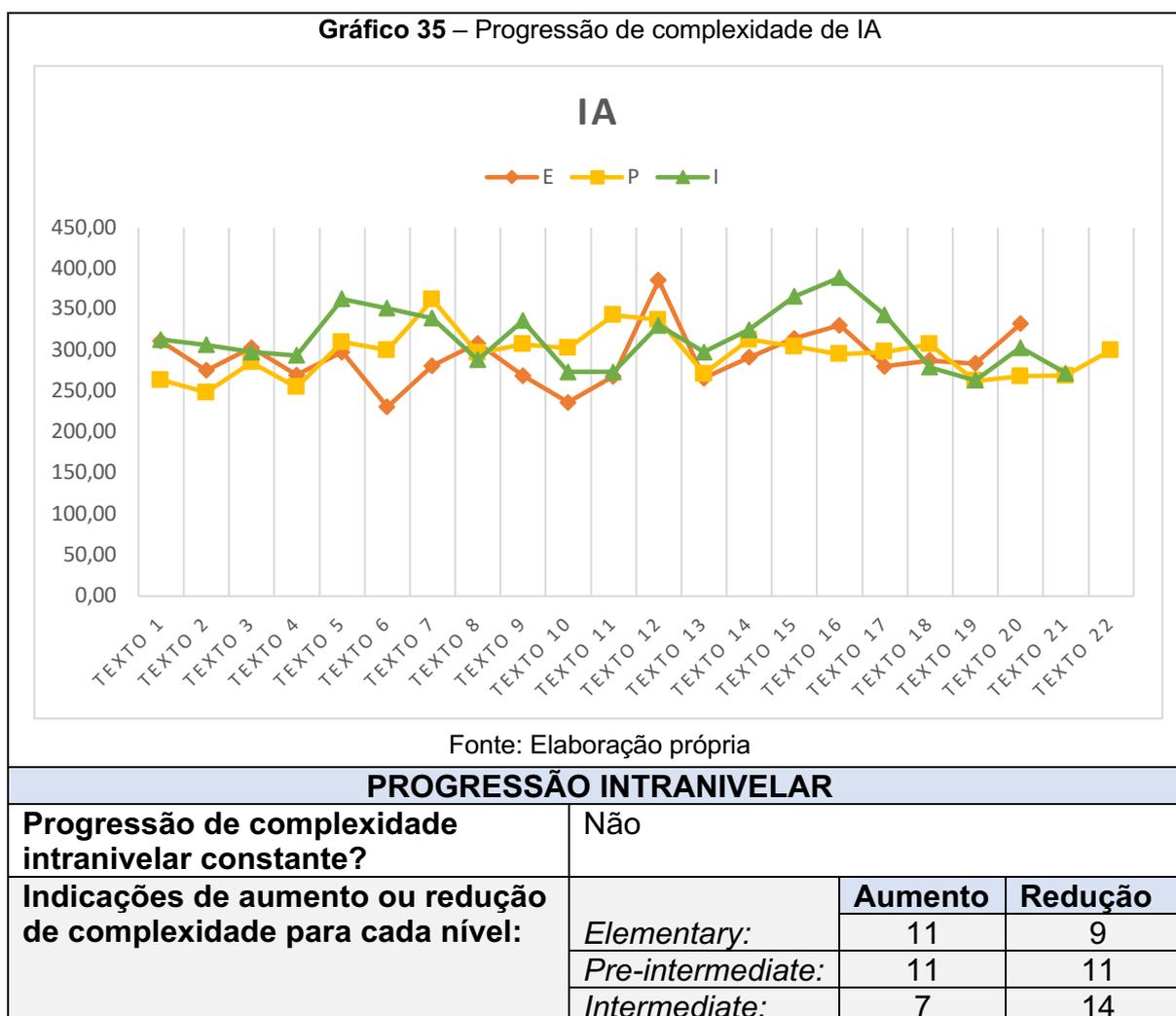
Em suma, os intervalos que demonstram progressão de complexidade estimada pela PS são curtos e inconstantes, indicando que a presença de palavras polissêmicas não é frequente e nem controlada. Além disso, o texto considerado mais complexo e o texto indicado como o mais acessível estão no livro *Pre-intermediate*.

Se buscássemos coerência em progressão de complexidade, o texto mais simples – por qualquer métrica – deveria estar no *Elementary* e o mais complexo no *Intermediate*.

Embora tenhamos concluído que não há progressão de complexidade constante por meio da frequência de palavras polissêmicas, é importante lembrar que, durante nossa pesquisa, compreendemos que os indícios de complexidade provenientes da PS são fracos. Caso tivéssemos o objetivo de progredir a complexidade, seria necessário “forçar” cenários nos quais palavras polissêmicas pudessem assumir mais de um sentido (estratégia que não beneficiaria, de forma alguma, a manutenção linguística).

A seguir, interpretaremos o gráfico de progressão de complexidade da métrica IA por meio do Quadro 62:

Quadro 62 – Análise de progressão de complexidade da métrica IA



Maior(es) intervalos de progressão de complexidade para cada nível:	<i>Elementary:</i>	E6 – E8, E10 – E12 e E13 – E16
	<i>Pre-intermediate:</i>	P19 – P22
	<i>Intermediate:</i>	I13 – I16
Observações:	<ul style="list-style-type: none"> ▪ O livro <i>Elementary</i> é o que apresenta intervalos mais curtos de progressão estimada por IA; ▪ O livro <i>Intermediate</i> é o que apresentou maior avanço de complexidade entre os textos (no intervalo I13-I16). 	
PROGRESSÃO INTERNIVELAR		
Progressão de complexidade internivelar constante?	Não	
Nível com o texto mais complexo pela métrica (ponto mais alto no gráfico):	<i>Intermediate</i>	
Nível com o texto mais acessível pela métrica (ponto mais baixo no gráfico):	<i>Elementary</i>	
Observações:	O único intervalo no qual as métricas são coerentes em questão de progressão de complexidade internivelar (ou seja, indicando que os textos do <i>Elementary</i> são mais acessíveis do que os do <i>Pre-intermediate</i> que, por sua vez, tem textos mais acessíveis que no livro <i>Intermediate</i>) foi entre os textos 5 e 6 de cada nível. Para os demais intervalos, não há coerência.	

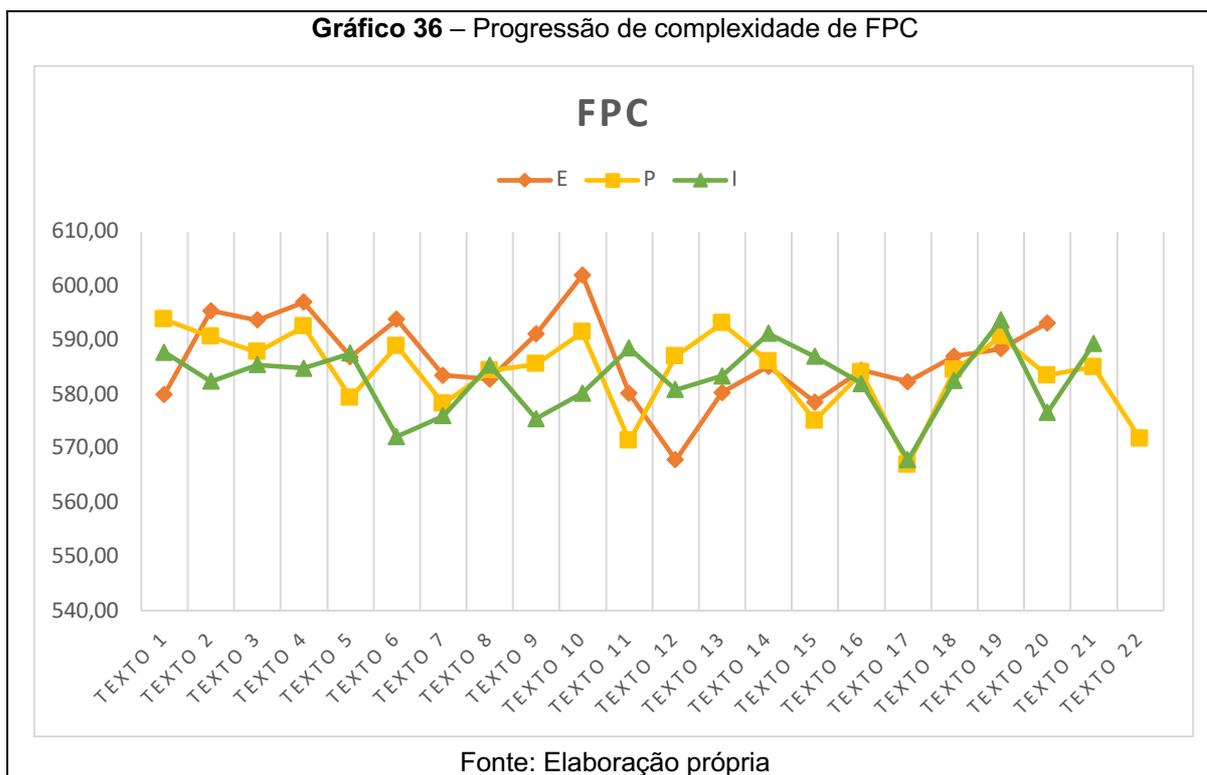
Fonte: Elaboração própria

Em suma, não é possível perceber progressão constante nem intranivelmente, nem entre os níveis. Além disso, os intervalos que relatam progresso de complexidade são muito curtos, indicativos de inconstância.

Uma característica positiva que pode ser percebida pelo gráfico de progressão de complexidade da IA é que o texto considerado mais simples é parte do livro *Elementary*, enquanto o mais complexo, do *Intermediate*. Conforme citamos anteriormente, o resultado ideal é exatamente o que ocorreu para a IA: texto mais acessível pertencente ao *Elementary* e texto mais complexo, ao *Intermediate*.

A seguir, continuamos nossa análise de progressão de complexidade por meio do gráfico-linha da métrica FPC que é apresentado e interpretado no Quadro 63 abaixo:

Quadro 63 – Análise de progressão de complexidade da métrica FPC

**PROGRESSÃO INTRANIVELAR**

Progressão de complexidade intranivelar constante?	Não		
Indicações de aumento ou redução de complexidade para cada nível:		Aumento	Redução
	<i>Elementary:</i>	8	12
	<i>Pre-intermediate:</i>	11	11
	<i>Intermediate:</i>	8	13
Maior(es) intervalos de progressão de complexidade para cada nível:	<i>Elementary:</i>	E10 - E12	
	<i>Pre-intermediate:</i>	P13 - P15 e P19 - P22	
	<i>Intermediate:</i>	I14 - I17	
Observações:	Nos intervalos E17 - E20, P17 - P19 e I17 - I19, o comportamento da métrica é inverso do que seria considerado ideal. Visto que são textos posicionados no final do livro, os valores dos índices deveriam indicar que houve aumento na complexidade, mas indicam diminuição.		

PROGRESSÃO INTERNIVELAR

Progressão de complexidade internivelar constante?	Não
Nível com o texto mais complexo pela métrica (ponto mais baixo no gráfico):	<i>Elementary</i>
Nível com o texto mais acessível pela métrica (ponto mais alto no gráfico):	<i>Pre-intermediate</i>

<p>Observações:</p>	<ul style="list-style-type: none"> ▪ Nos intervalos de textos 2 – 4 e 9 - 10 observamos que existe coerência de complexidade entre os níveis; ▪ O livro <i>Elementary</i> é o que apresenta textos menos complexos para a métrica na parte inicial do livro; ▪ Ainda que de forma inconstante, são muitos os intervalos em que o livro <i>Intermediate</i> apresenta maior complexidade ou uma diferença muito pequena para os demais níveis; ▪ O maior aumento de complexidade para cada nível foi entre os textos E9 - E12 para o <i>Elementary</i>, P13 – P15 para o <i>Pre-intermediate</i> e I14 – I17 para o <i>Intermediate</i>.
----------------------------	---

Fonte: Elaboração própria

Para a métrica FPC, também não é possível perceber coerência e constância na progressão de complexidade. Os intervalos nos quais se observa progresso de dificuldade são muito curtos.

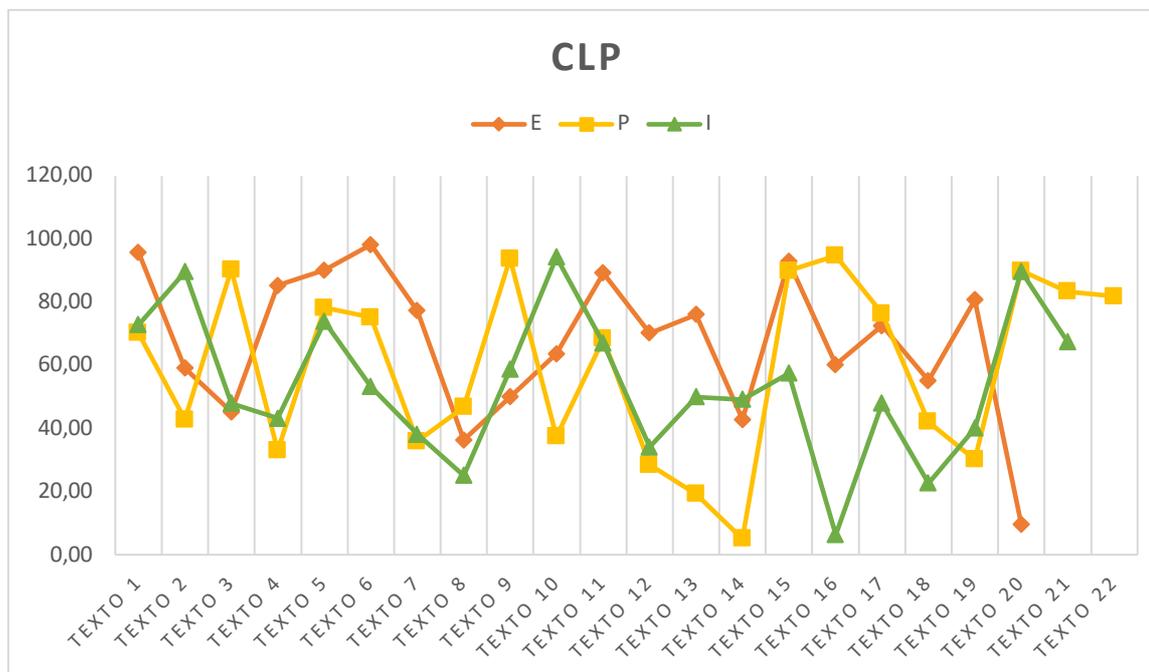
Ao contrário da métrica IA e do que seria ideal, o texto mais complexo pela FPC está no livro *Elementary*, enquanto o mais simples, no *Pre-intermediate*.

Como sabemos, a falta de constância de progressão de complexidade para a métrica IA e FPC está diretamente relacionada ao léxico e à semântica. Caso uma professora de ILA tenha por objetivo progredir a complexidade do léxico/semântica com seus alunos (caminho oposto à acessibilidade), existem estratégias que podem ser adotadas para tal finalidade. Uma das estratégias é a substituição lexical baseada em frequência em *corpus* (sobre a qual falaremos no capítulo 10, 'Orientações para Acessibilidade Textual Linguodidática') de forma oposta – substituindo unidades lexicais muito simples (com alta frequência em *corpus*) por sinônimos mais complexos (com baixa frequência em *corpus*).

Com a finalidade de continuar investigando se existe progressão de complexidade relacionada ao léxico e semântica, passamos agora para a análise da métrica CLP pelo Quadro 64 abaixo:

Quadro 64 – Análise de progressão de complexidade da métrica CLP

Gráfico 37 – Progressão de complexidade de CLP



Fonte: Elaboração própria

PROGRESSÃO INTRANIVELAR

Progressão de complexidade intranivelar constante?	Não		
Indicações de aumento ou redução de complexidade para cada nível:	<i>Elementary:</i>	Aumento	Redução
	<i>Pre-intermediate:</i>	9	11
	<i>Intermediate:</i>	13	9
Maior(es) intervalos de progressão de complexidade para cada nível:	<i>Elementary:</i>	E1 – E3, E6 – E8 e E11 – E14	
	<i>Pre-intermediate:</i>	P5 – P7, P11 – P14, P16 – P19 e P20 – P22	
	<i>Intermediate:</i>	I2 – I4, I5 – I8 e I10 – I12	
Observações:	Os três níveis apresentam grandes “saltos” e “quedas” nos valores de CLP entre um texto e outro (vide a diferença entre os textos P14 e P15) e intervalos curtos de progressão constante.		

PROGRESSÃO INTERNIVELAR

Progressão de complexidade internivelar constante?	Não
Nível com o texto mais complexo pela métrica (ponto mais baixo no gráfico):	<i>Pre-intermediate</i>

Nível com o texto mais acessível pela métrica (ponto mais alto no gráfico):	<i>Elementary</i>
Observações:	O único intervalo no qual é possível observar coerência de complexidade (<i>Intermediate</i> mais complexo que o <i>Pre-intermediate</i> e assim sucessivamente) é entre os textos 4 e 7.

Fonte: Elaboração própria

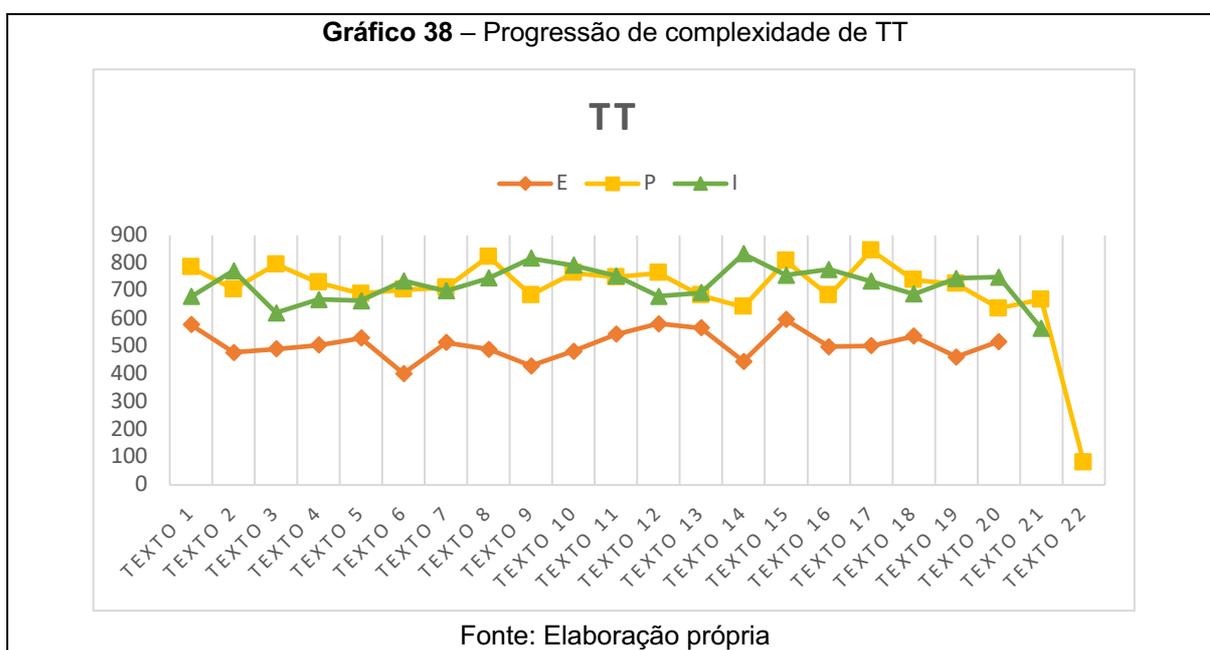
Dentre todas as métricas analisadas até o momento, a CLP é a que apresenta maior inconstância de progressão de complexidade intranívelar, especialmente por constantes “subidas” e “descidas” dos resultados da métrica e pela grande distância de valores entre um texto e outro.

Embora o texto mais acessível esteja no livro *Elementary* (resultado ideal), o mais complexo não é parte do *Intermediate*, mas do *Pre-intermediate*.

Em suma, podemos concluir que a CLP também não apresenta consistência em progressão de complexidade. Entretanto, é válido lembrar que durante nossa pesquisa, constatamos que essa métrica pode ser considerada altamente subjetiva, especialmente se levarmos em consideração o conceito de Escada de Abstração proposto por Hayakawa (1978).

Para finalizar a análise de progressão de complexidade lexical e semântica, prosseguimos para a observação da métrica TT pelo Quadro 65 abaixo:

Quadro 65 – Análise de progressão de complexidade da métrica TT



PROGRESSÃO INTRANIVELAR			
Progressão de complexidade intranivelar constante?	Não (Contudo, existem intervalos consideravelmente longos no nível <i>Elementary</i> .)		
Indicações de aumento ou redução de complexidade para cada nível:		Aumento	Redução
	<i>Elementary</i> :	11	9
	<i>Pre-intermediate</i> :	9	13
Maior(es) intervalos de progressão de complexidade para cada nível:	<i>Intermediate</i> :	10	9
	<i>Elementary</i> :	E2 – E5, E9 – E12 e E16 – E18	
	<i>Pre-intermediate</i> :	P5 – P7	
Observações:	<i>Intermediate</i> :	I3 – I5 e I7 – I9	
	<ul style="list-style-type: none"> No intervalo de textos E9 – E13 podemos ver o comportamento de TT de uma forma que seria ideal, progredindo constantemente e com uma distância de valores não muito alta entre um texto e outro; A maior queda de valores da métrica é entre o penúltimo (P21) e último (P22) texto do livro <i>Pre-intermediate</i>, comportamento que pode ser considerado peculiar. 		
PROGRESSÃO INTERNIVELAR			
Progressão de complexidade internivelar constante?	Entre <i>Elementary</i> e <i>Pre-Intermediate</i> , sim. Entre <i>Pre-Intermediate</i> e <i>Intermediate</i> , não.		
Nível com o texto mais complexo pela métrica (ponto mais alto no gráfico):	<i>Pre-intermediate</i>		
Nível com o texto mais acessível pela métrica (ponto mais baixo no gráfico):	<i>Pre-intermediate</i>		
Observações:	<ul style="list-style-type: none"> A distância entre os valores de TT do livro <i>Elementary</i> para os demais é evidente – característica que indica que esse nível apresenta textos menos complexos pela interpretação dessa métrica; Ainda que possam ser considerados como mais complexos do que os textos do <i>Elementary</i>, os livros <i>Pre-intermediate</i> e <i>Intermediate</i> apresentam constantes mudanças de posição nos valores de TT, indicando que não há constância de progressão de complexidade para essa métrica. 		

Fonte: Elaboração própria

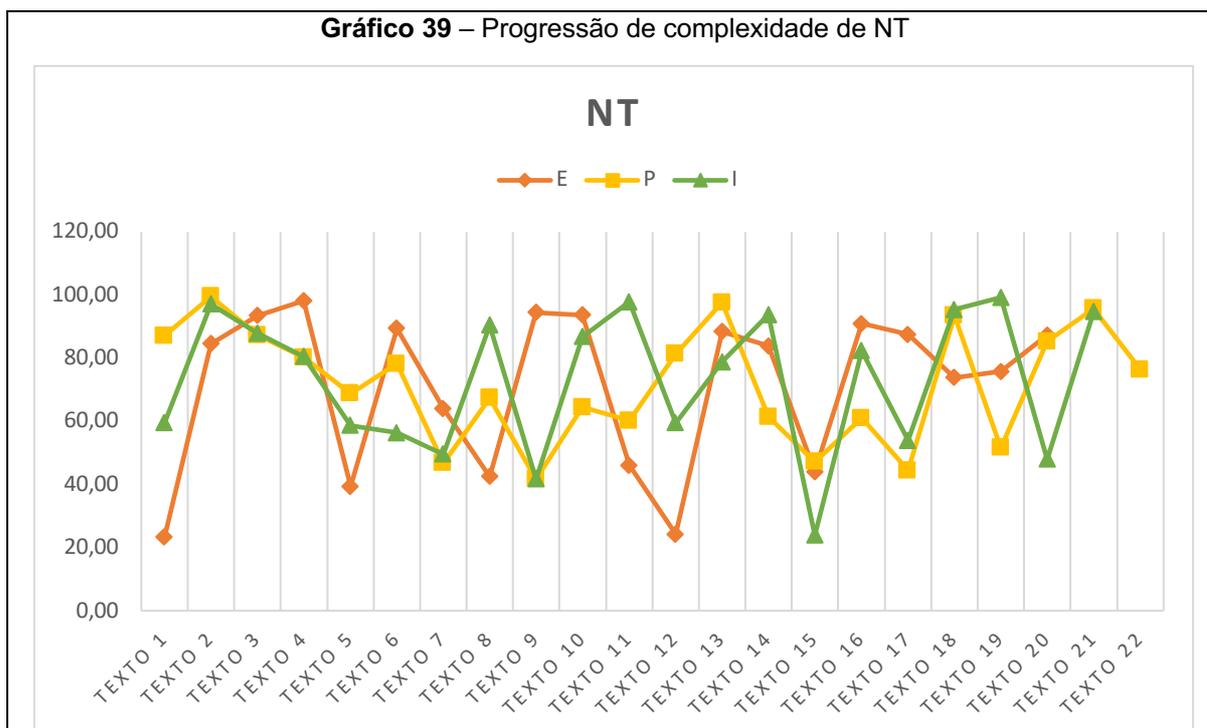
Em síntese, a métrica TT é a que apresenta um intervalo de progressão intranívelar considerado como ideal para o movimento de qualquer métrica (intervalo E9 – E13 que demonstra uma complexidade aumentada gradualmente). Além disso, essa é a única métrica de léxico que evidenciou uma distância de complexidade coerente entre o livro *Elementary* e os demais (embora isso não tenha acontecido entre o *Pre-intermediate* e o *Intermediate*).

O texto considerado mais complexo pela TT (ou seja, aquele com maior variedade lexical) e o mais acessível (com menor variedade lexical) pertencem ao livro *Pre-intermediate*. Entretanto, se eliminássemos o último texto do nível pré-intermediário (que apresentou comportamento inusitado), o material com valor indicativo de maior acessibilidade estaria no livro *Elementary*, visto que o valor de TT de todos os textos desse nível são menores que os resultados dos demais livros.

Por fim, as métricas relativas a léxico e à semântica nos demonstram que não há progressões constantes e controladas nem dentro dos mesmos níveis, nem entre os níveis. Conforme já exemplificamos, existem estratégias que permitem aumentar a acessibilidade das palavras em textos indicados como complexos e que, quando adotadas de forma inversa, podem potencializar a complexidade. Quando aplicadas com coerência, as estratégias de simplificação podem uniformizar a progressão de complexidade, caso esse seja o objetivo linguístico (especialmente para grupos de alunos que progridem linguisticamente de acordo com os níveis).

A seguir, iniciaremos a análise de progressão das métricas relacionadas à sintaxe. Iniciaremos nossa análise pelo comportamento da NT no quadro a seguir (Quadro 66):

Quadro 66 – Análise de progressão de complexidade da métrica NT



Fonte: Elaboração própria

PROGRESSÃO INTRANIVELAR

Progressão de complexidade intranivelar constante?	Não		
Indicações de aumento ou redução de complexidade para cada nível:	<i>Elementary:</i>	Aumento 10	Redução 10
	<i>Pre-intermediate:</i>	11	11
	<i>Intermediate:</i>	10	11
Maior(es) intervalos de progressão de complexidade para cada nível:	<i>Elementary:</i>	E2 – E4 e E18 – E20	
	<i>Pre-intermediate:</i>	P11 – P13 e P19 – P21	
	<i>Intermediate:</i>	I2 – I7	
Observações:	<ul style="list-style-type: none"> ▪ O livro <i>Intermediate</i> é o que apresenta o maior intervalo constante de progressão de complexidade (I2 – I7). Entretanto, após o texto I7, é o nível que apresenta maior inconstância do comportamento métrico; ▪ A distância dos valores entre um texto e outro é geralmente muito grande em todos os níveis. 		

PROGRESSÃO INTERNIVELAR

Progressão de complexidade internivelar constante?	Não
Nível com o texto mais complexo pela métrica (ponto mais baixo no gráfico):	<i>Elementary</i>

Nível com o texto mais acessível pela métrica (ponto mais alto no gráfico):	<i>Intermediate</i>
Observações:	Não há nenhuma constância dos valores de NT entre os níveis. É possível observar que os níveis se intercalam entre os pontos mais altos e mais baixos no gráfico e que não existe progressão de complexidade gradual e constante.

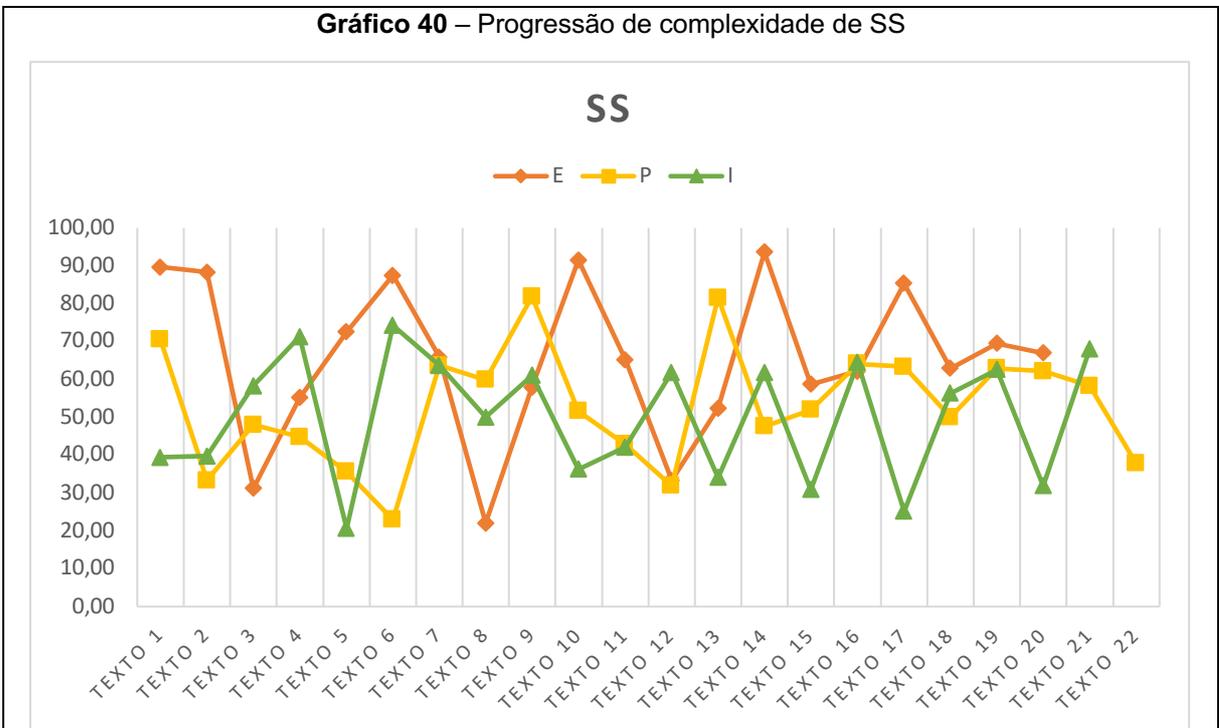
Fonte: Elaboração própria

Os valores de NT indicam alta inconstância tanto dentro de cada nível, quanto de forma internivelar. A variação nos valores é frequente e existe apenas um intervalo relativamente longo que indica progressão constante (I2 – I7).

De forma contrária ao que seria ideal, o texto mais complexo pela NT é parte do livro *Elementary*, enquanto o mais simples, do *Intermediate*.

A seguir, observaremos o comportamento da SS para verificar se há progressão de complexidade. No quadro abaixo (Quadro 67) apresentamos e discutimos pontos importantes do gráfico de progressão dessa métrica:

Quadro 67 – Análise de progressão de complexidade da métrica SS



Fonte: Elaboração própria

PROGRESSÃO INTRANIVELAR		
Progressão de complexidade intranivelar constante?	Não	
	Aumento	Redução

Indicações de aumento ou redução de complexidade para cada nível:	<i>Elementary:</i>	9	11
	<i>Pre-intermediate:</i>	13	9
	<i>Intermediate:</i>	8	13
Maior(es) intervalos de progressão de complexidade para cada nível:	<i>Elementary:</i>	E6 – E8 e E10 – E12	
	<i>Pre-intermediate:</i>	P3 - P6, P9 – P12, P16 – P18 e P19 – P22	
	<i>Intermediate:</i>	I6 – I8	
Observações:	<ul style="list-style-type: none"> A partir dos textos 10, todos os níveis apresentam grande inconstância e diferenças grandes entre os valores da métrica. 		
PROGRESSÃO INTERNIVELAR			
Progressão de complexidade internivelar constante?	Não		
Nível com o texto mais complexo pela métrica (ponto mais baixo no gráfico):	<i>Intermediate</i>		
Nível com o texto mais acessível pela métrica (ponto mais alto no gráfico):	<i>Elementary</i>		
Observações:	<ul style="list-style-type: none"> Na maioria dos intervalos, o livro <i>Intermediate</i> apresenta maior complexidade, enquanto o <i>Pre-intermediate</i> a segunda maior complexidade e o <i>Elementary</i>, a menor complexidade. Entretanto, esse comportamento é inconstante e, em diversas instâncias, os valores das três métricas são muito próximos (como no caso dos textos 7, 16 e 19). 		

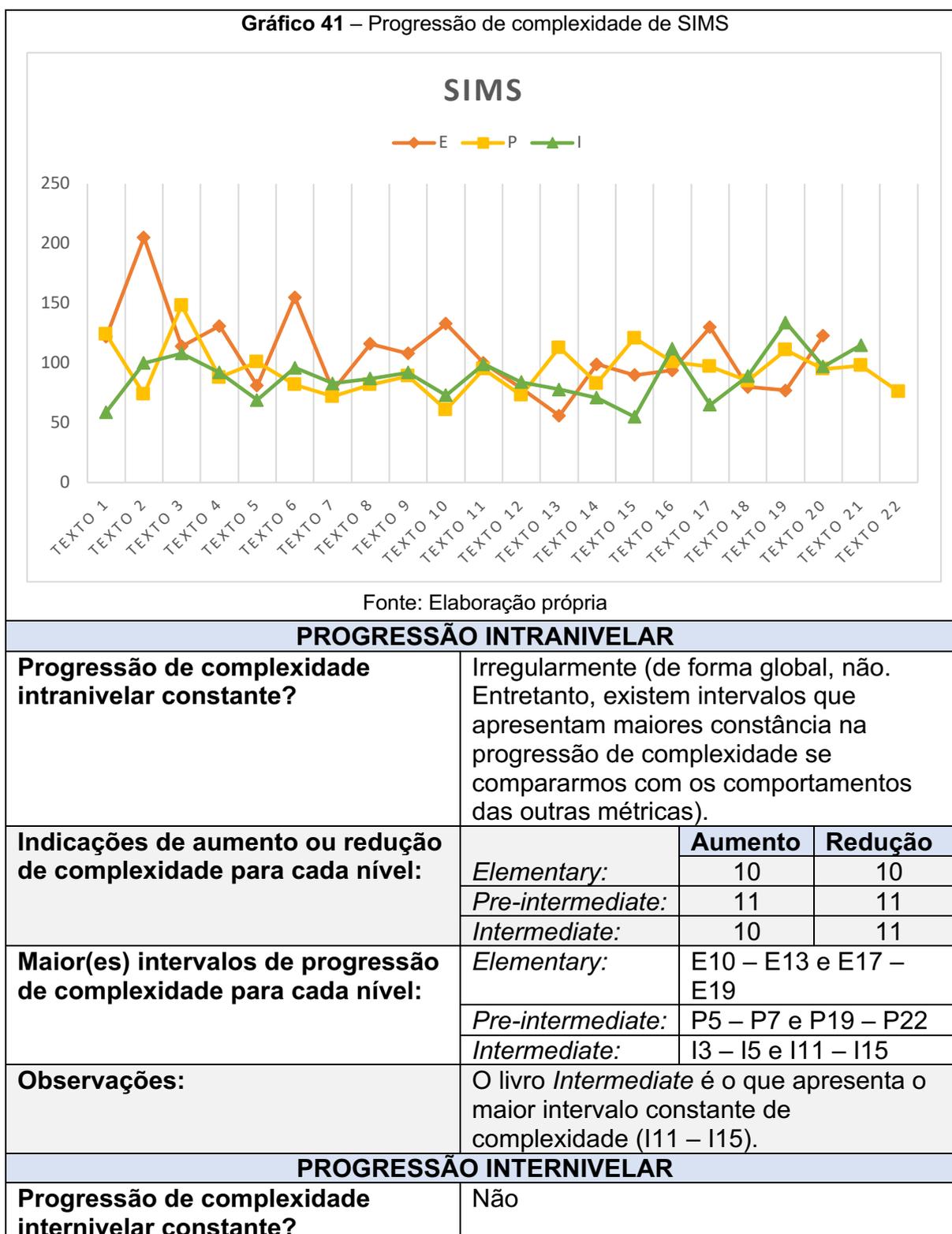
Fonte: Elaboração própria

De forma similar à NT, a SS também demonstra grande inconstância de progressão de complexidade tanto intra quanto internivelarmente. Além disso, os intervalos nos quais podemos perceber graduação de complexidade são curtos e seguidos, geralmente, por um crescimento métrico indicativo de maior acessibilidade (movimento contrário do que seria ideal).

Um ponto coerente na observação de complexidade entre os níveis é que o texto mais complexo é parte do livro *Intermediate*, enquanto o texto mais acessível, do *Elementary*.

Abaixo, continuamos a análise de progressão sintática pelo quadro abaixo (Quadro 68), que apresenta e discute a gradação (ou diminuição) de complexidade da métrica SIMS:

Quadro 68 – Análise de progressão de complexidade da métrica SIMS



Nível com o texto mais complexo pela métrica (ponto mais baixo no gráfico):	<i>Intermediate</i>
Nível com o texto mais acessível pela métrica (ponto mais alto no gráfico):	<i>Elementary</i>
Observações:	Somente nos textos 19 e 20 podemos observar complexidade menor no livro <i>Elementary</i> , média no livro <i>Pre-intermediate</i> e mais alta no <i>Intermediate</i> .

Fonte: Elaboração própria

Embora não seja possível observar coerência na progressão de complexidade pela SIMS, existem intervalos (curtos) que demonstram avanço em todos os níveis.

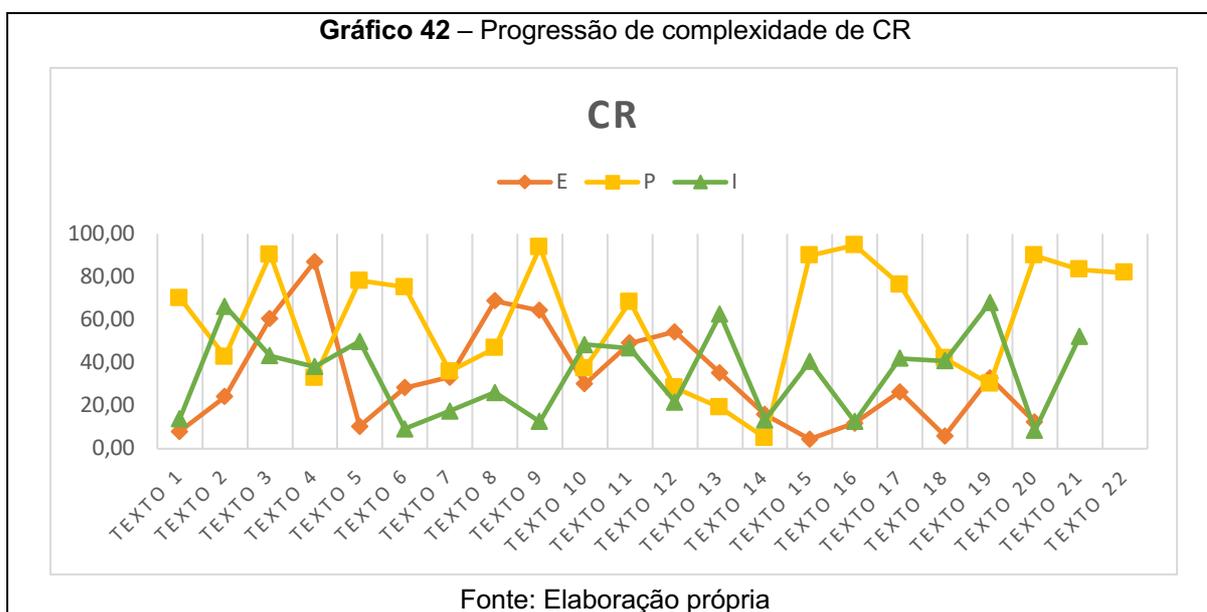
No que diz respeito às diferenças de complexidade entre níveis, percebemos que em alguns intervalos da parte inicial dos livros, os textos do *Elementary* se destacam dos demais sendo classificados como mais acessíveis pela SIMS (fator positivo). Contudo, a partir do texto 11 de todos os níveis, a progressão passa a ser irregular e os níveis intercalam na posição de maiores e menores complexidades.

Similarmente à SS, para a métrica SIMS o texto mais acessível também é parte do livro *Elementary* e o mais complexo, do *Intermediate*.

A penúltima métrica que nos auxiliará a observar se existe progressão coerente de complexidade em nível sintático é a CR, que apresentamos e discutimos no Quadro 69 a seguir:

Quadro 69 – Análise de progressão de complexidade da métrica CR

Gráfico 42 – Progressão de complexidade de CR

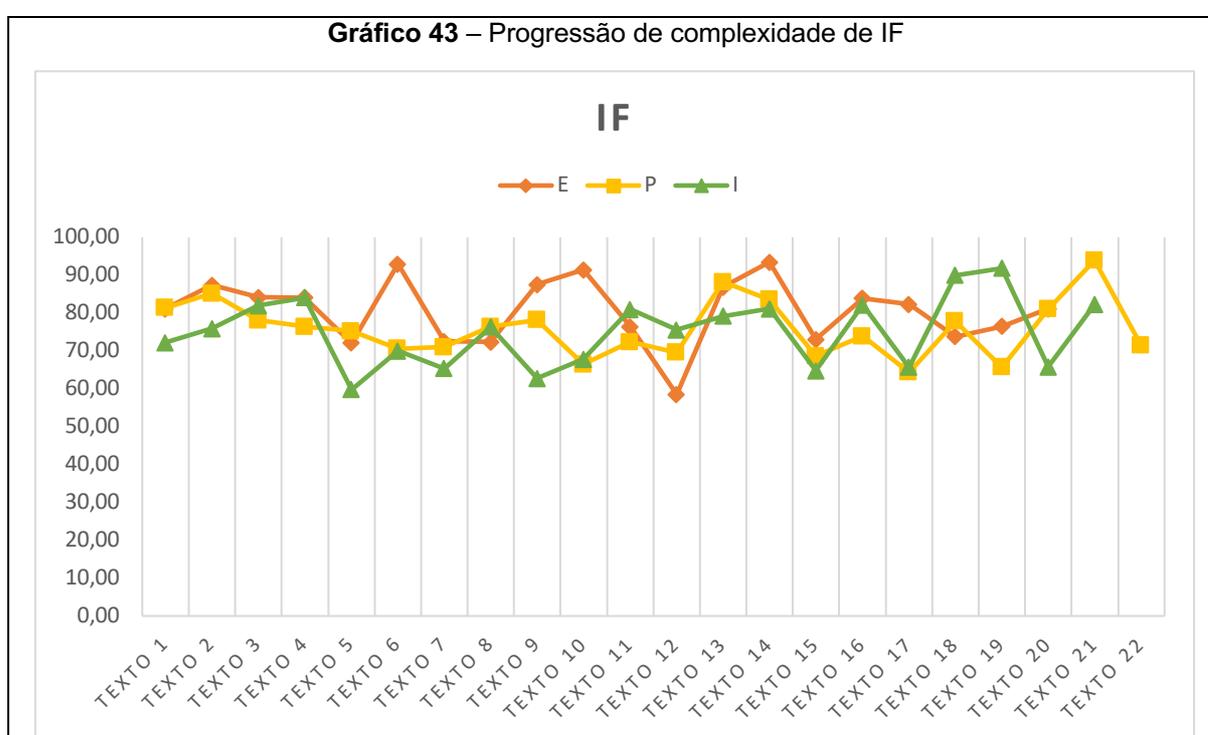


PROGRESSÃO INTRANIVELAR			
Progressão de complexidade intranivelar constante?	Não		
Indicações de aumento ou redução de complexidade para cada nível:		Aumento	Redução
	<i>Elementary:</i>	8	12
	<i>Pre-intermediate:</i>	13	9
	<i>Intermediate:</i>	10	11
Maior(es) intervalos de progressão de complexidade para cada nível:	<i>Elementary:</i>	E8 – E10 e E12 – E15	
	<i>Pre-intermediate:</i>	P5 – P7, P11 – P14, P16 – P19 e P20 – P22	
	<i>Intermediate:</i>	I2 – I4 e I10 – I12	
Observações:	<ul style="list-style-type: none"> ▪ O livro <i>Pre-intermediate</i> é o que apresenta mais intervalos de progressão, embora curtos; ▪ A distância de valores entre um texto e outro é, por vezes, muito grande (como entre o texto P14 e P15), fato que demonstra que além de não existir constância na progressão de complexidade pela CR, não há equilíbrio no aumento/diminuição de complexidade nem entre textos próximos. 		
PROGRESSÃO INTERNIVELAR			
Progressão de complexidade internivelar constante?	Irregularmente (ainda que a constância não seja regular entre os níveis, podemos observar intervalos nos quais as distâncias de valores são consideravelmente grandes, especialmente nos textos finais.).		
Nível com o texto mais complexo pela métrica (ponto mais baixo no gráfico):	<i>Elementary</i>		
Nível com o texto mais acessível pela métrica (ponto mais alto no gráfico):	<i>Pre-intermediate</i>		
Observações:	<ul style="list-style-type: none"> ▪ O livro <i>Pre-intermediate</i> tem sete textos com valores de CR superiores a todas as outras métricas dos demais livros (indicando que existem mais textos menos complexos por essa métrica nesse livro); ▪ Não existem intervalos nos quais a ordenação da progressão de complexidade segue a configuração ideal (<i>Elementary</i>, <i>Pre-Intermediate</i> e <i>Intermediate</i>). 		

O gráfico de progressão da métrica CR nos auxiliou a constatar que, de forma similar às demais métricas de sintaxe, não há progressão de complexidade de maneira constante nem dentro de um mesmo nível, nem entre os níveis. Além disso, de maneira oposta ao ideal, o texto mais complexo por essa métrica é parte do livro *Elementary*, enquanto o mais acessível, do *Pre-intermediate*.

Para finalizarmos a verificação da [falta] de progressão de complexidade a nível sintático, apresentamos no quadro a seguir (Quadro 70) o gráfico e comentários a respeito da métrica IF:

Quadro 70 – Análise de progressão de complexidade da métrica IF



Fonte: Elaboração própria

PROGRESSÃO INTRANIVELAR			
Progressão de complexidade intranivelar constante?	Não		
Indicações de aumento ou redução de complexidade para cada nível:		Aumento	Redução
	<i>Elementary:</i>	10	10
	<i>Pre-intermediate:</i>	11	11
Maior(es) intervalos de progressão de complexidade para cada nível:	<i>Elementary:</i>	E2 – E5, E10 – E12 e E16 – E18	
	<i>Pre-intermediate:</i>	P2 – P6 e P13 – P15	
	<i>Intermediate:</i>	Não se aplica	
Observações:	No nível <i>Intermediate</i> , não há nenhum intervalo com três ou mais textos que apresentem constância na progressão de		

	complexidade (somente na progressão da acessibilidade, comportamento contrário ao esperado).
PROGRESSÃO INTERNIVELAR	
Progressão de complexidade internivelar constante?	Não
Nível com o texto mais complexo pela métrica (ponto mais baixo no gráfico):	<i>Elementary</i>
Nível com o texto mais acessível pela métrica (ponto mais alto no gráfico):	<i>Elementary</i>
Observações:	<ul style="list-style-type: none"> ▪ Em alguns pontos específicos (textos 2, 6, 9 e 14) os valores de IF se comportam da forma ideal, indicando textos mais acessíveis no livro <i>Elementary</i> e menos acessíveis no <i>Intermediate</i>; ▪ Ainda que não exista constância na progressão de complexidade indicada por IF, é importante lembrar que os valores dessa métrica (previamente analisados) destacam poucos textos como indicativos de complexidade que sejam incoerentes para os níveis para os quais são propostos.

Fonte: Elaboração própria

O IF também não demonstrou constância de progressão de complexidade nem intraniverlamente, nem internivelarmente – embora tenha sido adequado em quatro posições, conforme mencionado nas observações do quadro acima.

Tanto o texto considerado mais complexo quanto o mais simples estão no livro *Elementary* – fato que indica que, em termos de extensão, talvez devêssemos considerar adotar estratégias de simplificação para potencializar a acessibilidade do texto marcado como mais complexo pela IF. Entre as estratégias, é possível adotar a substituição de palavras por outras unidades mais curtas e/ou a quebra de orações em duas ou mais (conforme iremos explicar no capítulo 10, ‘Orientações para Acessibilidade Textual Linguodidática’).

Por fim, a análise das métricas de sintaxe dispostas em gráficos-linha concluiu que não existe coerência na progressão de complexidade sintática nem dentro de cada um dos níveis, nem entre níveis diferentes. Novamente, assim como orientamos

para o léxico e semântica, existem estratégias que podem ser adotadas para adequar a acessibilidade de textos destacados como complexos por alguma métrica e que estejam posicionados ou em pontos iniciais de cada livro, ou em níveis mais elementares, com a finalidade de aprimorar a progressão de complexidade (que deveria ser relativamente constante e gradual). Da mesma forma, a adoção inversa das estratégias pode potencializar a complexidade dos textos que possam ser considerados muito simples e que não estejam oferecendo insumos para a manutenção linguística (que devem ser adotadas somente para casos particulares e de forma muito cautelosa).

Diante dos resultados das nossas análises (tanto de léxico/semântica, quanto de sintaxe), existem sugestões específicas para alguns textos de cada um dos livros que podem ser seguidas com o objetivo de aumentar a acessibilidade dos textos. Durante a análise das métricas realizada neste capítulo, destacamos alguns componentes e características de textos que foram sinalizados pelas métricas como potencialmente complexos e validados por nossa interpretação linguística como tal. Sendo assim, para alunos com limitações linguísticas e/ou não neurotípicos, algumas estratégias de simplificação podem potencializar a acessibilidade linguística, reduzir o filtro afetivo e, conseqüentemente, tornar a manutenção linguística mais fácil e eficiente. No próximo capítulo, apresentaremos algumas sugestões pontuais.

9 SUGESTÕES DE ESTRATÉGIAS DE ACESSIBILIDADE PARA OS TEXTOS DO *CORPUS* DE ESTUDO

Neste capítulo, apresentamos algumas sugestões específicas de estratégias de acessibilidade que podem ser adotadas nos textos do nosso *corpus* que foram sinalizados por alguma métrica como tendo potencial de complexidade inadequada para o nível a qual se destinam e que, diante da sinalização das métricas, foram confirmados pela nossa análise individualizada como tendo traços potencialmente complexos. Essas sugestões de estratégias são provenientes de diversos estudos sobre a área (conforme citamos no capítulo 2 ‘Complexidade, Acessibilidade e Simplificação Textual’) e podem auxiliar a tornar os textos mais acessíveis para alunos com limitações linguísticas e/ou estudantes neuroatípicos. Embora não tenhamos a pretensão de apresentar estratégias para todos os textos (somente para aqueles cujas indicações de alta complexidade das métricas foram confirmados por nós), as sugestões podem ser “replicadas” para os outros textos – caso um professor ache necessário para algum aluno ou grupo. Além disso, apresentaremos outras estratégias de simplificação no capítulo 10 desta tese (‘Orientações para a acessibilidade textual linguodidática’).

É importante lembrar que embora algumas estratégias possam ser “simplórias” e que algum docente possa discordar de uma sugestão ou de outra, a finalidade dessa apresentação não é ditar de forma autoritária o que deve ou não ser aplicado. Conforme o nome orienta, os procedimentos que apresentaremos nesta seção são sugestões que podem, ou não, ser adotadas de acordo com as necessidades dos alunos ou das turmas. Ainda enfatizamos que, conforme explicamos anteriormente, a adoção de estratégias de simplificação deve ser conduzida de acordo com o perfil do leitor ou da turma para quem iremos adaptar um texto e que, por existirem diversos perfis de leitores/turmas, também existem diversos possíveis conjuntos de estratégias de acessibilidade que podem ser aplicadas. Logo, destacamos mais uma vez, as sugestões apresentadas aqui não esgotam todas as possibilidades de outras estratégias que podem ser úteis para promover acessibilidade textual.

É válido lembrar, também, que professores com mais experiência geralmente têm mais facilidade para identificar características linguísticas que possam ser complexas para os seus alunos, mas essa habilidade geralmente se desenvolve com o tempo. Logo, ainda que algumas das sugestões possam parecer “muito simples” ou

até previsíveis, elas podem ter grande validade para professores que estejam no início das suas carreiras.

Além disso, gostaríamos de destacar que as sugestões são pontuais e que, caso uma professora deseje tornar seu texto mais acessível, nem todas as estratégias precisam ser adotadas simultaneamente. É essencial que professores conheçam as limitações e necessidades dos seus alunos e façam escolhas de estratégias de simplificação textual que sejam coerentes com os objetivos propostos e com o perfil dos alunos.

Também lembramos que é preciso tomar um cuidado muito grande na adoção de estratégias para que elas não eliminem os conteúdos que são propostos pelas unidades. Sendo assim, o bom senso do professor é primordial.

Para fins de organização e espaço, apresentamos quadros (similar ao Quadro 71 de exemplo abaixo)⁷² dos textos escolhidos destacando trechos⁷³ que tenham elementos potencialmente complexos e sugestões de reescrita. Além da sugestão, apresentaremos uma etiqueta que descreve quais estratégias foram adotadas para cada trecho. Essas etiquetas poderão ser consultadas no quadro de estratégias que apresentaremos posteriormente (Quadro 72).

Quadro 71 - Exemplo de apresentação de estratégia de reescrita

ORIGINAL	SUGESTÃO DE REESCRITA	ESTRATÉGIA(S)
Mary was joyful.	Mary was happy.	L1
If you are tired, go to bed.	Go to bed if you are tired.	S1

Fonte: Elaboração própria

Nos exemplos apresentados no quadro acima, demonstramos duas estratégias de reescrita simplificada. Para a primeira frase, substituímos uma palavra potencialmente complexa (*'joyful'*) por outra mais simples (*'happy'*) e, ao lado da sugestão de reescrita, adicionamos o código L1 que, no quadro de referências de estratégia, é descrito como 'substituição de léxico por sinônimo mais simples'. Já no segundo exemplo, nossa sugestão foi a 'alteração de ordem das frases' (estratégia S1).

⁷² Por questões legais, não podemos apresentar todos os textos do *corpus* de estudo, mesmo que editados. Portanto, apresentaremos na íntegra (mas com estratégias de reescrita aplicadas) somente os textos que "exigirem" muitas estratégias de reescrita cuja apresentação por meio dos quadros tornaria a leitura pouco acessível.

⁷³ Na apresentação dos trechos nos quadros, não marcaremos as palavras em língua inglesa em itálico.

Para um caso particular (texto E15 que “exigiu” uma grande quantidade de estratégias e muitas alterações), apresentaremos um quadro um pouco diferente, dispondo o texto reescrito na primeira linha e, nas linhas posteriores, as estratégias utilizadas.

Acreditamos que essa forma de apresentação facilitará a leitura e a consulta de estratégias para textos pontuais. Para promover ainda mais a acessibilidade, faremos aqui um recurso similar ao que apresentamos anteriormente para o resumo das métricas: deixaremos o quadro de estratégias propositalmente separado em uma página para ele que possa ser destacado para consultas (em versões impressas da tese) ou para que leitores do trabalho em versão digital possam salvar/imprimir somente essa página ou utilizar o recurso de busca (*control* + F no Windows, *command* + F no MacOS), pesquisar por ‘Quadros de Estratégias’ e vir diretamente para a página correta.

Na próxima página, apresentamos o Quadro de Estratégias (Quadro 72). Logo após o quadro, passaremos para a seção de sugestões de estratégias de acessibilidade para os textos do livro *Elementary*.

Quadro 72 – Quadro de estratégias

ESTRATÉGIAS DE LÉXICO E SEMÂNTICA	
L1	substituição de léxico por sinônimo mais simples ⁷⁴
L2	substituição de pronome por referencial
L3	remoção de palavras desnecessárias
L4	substituição de palavra complexa por frase explicativa
L5	adição de aposto para palavra que não pode ser substituída ou alterada por outra mais simples
L6	substituição de pronome demonstrativo por adjetivo pessoal
L7	substituição de palavra por outra cognata verdadeira ao português (ou similar)
L8	eliminação de palavra complexa não-essencial
L9	redução de advérbios
L10	redução de adjetivos
L11	substituição de palavra por sintagma mais simples
L12	substituição de um elemento dêitico por um equivalente não-dêitico
L13	remoção de contração ou abreviação de palavra
ESTRATÉGIAS DE SINTAXE	
S1	alteração de ordem das frases
S2	divisão de orações
S3	alteração de pontuação
S4	substituição da voz passiva por voz ativa (adicionando um sujeito)
S5	reposicionamento do objeto para depois da partícula do <i>phrasal verb</i>
ESTRATÉGIAS GERAIS	
G1	eliminação de informação repetida ou não-essencial
G2	separação de título de obras, nomes de locais ou palavras em destaque por meio de aspas únicas
G3	eliminação de conjunção
G4	preenchimento de zeugmas ou elipses
G5	reescrita de frase complexa
G6	reordenação de informações
G7	eliminação de informação não-necessária
G8	inserção de tópicos frasais
G9	substituição de exemplos por outros mais comuns para a realidade dos estudantes
G10	substituição de palavras por símbolos
G11	divisão de parágrafos com muitas informações sequenciais em lista

Fonte: Elaboração própria

-- Fim da página destacada para consulta --

⁷⁴ Achamos importante destacar, mais uma vez, que sabemos que a apresentação de palavras novas contribui para que os alunos desenvolvam vocabulário. Entretanto, visto que o léxico tende a ser a maior característica de complexidade, é essencial que palavras complexas sejam apresentadas nos momentos corretos. Sendo assim, essa estratégia é sugerida quando as palavras forem consideradas potencialmente difíceis para o nível onde são apresentadas (fato que não elimina sua apresentação em um momento posterior).

9.1 SUGESTÕES DE ESTRATÉGIAS DE ACESSIBILIDADE PARA OS TEXTOS DO LIVRO *ELEMENTARY*

Nesta seção, apresentamos os quadros de sugestões de estratégia para os textos do livro *Elementary* seguindo os padrões apresentados no início do capítulo. É importante lembrar que não apresentaremos estratégias para todos os textos, somente para aqueles indicados pela análise com um ou mais traços que podem ser potencialmente complexos para os alunos. Contudo, ainda que sugestões não sejam apresentadas para todos os textos, as que serão apresentadas aqui podem ser utilizadas como parâmetros para os demais materiais. Além disso, no próximo capítulo desta tese (capítulo 10, 'Orientações para Acessibilidade Textual Linguodidática') explicaremos como utilizar as métricas textuais para estimar complexidade e quais são as estratégias que podem ser utilizadas para aumentar a acessibilidade não só dos textos do nosso *corpus* de estudo, mas outros utilizados para ensino de ILA.

Considerando que os textos estão presentes no livro voltado para o nível elementar, apresentaremos grandes quantidades de sugestões (tomando por princípio de que os alunos estão nos seus primeiros passos com o ILA e, conseqüentemente, necessitam de maior acessibilidade). Evidentemente, nem todas as sugestões precisam ser seguidas, mas achamos importante apresentá-las para esse nível para que sirvam de parâmetro para outros textos, caso uma professora julgue necessário.

A partir deste ponto, os quadros serão apresentados em sequência, iniciando pelo E1.

SUGESTÕES PARA O TEXTO E1 (*Elementary* – página 21)

Quadro 73 – Sugestões de estratégias para o texto E1

ORIGINAL	SUGESTÃO DE REESCRITA	ESTRATÉGIA(S)
The Statue of Liberty is an American icon, but there are always long queues for tickets.	The Statue of Liberty is an American icon, but there are always long lines for tickets.	L1
Buy them on the internet before you go.	Buy the tickets on the internet before you go.	L2
The best thing is to walk - but good shoes are very important!	The best thing is to walk, but good shoes are very important!	L3
But if you are tired, take a bus or the subway!	Take a bus or the subway if you are tired.	S1

If it's hot, go for a swim in the Atlantic.	Go for a swim in the Atlantic if it's hot.	S1
Go to Coney Island beach by subway (an hour) and have a delicious New York hot dog.	Go to Coney Island beach by subway. The trip takes an hour. Have a delicious New York hot dog. Ou Go to Coney Island beach by subway. The trip takes an hour. The hot dog is delicious in New York. Eat one!	S2
Good Enough to Eat on 83rd Street is a great traditional American restaurant serving breakfast, lunch, and dinner.	There is a nice restaurant called 'Good Enough to Eat'. It is on 83rd Street. It is a great traditional American restaurant. It has salads, sandwiches, steak and more.	S1

Fonte: Elaboração própria

SUGESTÕES PARA O TEXTO E2

(*Elementary* – página 23)

Quadro 74 – Sugestões de estratégias para o texto E2

ORIGINAL	SUGESTÃO DE REESCRITA	ESTRATÉGIA(S)
Khalal, 39, is from Iraq.	Khalal is 39 years old. He is from Iraq.	S2, G4 ⁷⁵
People in the street often call me 'love', 'sweetheart', and 'darling'.	People in the street call me 'love' and 'sweet'.	L1, L3
In my country, if I want to go to the bank I take my brother and he does everything for me.	In my country, if I want to go to the bank I take my brother. He does everything for me there.	S2, G4
Everybody says that British food is awful, but I love it, especially in pubs.	Everybody says that British food is bad, but I love the food, especially in pubs.	L1
I also love the beautiful scenery, for example in Cornwall and in Scotland.	I also love the beautiful places, for example in Cornwall and in Scotland.	L1

Fonte: Elaboração própria

⁷⁵ Essas estratégias também devem ser aplicadas nos subtítulos que apresentam os demais personagens e que seguem a mesma estrutura do primeiro.

SUGESTÕES PARA O TEXTO E3

(Elementary – página 32)

Quadro 75 – Sugestões de estratégias para o texto E3

ORIGINAL	SUGESTÃO DE REESCRITA	ESTRATÉGIA(S)
Marjan Jahangiri, originally from Iran, is one of the only women professors of cardiac surgery in Europe.	Marjan Jahangir is originally from Iran. She is one of the only women professors of heart surgery in Europe.	L1, S2
I get up between 6.00 and 6.30 a.m., I get to work at 7.00, and my meetings usually start at 7.30.	I get up between 6.00 and 6.30. I get to work at 7.00. My meetings usually start at 7.30.	S2
After that, I don't have a break.	After my meetings, I don't have a break.	L2
I often do two operations a day, and I also have lectures and more meetings.	I often do two operations a day, and I also have classes and more meetings.	L1
Between 9.30 and 11.30 p.m., I do research and I watch the news on TV.	I do research and I watch the news on TV between 9.30 and 11.30 p.m.	S1
One or two nights a week I'm on call, so I probably need to do operations during the night.	I need to go to work if the hospital calls me (one or two nights per week). I need to do operations during the night if they call.	L4, S1
I often work at weekends, too. But that's OK - I think I have a fantastic life because I love my work.	I often work at weekends, too, but that's OK. I think I have a fantastic life because I love my work.	L10
I play the piano for an hour every day, late at night. I think it helps me with my operations - it's technical in the same way.	I play the piano for an hour every day, late at night. I think the piano helps me with my operations because it's technical in the same way.	L2, L10
I also go to the hairdresser twice a week.	I go to the hair salon two times a week.	L1
They turn the music off for me and I use the time to read all my academic papers.	They turn off the music for me, and I use the time to read all my university texts.	L1, S5

Fonte: Elaboração própria

SUGESTÕES PARA O TEXTO E5

(Elementary – páginas 42 e 43)

Quadro 76 – Sugestões de estratégias para o texto E5

ORIGINAL	SUGESTÃO DE REESCRITA	ESTRATÉGIA(S)
One of Europe's biggest marathons takes place in spring every year, usually in April.	One of the biggest marathons in Europe happens in the spring every year. It is usually in April.	L1, S2

About 40,000 people, including top runners and celebrities, run 26.2 miles through the city.	About 40,000 people go. Many top runners and celebrities run 26.2 miles around the city.	L1, S2
You can watch anywhere along the route, but the Tower of London and the finish, near Buckingham Palace, are my favourite places.	You can watch the race in any place along the way. My favourite places are The Tower of London and the finish line (near Buckingham Palace).	L1, S2
In April in London doesn't always mean good weather - be prepared for sun, showers, and wind!	London doesn't always have a good weather in April. Be prepared for sun, rain, and wind!	L1, S2
Go to the Open Air Theatre in Regent's Park. Open from May to September, this is a great place to see Shakespeare and other classics.	Go to the Open Air Theatre in Regent's Park. The theater is open from May to September. It is a great place to see Shakespeare and other classics.	S2, G4
If it rains - and it can – you can buy plastic raincoats at the theatre.	It can rain. You can buy plastic raincoats at the theatre if it rains.	S1, S2
In bad weather they sometimes cancel a performance, but you can use your tickets on a different day.	They sometimes cancel a performance if the weather is bad. Don't worry, you can use your tickets on a different day.	S1, G3
You can buy a picnic or take your own, and you can buy dinner, too (quite expensive!).	You can buy a picnic or take your own food. You can buy dinner, too (but it is expensive!),	S2, G4
Go to the Treetop Walkway at Kew Gardens. This is one of the highlights of Kew, London's botanical gardens.	Go to the Treetop Walkway at Kew Gardens. The walkway is one of the best parts of London's botanical gardens (Kew).	L1, L2, S2
It's open all year round, and it's beautiful in autumn when the trees are changing colour.	The botanical garden is open all year. It's beautiful in the autumn when the trees are changing colour.	L2, S2
The Walkway is 18 metres above the ground, and you can see all the gardens and a lot of famous buildings, including Wembley Stadium.	The Walkway is 18 metres above the ground. There, you can see all the gardens and a lot of famous buildings, including Wembley Stadium.	S2
You can also see birds and insects very close up, and find out all about the trees around you.	You can see birds and insects very near you. You can learn about the trees around you.	L1, S2
But if the weather is very windy, the Walkway sometimes closes.	The Walkway sometimes closes if the weather is very windy.	S1
My favourite, which not all tourists know about, is the Victoria and Albert Museum, or V&A.	My favourite museum is the Victoria and Albert Museum (V&A). Not all tourists know about this place.	S2, G4
It has special activities for children, and a great café open daily until 17.15. Entrance is free, but some exhibitions charge a fee.	The museum has special activities for children, and a great café. The coffee shop opens every day until 17.15. Entrance is free, but you need to pay to see some exhibitions.	L4, S1

SUGESTÕES PARA O TEXTO E6

(Elementary – páginas 46 e 47)

Quadro 77 – Sugestões de estratégias para o texto E6

ORIGINAL	SUGESTÃO DE REESCRITA	ESTRATÉGIA(S)
They live in a small house in the desert, near the mountains, a long way from a town or village.	The brothers live in a small house in the desert. The house is near the mountains and distant from a town or village.	L1, S2
They're very poor. They have no money and they have no animals.	Hassan and Walid are very poor. They have no money, and they have no animals.	L2, S2
Then Hassan works in the fields - but Walid just sits and looks at the hot, empty desert and the hot, empty sky.	Then Hassan works in the fields, but Walid just sits and looks at the hot, empty desert and sky.	S3, G1
She looks after them, but she's old and tired.	Their mother takes care of them, but she's old and tired.	L1, L2
'If you want some money, you have your silver ring - you can sell that,' answers Walid.	'You can sell your silver ring if you want some money,' answers Walid.	L2, S1
At night, it's very cold.	It's very cold at night.	S1
He gives him food and drink, and Hassan sleeps in a comfortable bed.	The prince gives Hassan food and drink. Hassan sleeps in a comfortable bed.	L2, S2
He wants to thank the prince and give him something, but he only has his empty water bottle.	Hassan wants to thank the prince and give him something. His empty water bottle is the only thing he has.	L2, S2
He gives it to him.	Hassan gives the bottle to the prince.	L2
He is very happy, and in return he gives Hassan a box.	The prince is very happy. He gives to Hassan a box.	L2, S2
Hassan walks through the mountains again, and after five days he arrives home.	Hassan crosses the mountains again. He arrives home after five days.	L1, S2
The next night, Walid makes a plan. '100 gold coins in return for a glass bottle,' " he thinks. 'If I give the prince my mother's silver ring, perhaps I can get 1,000 gold coins!'	Walid makes a plan the next night. He thinks '100 gold coins for a glass bottle. Maybe I can get 1,000 gold coins for my mother's silver ring.'	S1, G1

Fonte: Elaboração própria

SUGESTÕES PARA O TEXTO E8

(Elementary – página 55)

Quadro 78 – Sugestões de estratégias para o texto E8

ORIGINAL	SUGESTÃO DE REESCRITA	ESTRATÉGIA(S)
He was a soldier during World War II, and was in a prison camp in Germany from 1944 to 1945.	He was a soldier during World War II. He was in a prison camp in Germany from 1944 to 1945.	S2
His first novel was Piano Player, published in 1952, but it wasn't a success. However, his novel Slaughterhouse-Five, an anti-war book, was a best-seller, and was made into a film in 1972.	Kurt's first book was 'Piano Player'. The book was published in 1952, but it wasn't a success. His book 'Slaughterhouse-Five' was a best-seller. It was an anti-war book. A director made a movie about the book in 1972.	L2, S2, G2, S4
In later life, he was interested in politics and painting.	Kurt was interested in politics and painting when he was older.	L1, S2
His father was a caretaker and his mother was a lift operator.	Billy's father was a caretaker (someone who takes care of a place). Billy's mother was an elevator operator.	L1, L5, L7, S2
He was a theatre actor for 15 years, and in 1959 he was in his first film.	Billy was a theatre actor for 15 years. He was in his first film in 1959.	L2, S1, S2
He was very successful in the 1980s, when he was in two Star Wars films, The Empire Strikes Back and Return of the Jedi.	Billy was very successful in the 1980s. He was in two Star Wars films: The Empire Strikes Back and Return of the Jedi.	L2, S2
Before he was an actor, Williams was at a fine arts academy in New York, and this self-portrait is in the National Portrait Gallery in Washington D.C.	Billy studied at a fine arts academy in New York before he was an actor. His self-portrait is in the National Portrait Gallery in Washington D.C.	L2, S1, S2
She was born Adele Laurie Blue Adkins in London in 1988.	Her original name is Adele Laurie Blue Adkins. She was born in London in 1988.	S2, G5
When she was 14, her musical tastes were very different from most teenagers - she was a big fan of Billie Holiday and Eminem.	When Adele was 14 years old, her musical tastes were very different from most teenagers. She was a big fan of Billie Holiday and Eminem.	L2, S2
She was a performing arts student at the BRIT School in London, Where she was a classmate of singers Leona Lewis and Jessie J.	Adele studied performing arts at the BRIT School in London. Adele was classmates of singers Leona Lewis and Jessie J.	L2, S2, G5
Her first album, 19, in 2008, was an immediate success and she is now a major star.	Adele's first album is called '19'. The album came out in 2008. It was an immediate success and she is now a big star.	L1, L2, S2, G4
Adele isn't a painter; this self-portrait was for an online art project called Face Britain.	Adele isn't a painter. Her self-portrait was for an online art project called 'Face Britain'.	L6, G2

Fonte: Elaboração própria

SUGESTÕES PARA O TEXTO E14

(*Elementary* – página 83)

Quadro 79 – Sugestões de estratégias para o texto E14

ORIGINAL	SUGESTÃO DE REESCRITA	ESTRATÉGIA(S)
'Come in,' said a voice.	A voice said, 'come in.'	S1
'Madame Yolanda is not here today,' said the man. 'But don't worry. I'm going to tell you about your future. What questions do you want to ask?'	The man said, 'Madame Yolanda is not here today. But don't worry. I'm going to tell you about your future. What questions do you want to ask?'	S1
He turned over the second card.	He flipped the second card.	L1
'Let's look at the next card,' said the fortune teller.	The fortune teller said, 'Let's look at the next card.'	S1

Fonte: Elaboração própria

SUGESTÕES PARA O TEXTO E15

(*Elementary* – página 85)

O texto E15 é considerado um dos mais complexos no nosso *corpus* de estudo (não só pelas métricas, mas também comprovadamente em sala de aula). Por esse motivo, especificamente para esse texto, apresentaremos uma proposta de reescrita completa de forma diferente ao que estamos fazendo até o momento. No quadro abaixo (Quadro 80) apresentamos uma proposta de reescrita simplificada, seguida por uma lista das estratégias adotadas. Em negrito, destacamos na proposta de reescrita as palavras que apresentaríamos em um glossário inglês-português.

Quadro 80 – Proposta de reescrita e estratégias utilizadas para o texto E15

PROPOSTA DE REESCRITA SIMPLIFICADA
<p>THE WORLD'S MOST UNUSUAL HOTELS</p> <p><i>Do you like nature?</i></p> <p>'Verana' is a spa hotel in Mexico. It is a place for people who love nature. The spa hotel has beautiful views on all sides.</p> <p>There are ten houses for guests and four new buildings. The rooms have balconies above the trees. You can see 'Bay of Banderas' from the balconies.</p> <p>The hotel has an infinity pool. 'Verana' offers a variety of different spa treatments, as well as yoga classes.</p> <p>You can go watch whales or fish for tuna. You can also explore the area on foot.</p> <p>Verana is a perfect place for a wedding, honeymoon, or both. But remember, you can only get there by boat.</p> <p><i>Would you like to sleep in a mine?</i></p> <p>You can book a suite in 'Sala Silvermine' in Sweeden. It is a room under the ground. There, you can enjoy the deepest bedroom in the world. It is 155 metres below the</p>

surface. The corridors are cold and dark, but the rooms are warm (18°C). The light comes from candles in beautiful silver **candlesticks**.

You can explore the caves and magical lakes. A guide can help you. There are no other guests, so it's not for nervous people. Phones don't work, but you have a radio for emergencies.

The **hotel crew** brings you breakfast to your room in the morning. 'Sala Silvermine' is a perfect place for romantic couples who like a bit of adventure.

Do you like dogs?

You don't need to travel to the Arctic Circle to spend time with Siberian huskies.

People who love dogs can sleep in cabins next to dog's **kennels** at the 'Husky Lodge' in Swiss. The cabins are warm because they are **heated** with wood fires.

During the day, the dogs train their running. You can join them. The dogs pull **sledges** in the winter and bikes in the summer.

The 'Husky Lodge' has an excellent restaurant and three rooms for small meetings (if you need to work).

If the cabins are too expensive for you, the 'Husky Lodge' has a campsite too (but only in the summer).

ESTRATÉGIA(S)

L1	substituição de léxico por sinônimo mais simples
L2	substituição de pronome por referencial
L3	remoção de palavras desnecessárias
L8	eliminação de palavra complexa não-essencial
S1	alteração de ordem das frases
S2	divisão de orações
S3	alteração de pontuação
G2	separação de título de obras ou nomes de locais por meio de aspas únicas
G5	reescrita de frase complexa
G6	reordenação de informações
G7	eliminação de informação não-necessária
G8	inserção de tópicos frasais

Fonte: Elaboração própria

SUGESTÕES PARA O TEXTO E16

(*Elementary* – página 86)

Quadro 81 – Sugestões de estratégias para o texto E16

ORIGINAL	SUGESTÃO DE REESCRITA	ESTRATÉGIA(S)
When I eat out with my husband or friends, we usually share, because food portions are huge and we don't want all that food.	When I eat out with my husband or friends, we usually share our food. The food portions are huge, and we don't want all that food.	S2
In the winter, you often see people in the street or on the subway who are wearing just a T-shirt when it's really cold outside.	You often see people in the street or in the subway who are wearing just a t-shirt when it's winter and it's really cold outside.	S1, S2

When people speak to their boss (and to their teachers) they call them by their first name.	People call their bosses and teachers by their first name.	G5
The American President's salary is about the same as a doctor's.	The American President's salary is similar to a doctor's salary.	L1, G4
Some hospitals are like smart hotels - but a lot of people can't go there because they don't have medical insurance.	Some hospitals are similar to smart hotels. Lots of people can't go to these hospitals because they don't have medical insurance.	S2, L2
Here, they tell you everything.	In America, the doctors tell you everything.	L2, L12
I couldn't believe the tipping system, for example, in places like the hairdresser's.	I couldn't believe the tipping system (when you give extra money to people who serve you). One example is at the hair salon.	L5, S2
But I was on the subway here in New York, going home late at night, when a group of people got on. They were very loud - I think they were drunk - and we were quite frightened.	One day I was on the subway here in New York. I was going home, and it was late at night. A group of people got on. These people were very loud. I think they were drunk. I was really scared.	L1, L2, S1
My husband is from Ecuador, and when he first arrived, he was really surprised by his salary.	My husband is from Ecuador. He was really surprised by his salary when he first arrived.	S1, S2

Fonte: Elaboração própria

SUGESTÕES PARA O TEXTO E17

(*Elementary* – página 88)

Quadro 82 – Sugestões de estratégias para o texto E17

ORIGINAL	SUGESTÃO DE REESCRITA	ESTRATÉGIA(S)
During my studies, I worked in the Canary Islands and in Argentina and I studied for two semesters in the UK.	During my studies, I worked in the Canary Islands and in Argentina. I also studied for two semesters in the UK.	S2
After I finished, I decided that I didn't want to work for the rest of my life.	I decided that I didn't want to work for the rest of my life after I finished my studies.	S1, G4
I was 21 and couldn't imagine sitting in an office for the next 40 years and more.	I was 21 years old and couldn't imagine sitting in an office for the next 40 years and more.	G4
I spent Christmas on St Kilda Beach in Melbourne, I worked in a pub and then I travelled all over Australia and New Zealand.	I spent Christmas on St. Kilda Beach in Melbourne. I worked in a pub and then I travelled around Australia and New Zealand.	L1, S2
After coming back I got a job as a project assistant in an IT company - an office job.	I got a job as a project assistant in an Information Technology (IT) company after I came back to Germany. It was an office job.	L1, S1, S2

I was happy to have a routine again, but I spent all my holidays travelling - to Amsterdam, Istanbul, Portugal, Switzerland, Chile, Prague, and Morocco.	I was happy to have a routine again, but I spent all my holidays travelling. I went to places like Amsterdam, Istanbul, Portugal, Switzerland, Chile, Prague, and Morocco.	S2
Some months passed, but finally I realized what I wanted to do with my life, at least for now.	After some months, I finally realized what I wanted to do with my life, at least for now.	G5
Having experiences is very important to me – more important than working in an office all week to pay for a flat that I hardly ever spend time in, and to buy things I don't really need.	Experiences are very important to me. They are more important than working in an office all week to pay for an apartment that I am almost never in. Experiences are more important than buying things I don't really need.	L3, L11, S2, G4
Three months ago, I left my job. I decided to work in different jobs, sometimes here, sometimes abroad, sometimes as a volunteer, sometimes for money.	I left my job three months ago. I decided to work in different jobs, sometimes in Germany, sometimes out of Germany, sometimes as a volunteer, and sometimes for money.	S1, S2, L2
At the beginning of every year I make a bucket list' of 100 things I would like to do. These are some of them.	I make a list of things I want to do when every new year starts. I always write 100 things I would like to do.	S1, S2, G5
You can probably do them in your country - you don't need to travel.	Maybe, you can do these things in your country. You don't need to travel to do them.	L1, L2, S2, G4

Fonte: Elaboração própria

SUGESTÕES PARA O TEXTO E18

(Elementary – página 91)

Quadro 83 – Sugestões de estratégias para o texto E18

ORIGINAL	SUGESTÃO DE REESCRITA	ESTRATÉGIA(S)
"What's that?' our intern Francesca asked.	Francesca, our intern, asked, "What's that?"	S2
For all of you who are too young to remember, teletext was sort of like a very, very basic internet on your TV.	If you are too young to remember, 'teletext' was like a very basic internet on your TV.	L8, G2, G4
Sending an email is so easy these days, but some years ago, when you needed to send something to somebody urgently and the post was too slow, you sent a fax.	Sending an email is so easy these days. You had to use a fax machine if you wanted to sent something to someone urgently and the post was too slow.	G5
When they were introduced in the 1970s people thought they were a technological miracle.	People thought the fax machines were a technological miracle when they were introduced in the 1970's.	S1, L2

If you needed to book an appointment with your hairdresser or phone a local Chinese takeaway, what did you do?	What did you do if you needed to book an appointment with your hairdresser? Or if you needed to call a local Chinese restaurant?	L1, S1, S2
You looked up the phone number or address in a big fat yellow book - the Yellow Pages.	You searched the phone number or address in a book called 'The Yellow Pages'. It was a big fat yellow book.	L1, S1, G2
When you needed information for your studies, or to win an argument, you looked it up in an enormous encyclopedia - sometimes ten or twelve books of information in alphabetical order.	When you needed information for your studies, or to win an argument, you could use an encyclopedia. The encyclopedias were big books (sometimes ten to twelve volumes). They were organized in alphabetical order.	L1, S2
This is how we listened to music.	A 'Walkman' is what we used to listen to music.	L2, G5
And if you didn't have the cassette or CD of a song, you couldn't listen to it.	You couldn't listen to a song if you didn't have the cassette or CD.	L2, S1
If you had a meeting or job interview, or you wanted to go to a new restaurant in a part of town you didn't know, what did you always take with you?	What did you always take with you if you had a meeting in a part of the city that you didn't know? Or if you wanted to find a new restaurant?	S1, S2, G5

Fonte: Elaboração própria

SUGESTÕES PARA O TEXTO E19

(*Elementary* – páginas 98 e 99)

Quadro 84 – Sugestões de estratégias para o texto E19

ORIGINAL	SUGESTÃO DE REESCRITA	ESTRATÉGIA(S)
I hardly ever listen to music at home - I prefer going to concerts.	I hardly ever listen to music at home. I prefer going to concerts.	S2
If I am working, I get up one hour before I have to leave the house.	I get up one hour before I have to leave the house if I am working.	S2
If I am not working, and I went to bed late the night before, I get up at about 10.00 in the morning.	When I am not working and went to bed late the night before, I get up at about 10.00 a.m..	S2
I can very easily spend three or four hours on the internet, answering emails, reading the news, etc.	It is ok for me to spend three or four hours on the internet. I answer emails, read the news, etc.	S2
Perhaps the living room where I cook and eat, and from where I can see the River Thames in London.	Maybe the living room where I cook and eat. From there, I can see the River Thames in London.	L4 S2
Yes – many things, e.g. to sing well, to play the piano, and to speak foreign languages.	Yes! Many things! For example, to sing well, to play the piano, and to speak foreign languages.	L1, S2, S3

The first money I earned as a professional actor was when I was a student at Cambridge University in 1959.	The first money I got as a professional actor was in 1959. I was a student at Cambridge University.	S1, S2
--	---	--------

Fonte: Elaboração própria

SUGESTÕES PARA O TEXTO E20

(*Elementary* – página 101)

Quadro 85 – Sugestões de estratégias para o texto E20

ORIGINAL	SUGESTÃO DE REESCRITA	ESTRATÉGIA(S)
I go here quite often because it's very near my flat.	I go to 'Pizza West' quite often because it's close to my apartment.	L1, L12
The atmosphere is great, perfect for a Friday or Saturday night.	The place is great. It's perfect for a Friday or Saturday night.	L1, S2
Very expensive.	'Pizza West' is very expensive.	G4
In most pizza places we usually pay about £10 for a pizza; here it's €15 for more or less the same thing!	We usually pay about £10 for a pizza in most pizza places. The pizza is €15 for more or less the same thing at 'Pizza West' ⁷⁶ .	S1, L12
We've been here a lot and we always enjoy it.	We've been to 'Pizza West' a lot, and we always enjoy it.	L12
We went yesterday for an early dinner with our three-year-old daughter and they really made us feel at home - and they produced a delicious birthday cake for her.	We went to 'Pizza West' yesterday for an early dinner. We took our daughter. She is three years old. The restaurant workers made us feel at home. They made a delicious birthday cake for our daughter.	L2, S1, S2
We booked a table, but when we got there they asked us to sit with a lot of other people at a big table.	We booked a table. They made us sit with a lot of people at a big table when we got there.	S1, S2
We didn't want to because there were only two of us.	We didn't want to sit with other people because there were only two of us.	G4
Then they asked us to wait until a different table was free, so we waited at the bar for 45 minutes! But I recommend it because the food is excellent!	The waiter asked us to wait until a different table was free, so we waited at the bar for 45 minutes! But I recommend 'Pizza West' because the food is excellent!	L2, S1
The pizza here isn't the best I've tasted.	The pizza at 'Pizza West' isn't the best I've tasted.	L2
The service is a bit slow and the place is very busy - it can often be really noisy.	The service is a bit slow and the place is very busy. The restaurant can often be really noisy.	S2

⁷⁶ Recomendamos que a utilização das estratégias L2 e L12 sejam feitas com moderação. Apresentamos a aplicação da estratégia para todas as ocorrências de pronomes e elementos dêiticos, mas é necessário regular a aplicação da estratégia.

Not the place for a romantic dinner - it's more of a fun, exciting place.	'Pizza West' is not the place for a romantic dinner. The restaurant is for a fun and exciting experience.	L2, S2
---	---	--------

Fonte: Elaboração própria

Na próxima seção, apresentamos as sugestões para os textos do livro *Pre-intermediate*.

9.2 SUGESTÕES DE ESTRATÉGIAS DE ACESSIBILIDADE PARA OS TEXTOS DO LIVRO *PRE-INTERMEDIATE*

Visto que os textos da presente seção são componentes de um livro utilizado com estudantes que estão, supostamente, em um nível pré-intermediário, julgamos que as indicações de estratégias devam ser um pouco diferentes em frequência e em aplicações daquelas apresentadas para os textos do *Elementary*. Isso não quer dizer, que as estratégias serão outras, mas que a frequência de aplicação e a forma como serão aplicadas devem considerar a “bagagem” de conhecimento que os alunos têm; caso contrário, conforme já enfatizamos, a manutenção linguística não será eficiente.

Portanto, antes de apresentar as sugestões para os textos, destacamos alguns pontos importantes:

1. A estratégia L1 (substituição lexical por sinônimo mais simples) ainda será aplicada (uma vez que, conforme destacamos no início desta tese, o léxico tende a ser a maior característica de complexidade, conforme apontam diversos autores da área). Entretanto, ela será sugerida com menor frequência e para unidades lexicais que considerarmos, de fato, altamente complexas;

2. A substituição de pronomes por referenciais (estratégia L2) somente será sugerida quando o pronome estiver muito distante da palavra a qual se refere ou em caso de possível ambiguidade (caso a palavra possa “apontar” para outras unidades no parágrafo);

3. A redução de advérbios (estratégia L10) e/ou a substituição dessa classe de palavras só será sugerida até o texto P12, visto que ‘advérbios’ é o conteúdo lexical trabalhado na unidade 8 (texto P13 em diante). A partir do P13, essa estratégia será aplicada somente caso o advérbio seja pouco frequente;

4. Somente sugeriremos alteração de ordem de frase para casos nos quais uma frase subordinada seja apresentada antes da principal. Não iremos sugerir realocação de adjuntos para suas posições canônicas (conforme fizemos para o

Elementary), visto que esse conteúdo é trabalhado não só no nível elementar, mas também no início do pré-intermediário;

5. Para casos de frases coordenadas, somente iremos sugerir a estratégia de divisão de orações (estratégia S2) caso a oração seja muito extensa;

6. No caso de orações complexas (formadas simultaneamente por coordenação e subordinação), a estratégia de divisões de orações irá ser sugerida somente para casos muito longos e somente para separar a subordinada das demais.

Além disso, conforme apontamos na nossa análise das métricas dos textos do *Pre-intermediate*, os traços de complexidade são menos frequentes e, por vezes, acreditamos que seria possível aumentar a complexidade em um nível textual ou outro para potencializar a manutenção linguística.

Relembramos que mesmo diante das observações acima, as sugestões apresentadas para os textos do livro *Elementary* e as orientações do capítulo 10 desta tese ('Orientações para Acessibilidade Textual Linguodidática') podem ser aplicadas para casos de estudantes com limitações linguísticas, neuroatípicos ou sempre que o professor julgar necessário.

Iniciaremos a apresentação de sugestões pelo texto P1.

SUGESTÕES PARA O TEXTO P1

(*Pre-intermediate* – página 8)

Quadro 86 – Sugestões de estratégias para o texto P1

ORIGINAL	SUGESTÃO DE REESCRITA	ESTRATÉGIA(S)
You probably think this is a parent asking the question to their child.	Maybe, you think this is a parent asking the question to their child.	L1
Clint Bouchez, a 52-year-old businessman from Reigate, Surrey, got divorced ten years ago.	Clint Bouchez is 52 years old. He is a businessman from Reigate, Surrey. He got divorced ten years ago.	S2, G4
His daughter, 26-year-old Charlotte, doesn't want her dad to end up alone - her mum remarried five years ago, and Charlotte has a boyfriend.	Charlotte is Clint's daughter, and she is 26 years old. Charlotte doesn't want her dad to be alone. Her mum remarried five years ago, and Charlotte has a boyfriend.	L1, S2
They often go out together in the evening, they're planning a trip to India for later in the year, and they're both Bruce Springsteen fans.	They often go out together in the evening. They're planning a trip to India for later in the year. They're both Bruce Springsteen fans.	S2

But Charlotte thinks that Clint needs a new partner, and he agrees.	But Charlotte thinks that Clint needs a new partner. Clint agrees.	S2
'My dad is the perfect man,' she says. 'He's warm, he's generous, he knows how to look after a woman, how to treat a woman.'	Charlotte says, "My dad is the perfect man. He's warm and generous. My dad knows how to care for a woman and how to treat a woman."	S1
"I don't mind spending Friday and Saturday nights with my dad and that says a lot - I'm only 26."	"I don't mind spending Friday and Saturday nights with my dad and that says a lot because I'm only 26"	S3
It's difficult now just to walk over to a woman and ask her "Can I buy you a drink?" he says. Before, when you met people face to face, there was sometimes a "spark", a feeling of romantic destiny - you don't get that from someone's profile picture.'	Clint says, "It's difficult now just to approach a woman and ask her if I can buy her a drink. Before, when you met people face to face, there was sometimes a "spark", a feeling of romantic destiny. You don't get that spark from someone's profile picture."	L1, S1
I'm always hopeful,' he says. 'I really believe that sooner or later I'm going to find "the one".'	Clint says, "I am always hopeful. I really believe that sooner or later I'm going to find 'the one'."	S1

Fonte: Elaboração própria

SUGESTÕES PARA O TEXTO P2

(Pre-intermediate – página 14)

Quadro 87 – Sugestões de estratégias para o texto P2

ORIGINAL	SUGESTÃO DE REESCRITA	ESTRATÉGIA(S)
It's the mantra we always say to ourselves when we go on holiday to make sure we haven't forgotten anything.	It's the mantra we always say to ourselves when we go on holiday. We need to make sure we haven't forgotten anything.	S2
One day, we climbed a mountain - well, it wasn't really a climb, but it was a long walk - and it took about two hours to get to the top.	One day, we climbed a mountain - well, it wasn't really a climb, but it was a long walk. It took about two hours to get to the top.	S2
Then we went down again, and when we got back to the car, I wanted to take another photo, but I couldn't find my phone - it wasn't in my bag.	Then we went down again. When we got back to the car, I wanted to take another photo, but I couldn't find my phone because it wasn't in my bag.	S2, S3
My friends said 'OK, but you can't go on your own', so in the end we all went up again, which was another two hours.	My friends said, "OK, but you can't go on your own", so in the end we all went up again. It took us another two hours.	S2

Fonte: Elaboração própria

SUGESTÕES PARA O TEXTO P3

(Pre-intermediate – página 16)

Quadro 88 – Sugestões de estratégias para o texto P3

ORIGINAL	SUGESTÃO DE REESCRITA	ESTRATÉGIA(S)
The Guardian newspaper has a weekly feature called That's me in the picture, where people describe famous photos they were in. This photo was sent in by Jane Rangeley.	The Guardian newspaper has a weekly column called 'That's me in the picture'. People describe famous photos they were in. This photo was sent by Jane Rangeley.	L1, S2, G2
When I got home, I immediately started looking for a job in Paris.	When I got home, I quickly started looking for a job in Paris.	L1
There was an owl there, maybe escaped from the zoo, and some little birds were attacking it.	There was an owl there. Maybe it escaped from the zoo, and some little birds were attacking it.	S2, G4
And although my French boyfriend and I broke up in the end, we're still in touch.	My French boyfriend and I broke up in the end, but we're still in touch.	L1

Fonte: Elaboração própria

SUGESTÕES PARA O TEXTO P4

(Pre-intermediate – página 21)

Quadro 89 – Sugestões de estratégias para o texto P4

ORIGINAL	SUGESTÃO DE REESCRITA	ESTRATÉGIA(S)
Holiday time is the one time we all take photos. We're in a strange new place, the sun is shining, and we want to record our surroundings and happy memories.	Holiday time is the one time we all take photos. We're in a strange new place, the sun is shining, and we want to record the places we visit and happy memories.	L11
There are times, in the local market, for example, when having lots of people in the photo can be a good thing, but most of the time I prefer my photos to be less crowded.	For example, there are times in the local market when having too many people in the photo can be a good thing, but most of the time I prefer my photos to be with less people.	L1, S1
Although I am a very enthusiastic photographer, I often buy a nice postcard of the places I visit - much better than going home with pictures that aren't very good.	I am a very enthusiastic photographer, but I often buy a nice postcard of the places I visit. It's much better than going home with pictures that aren't very good.	L1, S2
We spent all day looking at the paintings, and we were walking back to the hotel when we saw someone feeding the birds.	We spent all day looking at the paintings. We saw someone giving food to the birds when we were walking back to the hotel.	L11

Fonte: Elaboração própria

SUGESTÕES PARA O TEXTO P5

(*Pre-intermediate* – página 26)

Quadro 90 – Sugestões de estratégias para o texto P5

ORIGINAL	SUGESTÃO DE REESCRITA	ESTRATÉGIA(S)
The aim of the game is to complete the alphabet wheel by saying the correct word for each letter in five minutes.	The objective of the game is to complete the alphabet wheel. You have to say the correct word for each letter in five minutes.	L1
The presenter gives the contestant definitions for words starting with the letters, for example, 'A - the place where you catch a plane' = 'Airport'.	The presenter gives the participants definitions for words starting with the letters. For example, "A - the place where you catch a plane" = "Airport".	L1
The presenter continues round the wheel to Z, then starts from the beginning again.	The presenter continues round the wheel to Z, then starts from the start again.	L1

Fonte: Elaboração própria

SUGESTÕES PARA O TEXTO P6

(*Pre-intermediate* – página 34)

Quadro 91 – Sugestões de estratégias para o texto P6

ORIGINAL	SUGESTÃO DE REESCRITA	ESTRATÉGIA(S)
Psychologist Judi James, one of the organizers of the survey, said, 'When some people read their friends' posts and see their photos on Facebook, Twitter, and Instagram, they begin to feel jealous of them. They think that their friends are having a much more exciting life. So they invent details about their own weekend.'	Psychologist Judi James is one of the organizers of the survey. She said, "When some people read their friends' posts and see their photos on social media, they start to feel jealous of them. They think that their friends are having a much more interesting life. So they invent details about their own weekend."	L1, S2
One of the main reasons people give for inventing these stories is to have something to talk about on Monday morning at work or at school. This is because they don't want other people to think that they have a boring life.	One of the main reasons people give for inventing these stories is to have something to talk about on Monday morning at work or at school. They invent stories because they don't want other people to think that they have a boring life.	L2
The most popular lies people tell are that they went to a party or went away for the weekend, and 3% of people even put on fake tan on Sunday night to make their colleagues think they've had a weekend away in the sun.	The most popular lies people tell are that they went to a party or went away for the weekend. 3% of people even put on fake tan on Sunday night to make their colleagues think they've had a weekend away in the sun.	S2

One person in ten invents a romantic break ⁷⁷ with their partner.	One person in ten invents a romantic time with their partner.	L1
Judi James said, "Social media is becoming increasingly important in our lives, and it seems we're living one life online and another in reality: So when you are feeling jealous on a Saturday night because your best friend is having a romantic dinner with her boyfriend in Venice, stop and ask yourself, 'Is she really there, or is she just sitting at home feeling bored like me?'"	Judi James said, "Social media is becoming more important every day. It seems we're living one life online and another in reality. You see your best friend having a romantic dinner on a Saturday night. Then, you feel jealous. Stop and ask yourself, "Is she really there?". Maybe, she is sitting at home feeling bored, too.	L11, S2

Fonte: Elaboração própria

SUGESTÕES PARA O TEXTO P7

(Pre-intermediate – página 37)

Quadro 92 – Sugestões de estratégias para o texto P7

ORIGINAL	SUGESTÃO DE REESCRITA	ESTRATÉGIA(S)
Every three months the OED (Oxford English Dictionary) publishes updates to its online dictionary. One recent update contained 900 new words, new expressions, or new meanings for existing words.	The Oxford English Dictionary publishes updates to its online dictionary every three months. One recent update had 900 new words, new expressions, or new meanings for words.	L1, S1
New words are created in many different ways. We can make a new word by combining two words, like gastropub (gastronomy + pub) or vlog (video + blog).	New words are created in many different ways. We can make a new word by combining two words, like Facebook (face + book) or vlog (video + blog).	L1, G9
Sometimes we put two words together in a new way, for example road rage or selfie stick.	Sometimes we put two words together in a new way, for example 'online call' or 'selfie stick'.	L1, G2
We also find that nouns can change into verbs. Take the word text. Text was always a noun (from about 1369, according to the OED), but it is now very common as a verb, to text somebody.	We also find that nouns can change into verbs. For example, the word 'text'. 'Text' was always a noun (from the year 1369, according to the Oxford English Dictionary). Now, 'text' is a very common verb. The meaning is 'to text somebody' (to send a message to somebody).	L1, L5, S2, G2, G4
Other new words already existed but with a different meaning. For example, tweet was the noise that	Other new words already existed but with a different meaning. For example, 'tweet' was the noise	G1, G2

⁷⁷ Ainda que a palavra 'break' não seja complexa, com grande frequência os estudantes interpretam como 'break up' na leitura em sala de aula.

a bird makes, but now we use it more often (as a verb or a noun) for a message that people put on the social networking site Twitter.	that a bird makes, but now we use it more often for a message that people put on the social media site Twitter.	
A lot of new words come from the names of brands or companies, for example we Skype each other and we google information.	A lot of new words come from the names of brands or companies. For example we have a 'Zoom' meeting and we 'Google' information.	S2, G2, G9
We also need more general words to describe new technology or new gadgets: wi-fi, ringtone, and smartphone are some examples.	We also need more general words to describe new technology or new gadgets: wi-fi, selfie, and smartphone are some examples.	G9
The word brunch (breakfast + lunch) first appeared in 1896, newspaper (news + paper) in 1667, and English speakers started to use the word café (from French) in the late 19th century.	The word 'brunch' (breakfast + lunch) first appeared in 1896. 'Newspaper' (news + paper), in 1667. English speakers started to use the word 'café' (from French) in the late 19th century.	S2, G2

Fonte: Elaboração própria

SUGESTÕES PARA O TEXTO P8

(Pre-intermediate – página 38)

Quadro 93 – Sugestões de estratégias para o texto P8

ORIGINAL	SUGESTÃO DE REESCRITA	ESTRATÉGIA(S)
Whatever happened to patience?	What happened to patience?	L1
With faster broadband, instant searches, and immediate downloads, we expect things to happen immediately, and if they don't we get impatient.	With faster internet, instant results, and immediate downloads, we want things to happen instantly. If things don't happen, we get impatient.	L1, S2
The growth of mobile phone apps like Tinder has changed the way we meet new people.	The great number of apps like 'Tinder' has changed the way we meet new people.	L3, L11, G2
British people's most hated time-wasting activities are (in order):	The things British people most hate waiting for are (in order):	G5
waiting for a replacement credit card	waiting for a new credit card	L1
waiting for a delayed flight	waiting for a late flight	L1

Fonte: Elaboração própria

SUGESTÕES PARA O TEXTO P10

(Pre-intermediate – página 47)

Quadro 94 – Sugestões de estratégias para o texto P10

ORIGINAL	SUGESTÃO DE REESCRITA	ESTRATÉGIA(S)
<p>But in fact there's a kind of pessimism - called 'defensive pessimism' - that can lead to very positive results, according to Julie K. Norem, a professor of psychology at Wellesley College, Massachusetts. 'Defensive pessimism is a strategy used in specific situations to manage anxiety, fear, and worry,' says Norem. Studies show that 30 to 35 per cent of Americans use it to help them in their lives, and they're often very successful people.</p>	<p>Julie K. Norem is a professor of psychology at Wellesley College in Massachusetts. She says that there's a type of pessimism that can lead to very positive results. It's called 'defensive pessimism'. It's a strategy used in specific situations to control anxiety, fear, and worry. Studies show that 30% to 35% Americans use 'defensive pessimism' to help them in their lives, and they're often very successful people.</p>	<p>L1, L2, S1, S2, G4, G10</p>
<p>Then they look at each possible problem and plan how to avoid it. So for the exam situation, they go to bed early and have a good night's sleep; they find out in advance exactly where the exam is; they eat a good breakfast, and take lots of pens and pencils, and a bottle of water; and they leave home early. That puts them in control, and it means that the exam will be better than for an optimist, who just thinks Oh, everything will be fine!' Because sometimes everything goes wrong, and it's good to be prepared.</p>	<p>Then 'defensive pessimists' look at each problem and plan how to avoid it. So for the test situation, they do the following:</p> <ol style="list-style-type: none"> 1. Go to bed early and have a good sleep. 2. Discover before the day of the test exactly where the exam. 3. They eat a good breakfast. 4. They take lots of pens and pencils, and a bottle of water. 5. They leave home early. <p>Doing these things puts them in control. It means that the exam will be better than for an optimist, who just thinks "Oh, everything will be fine!". Because sometimes everything goes wrong, and it's good to be prepared.</p>	<p>L1, L2, L3, S2, S3, G2, G11</p>

Fonte: Elaboração própria

SUGESTÕES PARA O TEXTO P11

(Pre-intermediate – página 53)

Quadro 95 – Sugestões de estratégias para o texto P11

ORIGINAL	SUGESTÃO DE REESCRITA	ESTRATÉGIA(S)
<p>Mahalia Jackson was a musical legend who helped to bring gospel music from the church to large audiences.</p>	<p>Mahalia Jackson was a musical icon. She helped to bring gospel music from the church to large audiences.</p>	<p>L1, S2</p>

She was Aretha Franklin's mentor, and in 1961 she was the first gospel singer to win a Grammy Award.	She was Aretha Franklin's mentor. In 1961, Mahalia was the first gospel singer to win a Grammy Award.	L2, S2
Mahalia often went with King on civil rights demonstrations and marches, including into the most hostile parts of the South, and she sang at the events.	Mahalia often went with Martin Luther King on civil rights protests and marches (including into the most aggressive parts of the South.) She sang at the events.	L1, S2, S3, G4
As a presenter from National Public Radio said, 'her voice became the soundtrack of the Civil Rights Movement.	A presenter from 'National Public Radio' said, "Mahalia's voice became the soundtrack of the 'Civil Rights Movement'."	L2, L8, G2
Mahalia was with King, who was going to give a five-minute speech. Before the day itself, he and his advisers decided what he was going to say.	Mahalia was with Martin Luther King. He was going to give a five-minute speech. Before the day of the march, Martin and his advisers decided what he was going to say.	L2, S2
King began speaking to an audience of more than 250,000 people. But towards the end, he felt that his speech was not going well. Suddenly Mahalia shouted 'Tell them about the dream. Martin. Tell them about the dream'.	Martin started speaking to an audience of more than 250,000 people. But almost at the end, he felt that his speech was not going well. Then, Mahalia shouted "Tell them about the dream, Martin. Tell them about the dream."	L1
She knew about it because she was at a previous event in Detroit, in June of that year, where King talked about his dream for African Americans.	Mahalia knew about the dream because she was at an earlier event in Detroit in June of that year. In Detroit, Martin talked about his dream for African Americans.	L1, L2, S2
It was, in the words of Kings adviser Clarence Jones, 'one of the world's greatest gospel singers shouting to one of the world's greatest preachers'.	Clarence Jones, one of Martin's advisers, said, "One of the world's greatest singers was shouting to one of the world's greatest ministers."	L1, S1, S2
King looked at Mahalia. Then he threw away his written speech, and looked at the audience. 'I have a dream...'	Martin looked at Mahalia. Then he discarded his written speech, and looked at the audience. He started, "I have a dream..."	L1, L7, G4

Fonte: Elaboração própria

SUGESTÕES PARA O TEXTO P12

(Pre-intermediate – página 58)

Quadro 96 – Sugestões de estratégias para o texto P12

ORIGINAL	SUGESTÃO DE REESCRITA	ESTRATÉGIA(S)
That's been true for a long time.	British people are usually bad at languages, and that has been true for a long time.	L13, G4

Sometimes they try to say a couple of phrases in the local language, but they stop making an effort as soon as they discover that the waiter knows a little English.	Sometimes, British people try to say some phrases in the local language, but they stop making an effort when they discover that the waiter knows a little English.	L1, S3
Some British people who live abroad often spend all their time with other Brits, and never learn the language at all.	Some British people who live out of the UK often spend all their time with other British people, and they never learn the language at all.	L1, G4
This is partly true. In many multinational companies, for example, employees have to speak English as it is the company's official language of communication.	This is true in parts. In many multinational companies, for example, people have to speak English because it is the company's official language of communication.	L1, L11
Children only have to learn a language until they are 14. After that, they don't have to continue if they don't want to.	British children only have to learn a different language until they are 14 years old. After that, they don't have to continue learning if they don't want to.	G4
Thirty per cent of young people say that they don't want to carry on with a foreign language because 'it's too difficult.	30% of young people say that they don't want to continue learning a foreign language because it's too difficult.	L11, G10
'I'd like to visit Spain and Latin America in the future. If I go, I don't want to be the typical Brit who expects everyone else to speak English.' He did a one-month intensive course at a language school in London.	Max said, "I would like to visit Spain and Latin America in the future. If I go, I don't want to be the typical British person who wants everyone else to speak English." Max did an intensive course at a language school in London for a month.	L1, L2, L13, S1, G4
When his course finished he went to Madrid for the weekend to do his tests.	Max went to Madrid for the weekend to do his tests when his course finished.	L2, S1

Fonte: Elaboração própria

SUGESTÕES PARA O TEXTO P13 (Pre-intermediate – páginas 66 e 67)

Quadro 97 – Sugestões de estratégias para o texto P13

ORIGINAL	SUGESTÃO DE REESCRITA	ESTRATÉGIA(S)
"I've found where she lives," said the detective quietly.	"I've found where the girl lives," said the detective quietly.	L2
"Here is the address." Hartley took the piece of paper. On it were the words...	"Here is the address." Hartley took the piece of paper. On the paper were the words...	L2
"She moved there a week ago," said the detective. "I can follow	The detective said, "She moved there a week ago. I can follow her	S1

her if you want. It will only cost you \$7 a day and expenses.”	if you want. It will only cost you \$7 a day and expenses.”	
“One day's work,” said the detective. “Ten dollars.”	The detective answered, “One day's work. Ten dollars.”	S1
He went up the stairs, into her apartment, and saw her standing by the window.	Hartley went up the stairs. He entered the girl's her apartment, and saw her standing by the window.	L2, S2
Vivienne was about twenty-one.	Vivienne was about twenty-one years old.	G4
“Vivienne,” said Hartley angrily, “you didn't answer my last letter. It took me a week to find your new address! Why didn't you answer me? You knew I was waiting to see you and hear from you.”	Hartley said angrily, “Vivienne, you didn't answer my last letter. It took me a week to find your new address! Why didn't you answer me? You knew I was waiting to see you and hear from you.”	S1
“Mr Hartley,” she said slowly, “I don't know what to say to you. I understand all the advantages of your offer, and sometimes I feel sure that I could be happy with you. But, then sometimes I am less sure. I was born a city girl, and I am not sure that I would enjoy living a quiet life in the suburbs.”	Vivienne said, “Mr. Hartley, I don't know what to say to you. I understand all the benefits of your offer. Sometimes I believe that I could be happy with you. But sometimes I am less sure. I was born a city girl, and I don't know if I would enjoy a quiet life in the suburbs.”	L1, S1, S2, G5
“My dear girl,” said Hartley, “You will have everything that you want. You can come to the city for the theatre, for shopping, and to visit your friends as often as you want. You can trust me, can't you?”	Hartley said, “My dear girl...You will have everything that you want. You can come to the city for the theater, for shopping, and to visit your friends as often as you want. You can trust me, can't you?”	S1, S3
“I can trust you completely,” she said, smiling at him. “I know you are the kindest of men, and that the girl who you get will be very lucky. I heard all about you when I was at the Montgomerys'.”	“I can totally trust you. I know you are a very kind man. I know that the girl who you get will be very lucky. I heard all about you when I was at the Montgomerys' house,” said Vivienne.	L11, S1, G4
“Ah!” exclaimed Hartley, “I remember so well the evening I first saw you at the Montgomerys. I will never forget that dinner. Come on, Vivienne, promise me. I want you. Nobody else will ever give you such a happy home.”	“I remember the evening I first saw you at the Montgomerys' house so well. I will never forget that dinner. Come on, Vivienne, promise me. I want you. Nobody else will ever give you such a happy home,” said Hartley.	S1, G1
“Tell me, Vivienne, is there,” he asked, “is there - is there someone else?”	Hartley asked, “tell me, Vivienne... Is there someone else?”	L2, S1, S3, G1
“You shouldn't ask that, Mr. Hartley,” she said. “But I will tell you. There is one other person - but I haven't promised him anything.”	Vivienne said, “You shouldn't ask that, Mr. Hartley. But I will tell you. There is one other person, but I haven't promised him anything.”	S1, S3

"Vivienne," said Hartley masterfully, "You must be mine."	"Vivienne, you must be my girl," said Hartley.	L8, S1, G1
"Do you think for one moment," she said calmly, "that I could come to your home while Héloïse is there?"	Vivienne said with calmly, "Do you think for one moment that I could come to your home while Héloïse is there?"	S1

Fonte: Elaboração própria

SUGESTÕES PARA O TEXTO P14

(Pre-intermediate – página 69)

Quadro 98 – Sugestões de estratégias para o texto P14

ORIGINAL	SUGESTÃO DE REESCRITA	ESTRATÉGIA(S)
Do you know why the queues at the other checkouts in the supermarket always seem to move faster than yours?	Do you know why the lines at the other cashiers in the supermarket always seem to move faster than your line?	L1, L2, L7
A new book by David Andrews, Why Does the Other Line Always Move Faster?, has the answer: because you only notice how fast the other queues are moving when yours is moving slowly.	'Why Does the Other Line Always Move Faster?' is a new book by David Andres. It has the answer: because you only notice how fast the other lines are moving when your line is moving slowly.	L1, S1
If your queue moves fast, then you won't notice the slower queues at all, because you're busy unloading your trolley, putting things into bags, and paying.	If your line moves fast, then you won't notice the slower lines at all, because you're busy with your things, putting things into bags, and paying.	L1, G5
If there are three queues in the supermarket and you join the middle one, there is a two in three chance that one of the other queues will be the fastest, whereas yours only has a one in three chance.	If there are three lines in the supermarket and you join the one in the middle, there is a two in three chance that one of the other lines will be the fastest. Your line only has a one in three chance.	L1, L11, S2

Fonte: Elaboração própria

SUGESTÕES PARA O TEXTO P15

(Pre-intermediate – página 74)

Quadro 99 – Sugestões de estratégias para o texto P15

ORIGINAL	SUGESTÃO DE REESCRITA	ESTRATÉGIA(S)
Janet Leigh (1927-2004) was one of film director Alfred Hitchcock' favourite actresses, and was in	Janet Leigh (1927-2004) was one the favorite actresses of film director Alfred Hitchcock. She	S1, S2

more than 50 films and many TV series.	was in more than 50 films and many TV series.	
Her daughter Jamie Lee Curtis (1958-) is also a successful actress.	Janet's daughter, Jamie Lee Curtis (1958-), is also a successful actress.	L2, S3
She had two children from her third marriage to actor Tony Curtis. The marriage lasted 11 years. She then married again, and this marriage lasted for the rest of her life.	Neste texto, a estratégia de substituição de pronome por referencial não pode ser aplicada visto que o exercício usa propositalmente o pronome 'she' para que os estudantes utilizem os tempos verbais para identificar cada referência.	
She has been in many different kinds of films, including the comedies Trading Places, Fish Called Wanda, and True Lies, for which she won a Golden Globe Award for Best Actress in a Musical or Comedy.	She has been in many different kinds of films, including the comedies 'Trading Places', 'Fish Called Wanda', and 'True Lies'. She won a 'Golden Globe Award' for 'Best Actress in a Musical or Comedy' for 'True Lies'.	S2, G2
She has also starred in the comedy-horror TV series Scream Queens. In one episode she recreated the famous scream from Psycho.	She has also starred in the comedy-horror TV series 'Scream Queens'. In one episode, she recreated the famous scream from 'Psycho'.	S3, G2
She became Lady Haden-Guest when her husband became Baron Haden-Guest after the death of his father.	She became 'Lady Haden-Guest' when her husband became 'Baron Haden-Guest' after his father died.	L1, G2
She has written several bestselling children's books. She says she finds the inspiration for her writing all around her - in the experiences of her children, her godchildren (one is actor Jake Gyllenhaal), her friends and of course in her own life.	She has written a lot of bestselling children's books. She says she finds the inspiration for her writing all around her: in the experiences of her children, her godchildren (one is actor Jake Gyllenhaal), her friends and of course in her own life.	L1, S3

Fonte: Elaboração própria

SUGESTÕES PARA O TEXTO P17

(Pre-intermediate – página 85)

Quadro 100 – Sugestões de estratégias para o texto P17

ORIGINAL	SUGESTÃO DE REESCRITA	ESTRATÉGIA(S)
The hardest lesson to learn in sport is how to lose like a true sportsperson, without blaming your defeat on others.	"How to lose like a true athlete without blaming your lost on others" is the hardest lesson to learn in sports.	L1, S1
In the 1982 German Grand Prix, Nelson Piquet was winning the race.	Nelson Piquet was winning the 'German Grand Prix' in 1982.	S1
South Korean footballer Ahn Jung-Hwan scored the goal that sent	Ahn Jung-Hwan is a South Korean footballer. He scored the	L1, S2, G1, G5

Italy out of the 2002 World Cup when they beat them 2-1.	goal that eliminated Italy from the 'World Cup'.	
After the match, the president of the club, Luciano Gucci, announced that the player's contract would not be renewed. 'That gentleman will never set foot in Perugia again, Gaucchi said. 'I have no intention of paying a salary to somebody who has ruined Italian football.' Gaucchi later apologized, but Ahn Jung-Hwan left the club and never went back to an Italian club.	Luciano Gucci was the president of 'Perugia'. After the match when Italy lost, Luciano announced that Ahn Jung-Hwan's contract would not be renewed. He said, "That man will never play in Perugia again. I have no intention of paying a salary to somebody who has ruined Italian football." Later, Luciano said he was sorry, but Ahn Jung-Hwan left the club and never went back to an Italian club.	L1, S1, S2, G4
In the 2003 Athletics World Championship, the 100 metres runner Jon Drummond was disqualified for a false start. Drummond lay down on the track and began to cry. Two hours later his coach told journalists: 'He's still crying. We're making him drink water because he's becoming dehydrated.'	Jon Drummond, who is a 100-metres runner, was disqualified in the 2003 'Athletics World Championship' because of a false start. Jon lay down on the track and started to cry. Two hours later, Jon's coach told journalists, "He's still crying. We're making him drink water because he's getting dehydrated."	L1, L2, S1
In the 2008 Beijing Olympics, Angel Matos of Cuba was trying to win a bronze medal in Tae Kwon Do when the referee disqualified him for a technical error. Matos was furious, and after several minutes of arguing he kicked the referee in the head, and then attacked a Swedish judge. He was immediately banned from all competitions for life.	Angel Matos, from Cuba, was a Tae Kwon Do fighter in the 2008 'Beijing Olympics'. He was trying to win a bronze medal. In one fight, the referee disqualified him for a technical error. Angel was furious and kicked the referee in the head after several minutes of discussing. Then, Angel attacked a Swedish judge. He was immediately banned from all competitions for life.	L1, S1, S2, G4
In 2016, at the Olympics in Rio de Janeiro, the German men's football team lost to the host nation in the final. Brazil won their first ever Olympic gold medal in the event and the local fans were delighted. As they celebrated, one very disappointed German player, Robert Bauer, decided to show 7 fingers to the fans, to remind them of the time Germany beat Brazil 7-1 in the 2014 World Cup semi-final.	The German men's football team lost to Brazil in the final match at the 'Olympics in Rio de Janeiro' in 2016. Brazil won their first Olympic gold medal in the event and the local fans were really happy. While the fans celebrated, Robert Bauer (a very angry German player) decided to show 7 fingers to the fans to remind them of the time Germany scored 7 goals against Brazil in the 2014 'World Cup' semi-final.	L1, L2, L3, S1, G2, G4, G5

Fonte: Elaboração própria

SUGESTÕES PARA O TEXTO P18
(*Pre-intermediate* – páginas 86 e 87)

Quadro 101 – Sugestões de estratégias para o texto P18

ORIGINAL	SUGESTÃO DE REESCRITA	ESTRATÉGIA(S)
Mr. Baker, his English teacher, really liked poetry, and I'm sure Alex was inspired by him because his song lyrics are incredible.	Mr. Baker, Alex's English teacher, really liked poetry. I'm sure Alex was inspired by Mr. Baker because his song lyrics are incredible.	L2, S2, G4
But he didn't use to be very interested in music, he was much more interested in sports - he was possibly the best in the school at basketball.	But Alex didn't use to be very interested in music, he was much more interested in sports. He was possibly the best in school at basketball.	L2, S2
When he was 14 he broke his arm in my PE lesson. He was in hospital for a week, and we collected money and bought him a CD, so he was probably beginning to get interested in music.	When Alex was 14, he broke his arm in my Physical Education class. He was in hospital for a week, and we collected money and bought him a CD. He was probably starting to get interested in music.	L1, L2, L13, S3
Joanna was about 12 when I taught her.	Joanna was about 12 years old when I taught her.	G4
The school was a bit like a prison, and then her mum, Anne, got seriously ill.	The school was similar to prison. Then, J.K.'s mom, Anne, got seriously ill.	L1, S2
So perhaps it's just a coincidence that I used to have long, black hair. But to be honest, I think Professor Snape, especially in the later books, is very like me.	So maybe it's just a coincidence that I used to have long, black hair. But to be honest, I think Professor Snape is very similar to me, especially in the later books.	L1, S1
Jude was bullied at his first secondary school so he changed schools when he was 14 and came to us.	Jude was bullied at his first secondary school, so he changed schools when he was 14 years old and came to us.	S3, G4
He was clever, but some teachers thought he was arrogant. Some of his classmates thought the same and he wasn't very popular with them. Other boys were jealous of him because the girls usually liked him.	Jude was smart, but some teachers thought he was arrogant. Some of his classmates thought he was arrogant too. He wasn't very popular with his classmates.	L1, L2, S2

Fonte: Elaboração própria

SUGESTÕES PARA O TEXTO P19

(Pre-intermediate – páginas 90 e 91)

Quadro 102 – Sugestões de estratégias para o texto P19

ORIGINAL	SUGESTÃO DE REESCRITA	ESTRATÉGIA(S)
One was English, and one was Irish, and they were both on Erasmus scholarships.	One woman was English, and the other was Irish, and they were both on the same scholarships.	G1, G4
Their hair is the same colour and length, they're the same age and size, and when you see them together, in the photo they put on social media, you would think that they were identical twins.	Their hair is the same colour and length and they're the same age and size. When you see them together in the photo they put on social media, you would think that they were identical twins.	S2
In fact, Cordelia Robert and Ciara Murphy are unrelated.	In fact, Cordelia Robert and Ciara Murphy are not related.	L1
Niamh Geaney, from Dublin, and two friends were so interested in trying to find their 'twins' that they set up an online project called Twin Strangers.	Niamh Geaney, from Dublin, and two friends were so interested in trying to find their "twins" that they created an online project called 'Twin Strangers'.	L1, G2
Very quickly, Niamh found a remarkably similar-looking stranger who lived just a few miles away.	Very quickly, Niamh found a stranger who looked a lot like him. This stranger lived just a short distance away.	L1, S2, G5
It's perhaps not so surprising, as both young women look typically Irish, with dark hair and very pale skin, but then Niamh found another lookalike - Luisa Guizzardi, who is from Genoa in Italy!	Maybe it's not so surprising, because both young women look typically Irish. They have dark hair and very pale skin. But then, Niamh found another person who looked like her. Her name was Luisa Guizzardi and she is from Genoa in Italy!	L1, S2, G4

Fonte: Elaboração própria

SUGESTÕES PARA O TEXTO P22

(Pre-intermediate – página 101)

Quadro 103 – Sugestões de estratégias para o texto P22

ORIGINAL	SUGESTÃO DE REESCRITA	ESTRATÉGIA(S)
An Italian couple ran out onto the runway of Malta's international airport to stop a Ryanair jet from leaving for Italy without them, a Maltese court heard on Thursday.	An Italian couple invaded the track of Malta's international airport. They wanted to stop a Ryanair plane from leaving for Italy without them.	L1, S2, G1
Matteo Clementi, 26, and Enrica Apollonio, 23, got stuck in terrible traffic on their way to the airport on Wednesday.	Matteo Clementi and Enrica Apollonio got stuck in terrible traffic on their way to the airport on Wednesday.	G1
They went to the next gate, forced open a security door, and ran towards the plane.	They went to the next gate, open a security door using force, and ran in direction of the plane.	L1, L11

The engines were running and the stairs had been removed, but the couple tried signalling to the pilots to let them get on.	The motor of the plane was on and the stairs had been removed, but the couple tried making a sign to the pilots to let them get on.	L1, L11
A lawyer defending them in court said that Wednesday was Erica's 23rd birthday.	The lawyer who defended Erica and Matteo said that Wednesday was Erica's 23rd birthday.	L2, G1, G4
Instead she spent the evening in prison, and the couple were fined €2,329.	But Erica had to spend the evening in prison, and the couple had to pay a €2,329 fine.	L1, L2, G5
A man who found \$20 in the street near San Francisco International Airport used it to play the California lottery and won \$1 million, spokesman Greg Parashak said on Monday.	A man found \$20 in the street near 'San Francisco International Airport'. He used the money to play the California lottery and won \$1 million.	S2, G1, G2
Hubert Tang used the \$20 to buy two lottery scratch cards at a store near the airport on Wednesday and won the top prize with one of them, Parashak told us.	Hubert Tang used the \$20 to buy two scratch lottery cards at a store near the airport on Wednesday. He won the biggest prize with one of the tickets.	L2, S2 Sugerimos, também, adicionar um glossário com o termo "scratch lottery cards".
Tang said, 'I scratched the ticket outside the store. I told my friend who I was with that I didn't know if it was real but I thought I had just won a million dollars.'	Tang said, "I scratched the ticket outside the store. I told my friend that I didn't know if it was real, but I thought I had just won a million dollars."	S3, G1
Tang, who works as a barman at the airport, had not played the lottery for the last ten years.	Tang works as a barman at the airport. He had not played the lottery for the last ten years.	S2, G4

Fonte: Elaboração própria

Lembramos, como fizemos no início deste capítulo e repetidas vezes, que as sugestões apresentadas anteriormente são, como o nome propõe, sugestões. É sempre preciso ter coerência com as escolhas de estratégias de acessibilidade que iremos adotar levando em consideração o perfil e a necessidade dos nossos estudantes.

No próximo capítulo, apresentamos um manual sucinto de como utilizar as métricas textuais para estimar complexidade e, a partir dos resultados, escolher estratégias de simplificação que se adequem aos propósitos.

10 ORIENTAÇÕES PARA ACESSIBILIDADE TEXTUAL LINGUODIDÁTICA

Neste capítulo da tese, temos por objetivo apresentar um compilado de informações sobre as metodologias e métricas que utilizamos para a pesquisa, de forma didática, a fim de auxiliar pesquisadores e professores que desejem replicar nossos procedimentos. Os processos e informações que descrevermos neste capítulo podem auxiliar docentes de inglês como língua adicional (ILA) a verificar potenciais de complexidade textual dos materiais utilizados em suas aulas (seja para aumentar ou diminuir a acessibilidade de acordo com as necessidades dos seus alunos) e, diante dos resultados, adotar estratégias de acessibilidade coerentes aos indicativos.

Para que possamos apresentar o capítulo como um *booklet* que possa ser utilizado de forma independente da tese (embora nosso conselho seja a leitura da pesquisa em sua íntegra para informações mais detalhadas), algumas informações (como, por exemplo, as interpretações dos valores métricos referentes à complexidade) serão repetidas.

Na primeira parte deste capítulo, instruímos sobre a utilização da plataforma Coh-Matrix, por meio da qual as métricas podem ser coletadas.

Na segunda parte, descrevemos e exemplificamos como utilizar os valores de cada uma das métricas que adotamos em nossa pesquisa com o objetivo de verificar características textuais que possam indicar complexidade. Esses procedimentos podem ser utilizados não só para análise dos dados apresentados por nós dos livros da coleção *English File*, mas também para qualquer outro texto escrito em língua inglesa⁷⁸ voltado para alunos de ILA.

Além das orientações de utilização das métricas, para cada uma delas apresentaremos sugestões de estratégias de acessibilidade que podem ser utilizadas para tornar os textos mais acessíveis. Caso o objetivo seja aumentar a complexidade dos textos (para aumentar o insumo linguístico para alunos avançados, por exemplo), nossas sugestões sobre o que “evitar” podem ser utilizadas de forma contrária. Por exemplo, se para determinada métrica sugerirmos como estratégia a substituição de

⁷⁸ Ainda que a nossa língua de pesquisa seja o inglês, os procedimentos metodológicos podem ser adotados e as métricas que utilizamos também podem ser obtidas para textos em língua portuguesa por meio da versão do Coh-Matrix para o português, que pode ser acessado em <http://nilc.icmc.usp.br/coh-matrix-dementia/>

um pronome pessoal por um substantivo com a facilidade de facilitar a referência para alunos básicos, pode-se fazer o oposto e substituir um substantivo por um pronome para aumentar a complexidade e proporcionar progresso para alunos avançados.

É extremamente importante destacar alguns pontos sobre as estratégias que iremos sugerir:

I. Elas têm por finalidade tornar textos mais acessíveis para alunos em níveis iniciais, ou para alunos com dificuldades, ou para alunos neurotípicos;

II. A adoção das estratégias deve seguir o bom senso do docente em relação aos perfis de seus alunos e, também, ser coerente com os conteúdos que estão sendo trabalhados. Se, por exemplo, o conteúdo da unidade na qual determinado texto se encontra é pronomes relativos, a estratégia sugerida de evitar apostos para melhorar os valores de SS deve ser evitada; caso contrário, estaríamos retirando do texto os insumos para praticar e desenvolver o foco principal daquela unidade;

III. As sugestões não dão conta de resolver todos os possíveis problemas de complexidade textual da língua inglesa (isso nunca será possível, conforme mencionamos no decorrer desta tese, uma vez que complexidade depende de inúmeros fatores);

IV. As orientações não devem ser vistas como reducionistas e devem ser seguidas para casos particulares;

V. Docentes devem ter o cuidado de equilibrar complexidade e insumo de aprendizado. É preciso lembrar que, conforme citamos na tese, é essencial que materiais voltados para o ensino apresentem vocabulários e conteúdos um pouco além daquele já internalizado pelo aluno, caso contrário, a manutenção linguística pode não ocorrer.

Iniciaremos, portanto, com orientações gerais sobre a utilização do Coh-Metrix e sobre a coleta de métricas.

10.1 COMO UTILIZAR O COH-METRIX E COLETAR MÉTRICAS TEXTUAIS

O Coh-Metrix (disponível em <http://cohmetrix.com/>) é uma ferramenta de Processamento de Linguagem Natural que permite a submissão de textos para

coletas de métricas textuais. Essas métricas podem ser utilizadas para diversas finalidades, incluindo, análise de complexidade textual.

Atualmente, a ferramenta disponibiliza 106 métricas.

Para que esses índices possam ser coletados, o primeiro passo é o acesso à ferramenta por intermédio do site, como ilustrado na imagem abaixo (Figura 35):

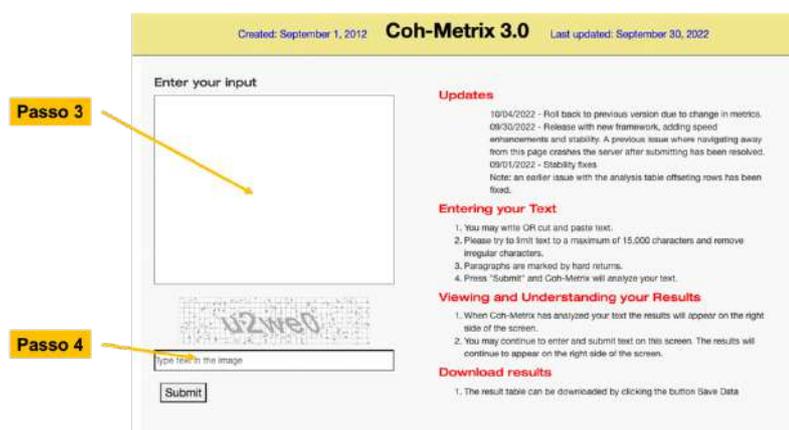
Figura 35 – Passos 1 e 2 para o uso do Coh-Metrix



Fonte: Elaboração própria (com base no site do Coh-Metrix)

Assim que a ferramenta for aberta, podemos colar o nosso texto no campo destinado para essa finalidade (Passo 3) e, em seguida, inserir as letras que aparecem no campo ilustrado pelo passo 4 (*captcha*), conforme mostra a imagem abaixo (Figura 36):

Figura 36 – Passos 3 e 4 para o uso do Coh-Metrix

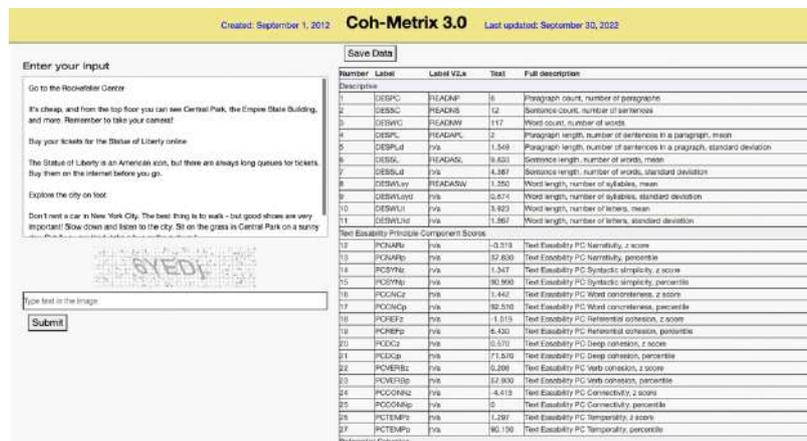


Fonte: Elaboração própria (com base no site do Coh-Metrix)

Assim que o texto e o *captcha* forem inseridos, é necessário clicar em “*submit*”. É comum que o Coh-Metrix demore alguns segundos (por vezes, minutos) para carregar as informações métricas. O tempo de carregamento é proporcional ao tamanho do texto, ou seja, textos mais longos costumam demorar mais tempo para a exibição dos resultados.

Quando a leitura do texto pelo *software* for concluída, ele exibirá as métricas no campo à direita, similarmente aos valores que são exibidos no exemplo da imagem abaixo (Figura 37):

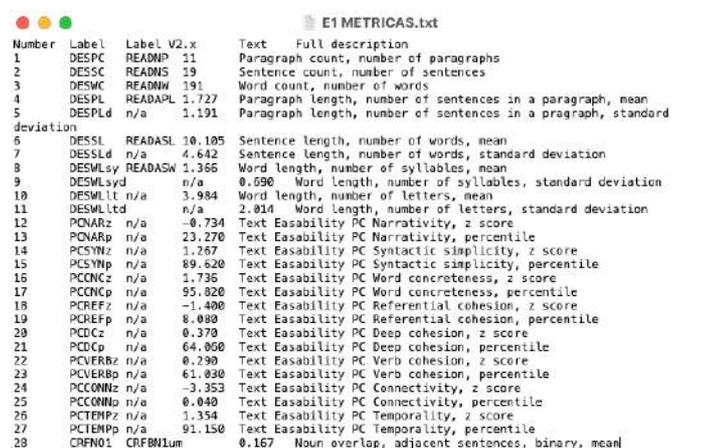
Figura 37 – Ilustração de resultados do Coh-Metrix



Fonte: Coh-Metrix

Se desejar, é possível clicar no botão “*Save Data*” para gerar um arquivo .txt com as métricas para que você salve os resultados no seu computador, como no exemplo da imagem abaixo (Figura 38):

Figura 38 – Ilustração de download de métricas do Coh-Metrix



Fonte: Elaboração própria (com base no site do Coh-Metrix)

A interpretação das métricas como indicativos de complexidade/acessibilidade está relacionada a diversos fatores apontados na condução desta tese. Nos próximos capítulos desse *booklet*, apresentaremos um compilado de instruções e orientações para utilização de algumas das métricas que podem ser obtidas pelo Coh-Metrix para a finalidade de análise de complexidade textual.

10.2 COMO UTILIZAR A NARRATIVIDADE (NT) PARA ESTIMAR COMPLEXIDADE

A métrica Narratividade (NT) é uma métrica percentual e sua relação com acessibilidade é proporcionalmente direta, ou seja, valores mais altos indicam maiores potenciais de complexidade.

A NT, como o nome sugere, indica o quanto um texto é parecido ou diferente de uma história narrativa. Essa tipologia textual é considerada simples por diversos motivos, dentre os quais:

- a. Por ser o primeiro tipo de texto com os quais temos contato desde a primeira infância. Sendo assim, estamos acostumados a “consumir” esse tipo de texto, seja em contos de fada, quadrinhos ou até mesmo desenhos animados;
- b. O desencadeamento dos acontecimentos tende a ser lógico;
- c. As descrições são, geralmente, detalhadas;
- d. As relações construídas no texto são, primordialmente, claras e fáceis de serem acompanhadas;
- e. A linguagem utilizada é simples e direta.

Para que possamos utilizar essa métrica (cujo número no Coh-Metrix na versão disponível no ano 2022 é 13 - *Text Easability PC Narrativity, percentile*) como indício de complexidade, podemos considerar que, conforme mencionado anteriormente, quanto mais altos os valores de NT, mais simples o texto tende a ser.

Por meio da nossa pesquisa proveniente da tese da qual esse *booklet* é parte, os resultados nos permitiram desenvolver o seguinte quadro classificatório (Quadro 104) que organiza os valores de NT de acordo com sua complexidade/acessibilidade:

Quadro 104 – Referências de complexidade de NT

VALORES DE REFERÊNCIA DE NT	
Fácil:	80 a 100
Médio:	40 a 79
Difícil:	0 a 39

Fonte: Elaboração própria

Para exemplificar a interpretação dos valores, considere os exemplos no quadro abaixo (Quadro 105):

Quadro 105 – Exemplos de NT

Exemplo 1: NT: 78,52	<p><i>My dad has depression. Sometimes he feels happy, sometimes he feels really sad.</i></p> <p><i>Depression is a mental disease that happens because the chemicals in our bodies are not regulated.</i></p> <p><i>Sometimes, my dad has episodes of depression. He told me, “Depression can happen any time. Sometimes, I know what makes an episode start. Other times, I don’t know what makes me feel that way.”</i></p> <p><i>Dad also told me that, like himself, people who suffer from depression can look for help. These people can go to a doctor or psychologists. These professionals can help depressed people to find the right treatments and have a normal life.</i></p>
Exemplo 2 (NT: 28,77)	<p><i>Depression is a mental issue related to chemical imbalance. People who suffer from this disease usually have some episodes of depression from time to time.</i></p> <p><i>These episodes may be triggered by a clear cause or not. Patients with depression can lead a normal life, and they can control this disease with the help of doctors and psychologists.</i></p>

Fonte: Elaboração própria

Ao compararmos os valores de NT entre os dois exemplos acima, podemos ver que o exemplo 1 tem um resultado mais alto do que o exemplo 2, que indica, por essa métrica, que o exemplo 1 é mais similar a uma narrativa e, portanto, mais acessível.

Pela leitura do texto, é fácil identificar quais são as características que tornam o exemplo 1 mais acessível e que impactam nos valores de NT. Nesse texto, podemos perceber que uma história está sendo contada, inclusive com diálogos.

Ao compararmos os exemplos 1 e 2, podemos ver que as mesmas informações estão sendo apresentadas, mas de formas diferentes. Ou seja, se o objetivo do texto é informar um leitor sobre a depressão, o exemplo 1 é mais fácil de ser assimilado e “carrega” as mesmas informações que o segundo.

É claro que esses exemplos são muito simples e propositalmente redigidos para “forçar” os resultados de NT, mas as estratégias que utilizamos para “causar” um valor alto no exemplo 1 podem ser adotadas para simplificar outros textos.

Logo, características textuais que influenciam na métrica NT e as estratégias que podem ser adotadas para melhorar a acessibilidade textual relacionada a essa métrica são apresentadas no quadro a seguir (Quadro 106). Destacamos que os exemplos que apresentaremos para determinadas estratégias são curtos e pontuais, servindo apenas como forma de ilustração da aplicação das técnicas de simplificação.

Quadro 106 – Estratégias para aumento de acessibilidade de NT

CARACTERÍSTICAS QUE INFLUENCIAM NA NT	ESTRATÉGIAS PARA AUMENTO DA ACESSIBILIDADE DE NT
1. Extensão de palavras e frases	a) Utilizar palavras mais curtas; b) Dividir frases e parágrafos em dois ou mais (quando necessário, repetir o sujeito ou retomá-lo por meio de um sinônimo ou pronome fácil de relacionar, ou repetir o verbo principal da oração); c) Eliminar informações desnecessárias com a finalidade de reduzir extensões.
2. Encadeamento e relação entre palavras	a) Verificar se a relação entre palavras (como entre substantivos e pronomes, sujeitos e verbos etc.) está clara; b) Aproximar um pronome do seu referencial. Se o pronome que retoma uma determinada palavra estiver muito longe de seu referencial, é preferível repetir a palavra ou utilizar um sinônimo para voltar a mencioná-lo; c) Observar a presença de apostos que possam separar um elemento referenciador (como um pronome, por exemplo) do elemento referenciado.
3. Posicionamento e quantidades de palavras em funções sintáticas	a) Contar a quantidade de palavras que ocupam cada posição sintática. Muitas palavras antes ou depois do verbo principal tendem a aumentar a complexidade dos textos.
4. Pontuações	a) Verificar se as pontuações estão adequadas; b) Evitar uso de pontuações não usuais (como utilizar um aposto entre hifens em vez de vírgulas ou um ponto e vírgula no lugar de ponto).
5. Escolha de palavras	a) Utilizar palavras mais curtas; b) Escolher palavras mais simples (para verificar qual entre dois sinônimos é o mais simples, Silva (2018) indica verificar a frequência dessas palavras em <i>corpus</i> e optar pela unidade que for mais frequente); c) Atentar para o gênero textual e para o público e escolher vocabulários coerentes (evitar utilizar a linguagem acadêmica em textos voltados para o ensino de línguas, por exemplo).

6. Lógica da progressão textual	<p>a) Verificar se a apresentação dos fatos segue uma estrutura lógica e de fácil construção mental durante a leitura;</p> <p>b) Aumentar a frequência de verbos. Em frases muito extensas e com mais de um objeto (função sintática) para o mesmo verbo, utilize sinônimos. Por exemplo: em vez de “<i>He bought a car, a house, and a new phone.</i>”, divida essa oração e utilize mais verbos com a mesma ideia ou repita o mesmo verbo (“<i>He bought a new car. He purchased a new house. He acquired a new phone.</i>”). Lembre-se, contudo, que a escolha desses verbos pode aumentar a complexidade lexical. Certifique-se de que os verbos escolhidos para representar a mesma ideia sejam parte do vocabulário dos alunos ou de simples inferência. Caso contrário, opte pela repetição do mesmo verbo;</p> <p>c) Questionar se o leitor precisará sair do texto em algum momento para pesquisar o significado de conceitos específicos (a saída do texto é destacada como fator de distração e de interrupção da construção do significado por meio da leitura).</p>
7. Remissividade	<p>a) Analisar se o texto dá “voltas”, ou seja, se as informações se repetem e se a apresentação dos fatos não está misturada;</p> <p>b) Organizar os textos em tópicos (especialmente textos informativos que não sejam narrativos);</p>
8. Coordenação e subordinação de orações	<p>a) Verificar se há muitas ocorrências de coordenação e ordenação de orações. No caso de alta frequência, adotar a estratégia de quebra de frases e parágrafos em estruturas menores, mesmo que seja necessário repetir palavras;</p> <p>b) Optar por conjunções e advérbios mais simples (como ‘<i>but</i>’ no lugar de ‘<i>however</i>’);</p> <p>c) Observar se a posição dos elementos sintáticos segue a ordem canônica e optar por formação canônica (sujeito, verbo, objeto/complemento), evitando deslocamentos.</p>
9. Personificação	<p>a) Quando os sujeitos do texto não forem claros, personificá-los por meio de palavras que sejam concretas (como ‘<i>people</i>’, ‘<i>students</i>’, ‘<i>the girl</i>’ etc.)</p>
10. Voz ativa/passiva	<p>a) Optar pela utilização da voz ativa, mesmo que seja necessário transformar o agente da passiva em um sujeito que é óbvio (como, por exemplo, “<i>A thief robbed the bank.</i>” no lugar de “<i>The bank was robbed.</i>”</p>

Fonte: Elaboração própria

As informações presentes no quadro acima nos permitem adequar textos com a finalidade de aumentar os valores de NT e, conseqüentemente, a acessibilidade de leitura. Entretanto, é válido observar que a NT aborda apenas alguns traços e níveis textuais que podem ser fatores de altos potenciais de complexidade. Com o intuito de realizar análises mais completas, sugerimos a utilização de múltiplas métricas, como apresentaremos nas próximas subseções.

10.3 COMO UTILIZAR A SIMPLICIDADE SINTÁTICA (SS) PARA ESTIMAR COMPLEXIDADE

A métrica Simplicidade Sintática (SS) está relacionada, conforme o nome sugere, às distintas formações sintáticas de todas as frases de um texto.

Esse índice, portanto, observa por meio de um processo de análise sintática, a quantidade de formações distintas possíveis em um texto. Ao relacionarmos essa característica a potenciais de complexidade textual, podemos afirmar que textos que têm muitas frases com formações sintáticas distintas tendem a ser mais complexos.

Imagine que você desconhece totalmente a língua inglesa e tenha começado a estudá-la. Seu vocabulário é limitado e seus conhecimentos sobre as posições das palavras, também.

Ao passo que você estuda, você percebe, por exemplo, que a formação mais comum das frases nessa língua segue a ordem 'sujeito, verbo, objeto' (mesmo que você nem precise conhecer a nomenclatura sintática). Contudo, por saber que línguas não têm uma natureza rígida e que, ao contrário, seus dinamismos, geralmente, permitem modificações, você começa a perceber que existem outras formas de estruturar frases, e que, além disso, funções sintáticas como sujeito, por exemplo, podem ter várias palavras, ou até mesmo uma frase, na sua posição.

Ainda que você tenha tido tal percepção, seu vocabulário continua limitado. Então, faz de conta, que você tenha percebido e estudado que a palavra "*reading*" costuma ser um verbo no gerúndio e, usualmente, é encontrada em exemplos como "*She is reading a book.*"

Depois de um determinado tempo, você está lendo um texto no seu livro didático de inglês e se depara com a seguinte frase: "*Reading is important.*" O deslocamento da palavra "*reading*" para frente do verbo "*is*" pode ser um fator que dificultará ou impedirá sua interpretação correta da frase, uma vez que a palavra "*reading*" não está posicionada no local onde você costumava percebê-la.

Mais adiante no texto, você encontra a frase: "*Reading fiction books every day is, according to teachers, a great way of improving one's vocabulary.*" Agora, além da posição da palavra "*reading*" ser diferente do habitual, a quantidade de palavras depois dela é muito maior do que os exemplos que você costumava ver, potencializando, ainda mais, a complexidade do texto.

Esse exemplo, embora simples, pode ilustrar algumas características a respeito de sintaxe que podem contribuir para que um texto se torne mais simples ou mais difícil.

Por ser uma métrica que analisa a formação sintática de cada frase do texto e verifica se elas são formadas de maneiras mais simples ou complexas, a SS pode ajudar a estimar a complexidade de um texto. Nossos estudos nos auxiliaram a desenvolver o quadro abaixo (Quadro 107) que pode ser utilizado como referência de interpretação dos valores dessa métrica em relação à potenciais de complexidade:

Quadro 107 – Referências de complexidade de SS

VALORES DE REFERÊNCIA DE SS	
Fácil:	75 a 100
Médio:	50 a 74
Difícil:	0 a 49

Fonte: Elaboração própria

No quadro abaixo (Quadro 108) apresentamos algumas ocorrências que impactam na métrica SS e estratégias que podem ser adotadas para simplificar textos destacados por essa métrica como complexos:

Quadro 108 – Estratégias para aumento de acessibilidade de SS

CARACTERÍSTICAS QUE INFLUENCIAM NA SS	ESTRATÉGIAS PARA AUMENTO DA ACESSIBILIDADE DE SS
1. Deslocamento de elementos	a) Optar pela ordem canônica (sujeito verbo objeto); b) Atentar para a posição de adjuntos, evitando o deslocamento para posições não-canônicas.
2. Presença de orações coordenadas e subordinadas	a) Verificar a quantidade de orações coordenadas e subordinadas; b) Dividir orações compostas em duas ou mais, quando possível (se for preciso, repetindo sujeitos ou verbos); c) Atentar para a complexidade semântica de conjunções e advérbios;
3. Voz ativa/passiva	a) Dar preferência para a voz ativa (além da estrutura sintática da voz passiva ser diferente, em língua inglesa esse tipo de voz é formado por verbos no particípio que, pelo fato de poderem ser regulares ou irregulares, também aumentam a complexidade do texto por características lexicais e semânticas.
4. Apostos	a) Evitar apostos e optar por apresentar as informações por meio de uma nova frase. Por exemplo, em vez de “ <i>My sister, who studies in Paris, speaks French really well.</i> ”, optar por: “ <i>My sister studies in Paris. She speaks French really well.</i> ”

Fonte: Elaboração própria

Na próxima seção, apresentaremos instruções sobre a métrica Similaridade Sintática (SIMS) que está fortemente relacionada com as orientações de SS.

10.4 COMO UTILIZAR A SIMILARIDADE SINTÁTICA (SIMS) PARA ESTIMAR COMPLEXIDADE

A similaridade sintática (SIMS) é um índice (e uma característica textual) que também influencia na acessibilidade do texto. Essa métrica é altamente relacionada à anterior (SS) e muitas das indicações a respeito da SS também se aplicam a ela.

A SIMS analisa a formação sintática de todas as frases de um texto e observa o quanto as frases seguem ordens parecidas ou distintas. McNamara et al. (2014) e outros pesquisadores de acessibilidade textual indicam que textos com alta frequência de frases que seguem formações sintáticas parecidas tendem a ser mais acessíveis.

Evidentemente, precisamos lembrar que ao lidarmos com textos voltados ao ensino, a presença de frases com deslocamentos de elementos, com coordenação e subordinação e com diferentes configurações sintáticas pode auxiliar e servir de insumo linguístico, mas, ao mesmo tempo, com cuidado. Conforme já mencionamos, a forma e o momento no qual alguns aspectos linguísticos são apresentados para os alunos podem ser fatores que irão auxiliar ou prejudicar a manutenção linguística.

Obviamente, mesmo para um aluno de níveis iniciais, textos que sigam somente a ordem canônica de formações de frases não potencializarão a manutenção linguística. Entretanto, estruturas sintaticamente complexas também não.

Uma vez que a sintaxe é caracterizada pela relação entre os elementos de uma frase/oração, não só a posição das palavras, mas também a quantidade de palavras em cada relação, influenciam na interpretação (como no exemplo da palavra “*reading*”, apresentado na seção anterior).

Nossa pesquisa nos auxiliou a elaborar o quadro abaixo (Quadro 109) que classifica os valores de SIMS de acordo com seu potencial de complexidade. Esse quadro pode ser utilizado para fins interpretativos dessa métrica:

Quadro 109 – Referências de complexidade de SIMS

VALORES DE REFERÊNCIA DE SIMS	
Fácil:	Acima de 101
Médio:	80 a 100

Difícil:	Abaixo de 80
----------	--------------

Fonte: Elaboração própria

Além das indicações sobre as características que influenciam a SS (seção anterior, 10.3), para obter melhores resultados de SIMS, devemos atentar para a estruturação sintática de cada frase do texto. Sendo assim, somando-se às informações apresentadas na seção anterior, apresentamos três estratégias para otimizar o índice SIMS, conforme o quadro abaixo (Quadro 110):

Quadro 110 – Estratégias para aumento de acessibilidade de SIMS

CARACTERÍSTICAS QUE INFLUENCIAM NA SIMS	ESTRATÉGIAS PARA AUMENTO DA ACESSIBILIDADE DE SIMS
1. Forma de deslocamento de elementos e de apresentação de subordinação e coordenação	<p>a) Ao deslocar elementos, dar preferência por seguir a mesma ordem em mais frases do texto. Se, por exemplo, você deslocar uma expressão de tempo para o início da frase, opte por fazer o mesmo tipo de deslocamento no decorrer do texto para esse tipo de expressão.</p> <p>Ao passo que o aluno avança de nível linguístico, apresente outros tipos de deslocação.</p> <p>b) Apresentar orações coordenadas e subordinadas de forma paralela no decorrer do mesmo texto, preferencialmente repetindo as conjunções e advérbios para comunicar a mesma ideia (como, por exemplo, utilizar <i>'but'</i> em vez de <i>'however'</i> em todas as instâncias de coordenação por oposição).</p>
2. Quantidade de palavras em cada função sintática	<p>a) Tentar padronizar a quantidade de palavras em cada função sintática. Evite que em uma frase o número de palavras que ocupa a posição de sujeito, por exemplo, seja muito superior à quantidade de palavras que exercem a mesma função em outras frases.</p>

Fonte: Elaboração própria

É necessário voltar a enfatizar que as estratégias apresentadas pelo quadro acima (bem como todas as demais) são indicações para serem adotadas com a finalidade de aumentar a acessibilidade de um texto em casos particulares. A aplicação dessas estratégias deve ser coerente com a finalidade didática de cada texto e deve ter o cuidado de não eliminar insumos linguísticos que sejam necessários ou propositalmente apresentados.

A seguir, apresentamos orientações para o uso da métrica coesão referencial para estimar complexidade textual.

10.5 COMO UTILIZAR A COESÃO REFERENCIAL (CR) PARA ESTIMAR COMPLEXIDADE

A métrica coesão referencial (CR) observa as relações estabelecidas no texto. De forma resumida, esse índice resulta em um número percentual que indica o quão fácil (valores mais altos) ou difíceis (valores mais baixos) as relações entre palavras no texto são feitas. Essa relação pode ser percebida, especialmente, entre pronomes e substantivos, mas também na conexão entre sujeito e predicado.

Observe os exemplos abaixo:

Exemplo 1: My friends like parties, which are always incredible.

Exemplo 2: My friends like parties. They are always incredible.

Exemplo 3: My friends like parties. The parties are always incredible.

Exemplo 4: My friends like parties. These events are always incredible.

No exemplo 1, podemos compreender, por meio do pronome relativo *'which'* que o adjetivo *'incredible'* caracteriza as festas. Entretanto, não podemos supor que um aluno em níveis iniciais (ou um aluno neuroatípico) também faria essa relação de forma fácil, especialmente ao observarmos que pronomes relativos são ensinados apenas no nível *Pre-intermediate*.

No exemplo 2, por outro lado, o pronome *'they'* causa ambiguidade. Quem são incríveis? As festas ou os amigos? Logo, a coesão referencial entre o pronome e o substantivo a qual ele se refere é baixa.

Já no exemplo 3, está extremamente claro a que o adjetivo se refere; visto que não há um pronome retomando um substantivo, mas a repetição do mesmo substantivo que garante a coesão referencial. Entre os quatro exemplos, o terceiro é o mais acessível; por outro lado, nele não há muito insumo para a manutenção linguística (como no caso 1, que pode ser utilizado para ensinar pronomes relativos ou no exemplo 4 que pode ser uma maneira de trabalhar com sinônimos).

Por fim, no exemplo 4 temos duas frases nas quais o uso do sinônimo *'events'* e do pronome demonstrativo *'these'* fazem referência ao substantivo *'parties'* e, ao mesmo tempo, garantem a relação do adjetivo *'incredible'* ao referente certo. Além da retomada de *'parties'* feita por *'these events'*, a adjetivação é mais simplesmente entendida por não termos um pronome que gere ambiguidade e pela proximidade do

substantivo ao adjetivo. Logo, o exemplo 4 tem alta coesão referencial, mas exige maiores conhecimentos linguísticos dos leitores (tanto lexical, para compreender que ‘events’ está sendo usado como sinônimo de ‘parties’, quanto referencial, na compreensão de que ‘these’ se refere simultaneamente às palavras ‘events’ e ‘parties’).

Por meio dos exemplos supracitados, ainda que simples, podemos compreender a importância de analisar coesão referencial nos textos voltados ao ensino, especialmente em textos redigidos em uma língua adicional.

Nossa análise nos auxiliou a desenvolver um quadro de referência dos valores de CR que pode ser utilizado na interpretação de complexidade textual. Esses valores são exibidos pelo quadro abaixo (Quadro 111):

Quadro 111 – Referências de complexidade de CR

VALORES DE REFERÊNCIA DE CR	
Fácil:	50 a 100
Médio:	31 a 49
Difícil:	0 a 30

Fonte: Elaboração própria

Agora que já compreendemos o que significa a CR e os motivos pelos quais ela deve ser levada em consideração em análises de potencial de complexidade textual, o quadro abaixo (Quadro 112) nos apresenta características que influenciam nos valores dessa métrica e estratégias que podem ser utilizadas para aumentar a acessibilidade relativa à coesão referencial:

Quadro 112 – Estratégias para aumento de acessibilidade de CR

CARACTERÍSTICAS QUE INFLUENCIAM NA CR	ESTRATÉGIAS PARA AUMENTO DA ACESSIBILIDADE DE CR
1. Quantidade de pronomes	<p>a) Verificar se a quantidade de pronomes (não-dêiticos) é grande. Muitos pronomes tendem a confundir leitores em níveis iniciais em relação aos referentes, especialmente se houver mais de um substantivo com gênero ou número que também poderia ser referido pelo mesmo pronome. Como, por exemplo na frase “<i>He loves his dad. He bought him a gift.</i>”, não é claro a quem os pronomes ‘he’ e ‘him’ se referem;</p> <p>b) Optar por uso de sinônimos para retomar conceitos ou repetir as palavras, especialmente em textos com muitos pronomes.</p>

2. Quantidade de verbos	a) Analisar a quantidade de verbos. Maiores presenças de verbos, especialmente próximos aos sujeitos, tende a aumentar a coesão, uma vez que os verbos devem concordar com os sujeitos.
3. Distância de pronomes dos seus referentes	a) Observar se um pronome não está muito distante de seu referente; caso estiver, adotar procedimento b) da característica 1 deste quadro.
4. Extensão textual	a) Textos mais extensos tendem a apresentar maiores quantidades de palavras para retomar outras já citadas. Textos menores (incluindo frases e parágrafos menores) reduzem essa necessidade.

Fonte: Elaboração própria

Voltamos a destacar que essas orientações fazem parte de um conjunto de outras instruções que podem auxiliar no aumento da acessibilidade textual, não sendo as únicas indicadas nem para diminuir a complexidade de um texto, nem como únicas que podem ser adotadas para melhorar os valores de coesão referencial, mas como as principais estratégias que podem aprimorar os valores dessa métrica.

Na próxima seção, apresentamos instruções para o uso de IF para pesquisas de acessibilidade textual.

10.6 COMO UTILIZAR O ÍNDICE FLESCH (IF) PARA ESTIMAR COMPLEXIDADE

O Índice Flesch (IF) é uma métrica consagrada em estudos de acessibilidade textual. Entretanto, visto que esse índice estima complexidade somente por extensão de palavras e de frases, ele deve ser utilizado com outros indícios de complexidade para que as análises de complexidade sejam abrangentes e efetivas.

Dentre as métricas apresentadas nesta tese, o IF é a única que foi especialmente desenvolvida para analisar complexidade. Todas as demais são índices textuais que, entre outras finalidades, podem auxiliar nos estudos de acessibilidade. No quadro abaixo (Quadro 113) estão as faixas de IF e suas interpretações relativas à complexidade textual:

Quadro 113 - Resultados do Índice Flesch

ÍNDICE	CLASSIFICAÇÃO
100-90	Muito fácil de ler. Facilmente compreendido por um aluno médio de 11 anos de idade.
90-80	Fácil de ler. Inglês conversacional.

80-70	Razoavelmente fácil de ler.
70-60	<i>Plain English</i> . Facilmente compreendido por alunos de 13 a 15 anos de idade.
60-50	Razoavelmente difícil de ler.
50-30	Difícil de ler
30-10	Muito difícil de ler. Melhor compreendido por estudantes de graduação.
10-0	Extremamente difícil de ler. Melhor compreendido por graduados e pós-graduados.

Fonte: Goldim (2003)

Conforme já mencionado, os parâmetros utilizados para o cálculo do IF são referentes à extensão das palavras e das frases. Por isso, as características que influenciam nos valores dessa métrica são intrinsicamente relacionadas à extensão, conforme explicamos no quadro abaixo (Quadro 114):

Quadro 114 – Estratégias para aumento de acessibilidade de IF

CARACTERÍSTICAS QUE INFLUENCIAM NO IF	ESTRATÉGIAS PARA AUMENTO DA ACESSIBILIDADE DE IF
1. Tamanho das palavras	a) Substituir palavras extensas por sinônimos mais curtos. Por exemplo, substituir o adjetivo da frase " <i>It's amazing.</i> " por outro com significado semelhante (como ' <i>great</i> ').
2. Tamanho das frases	a) Reduzir orações extensas em menores, mesmo que isso exija repetição de palavras. Em vez de: " <i>They look after their younger siblings and take care of the house.</i> ", optar por " <i>They look after their younger siblings. They also take care of the house.</i> " Mesmo que uma nova palavra tenha sido adicionada e outra repetida, o tamanho das frases foi reduzido, uma vez que dividimos uma oração com coordenação em duas outras orações.

Fonte: Elaboração própria

É importante lembrar que as estratégias descritas acima devem ser utilizadas quando buscamos reduzir a complexidade de um texto que sejam relativas à extensão. Conforme já destacamos repetidas vezes, a adoção de estratégias deve ser coerente com as finalidades e, preferencialmente, abranger outros indícios de complexidade.

Na próxima seção, apresentamos informações sobre como a métrica polissemia pode ser utilizada em estudos de complexidade.

10.7 COMO UTILIZAR A POLISSEMIA (PS) PARA ESTIMAR COMPLEXIDADE

A métrica polissemia (PS), como o nome sugere, observa a quantidade de palavras polissêmicas que estão presentes no texto em relação ao número total de palavras. Quanto mais palavras polissêmicas constarem no texto, maior pode ser o potencial de complexidade indicado pela métrica.

No quadro abaixo (Quadro 115), apresentamos uma organização dos valores de PS que auxiliam a interpretar potenciais de complexidade:

Quadro 115 – Referências de complexidade de PS

VALORES DE REFERÊNCIA DE PS	
Fácil:	Acima de 4,05
Médio:	3,49 a 4,04
Difícil:	Abaixo de 3,50

Fonte: Elaboração própria

Por ser uma métrica que analisa uma característica léxico-semântica apenas, poucas são as indicações de traços que influenciam no valor da PS, conforme demonstramos no quadro abaixo (Quadro 116):

Quadro 116 – Estratégias para aumento de acessibilidade de PS

CARACTERÍSTICAS QUE INFLUENCIAM NA PS	ESTRATÉGIAS PARA AUMENTO DA ACESSIBILIDADE DE PS
1. Quantidade de palavras polissêmicas	a) Verificar a quantidade de palavras que podem ter mais de um sentido. Quanto mais palavras polissêmicas estiverem presentes no texto, maior pode ser o potencial de complexidade.
2. Contexto imediato	b) Observar se o contexto imediato apresenta sugestões suficientes para a evocação correta do significado. Por exemplo, se a palavra <i>'tip'</i> estiver dentro de um contexto de restaurantes, por exemplo, é mais fácil compreendê-la como <i>'gorjeta'</i> na frase <i>"She left the tip on the table."</i> do que o outro significado da palavra (ponta).
3. Posição das palavras polissêmicas	a) Analisar a posição de palavras polissêmicas. Algumas palavras polissêmicas são conhecidas por alunos mesmo em níveis iniciais. <i>'Like'</i> , por exemplo, é uma palavra que estudantes brasileiros conhecem tanto o significado "gostar" quanto "como/tipo". Logo, em determinadas instâncias, a posição da palavra pode auxiliar na compreensão. No exemplo: <i>"Like me."</i> estudantes podem compreender a palavra <i>'like'</i> como parte de um imperativo (Goste de mim!) ou como um comparativo (Assim como eu). A posição da palavra poderá ajudar a ativar o sentido correto. (<i>I told</i>

	<i>him, "Like me!", I'm a perfect match for you!" ou "I love reading and, like me, he loves it too.").</i>
--	--

Fonte: Elaboração própria

Ainda que palavras polissêmicas possam ser fatores que potencializam a complexidade de textos, observamos, durante a nossa pesquisa e pela nossa experiência em sala de aula, que em língua adicional, polissemia pode se tornar uma característica de alta complexidade para alunos com vocabulários mais amplos.

Uma vez que, para compreender múltiplos significados de uma palavra é preciso certo domínio lexical, a possível ativação de significados não pretendidos de palavras polissêmicas é quase rara para alunos em níveis iniciais. O fator polissemia, então, passa a ser considerado como traço de potencial complexidade à medida que os estudantes desenvolvem seu vocabulário.

Por outro lado, quando estudantes têm maior domínio lexical, também – geralmente – têm maiores habilidades linguísticas que os auxiliarão a identificar os sentidos corretos propostos pelas palavras que podem ser polissêmicas reduzindo, assim, a possível dificuldade de interpretação.

Ainda que estejamos afirmando que a polissemia é uma característica com pouco potencial de complexidade para alunos de uma língua adicional, não estamos dizendo que ela não seja indicativa de complexidade. Evidentemente, palavras com múltiplo significado podem causar ambiguidade em determinados contextos, mas cabe aos docentes observar esses tipos de palavras nos textos e prever possíveis desvios de significado. Palavras como *'match'*, *'like'*, *'fish'*, *'mean'* e *'book'*, por exemplo, são unidades lexicais polissêmicas comumente conhecidas até por alunos de níveis iniciais e que, dependendo do contexto, podem causar ambiguidade. Já exemplos como *'rank'*, *'lead'*, *'dish'*, *'frame'* e *'face'* são tipos de palavras polissêmicas cuja significação não tão comum (fétido, chumbo, fazer fofoca, armar [para alguém] e enfrentar) é, geralmente, aprendida em níveis linguísticos mais avançados e que raramente irão causar ambiguidade em textos para alunos com limitações linguísticas ou em níveis iniciais (uma vez que seria necessário que não só conhecessem os outros significados, mas também não fossem capazes de utilizar pistas contextuais para ativar o conceito correto).

Sendo assim, orientamos a utilizar a métrica PS apenas como uma forma de sinalização. Ou seja, ao observarmos valores de PS que possam indicar potencial de complexidade, analisar detalhadamente as possíveis palavras polissêmicas para,

então, adotar estratégias que sejam adequadas (conforme sugerimos no quadro acima).

Na próxima seção, orientamos sobre a métrica IA e como utilizá-la para estimar complexidade.

10.8 COMO UTILIZAR A IDADE DE AQUISIÇÃO (IA) PARA ESTIMAR COMPLEXIDADE

A métrica Idade de Aquisição (IA) é baseada em uma lista de palavras desenvolvida por Gilhooly e Logie (1980) que classifica as unidades lexicais de acordo com a idade em que cada uma delas é, geralmente e neurotipicamente, aprendida por falantes nativos da língua inglesa. Sabe-se que, por exemplo, palavras como ‘*mother*’, ‘*water*’ e ‘*hungry*’ são aprendidas na primeira infância, enquanto outras como ‘*homework*’, ‘*chores*’ e ‘*recess*’ costumam ser “adquiridas” em início de idade escolar. Sendo assim, existe uma certa sequenciação lógica e natural no desenvolvimento de vocabulário desde o nascimento.

Com base nessa ideia, a métrica IA do Coh-Metrix utiliza a lista de Gilhooly e Logie (1980) e “pontua” as palavras de acordo com a idade em que elas são aprendidas. A premissa é clara: quanto mais cedo na infância uma palavra é aprendida por falantes nativos, maior tende a ser a acessibilidade dessa palavra. Ou seja, textos ricos em palavras que até crianças pequenas compreendem são, via de regra, mais fáceis de serem entendidos.

Por meio da nossa pesquisa, conseguimos mapear o comportamento de IA e, a partir desse mapeamento, desenvolver o quadro abaixo (Quadro 117) que organiza os valores da métrica de acordo com seus indicativos de complexidade:

Quadro 117 – Referências de complexidade de IA

VALORES DE REFERÊNCIA DE IA	
Fácil:	Abaixo de 275
Médio:	274 a 349
Difícil:	Acima de 350

Fonte: Elaboração própria

Como é possível observar pelo quadro acima, a relação entre os valores de IA e potenciais de complexidade é inversa; ou seja, valores mais altos indicam maior complexidade.

Visto que essa métrica é estritamente relacionada ao léxico, as características que influenciam em seus valores são fáceis de inferir, conforme apresentamos no quadro abaixo (Quadro 118):

Quadro 118 – Estratégias para aumento de acessibilidade de IA

CARACTERÍSTICAS QUE INFLUENCIAM NA IA	ESTRATÉGIAS PARA AUMENTO DA ACESSIBILIDADE DE IA
1. Tipos de palavras apresentadas	<p>a) Observar as palavras apresentadas tendo como “guia” a seguinte pergunta: “Com quantos anos um falante nativo começa a utilizar essa palavra?”</p> <p>Obviamente, a resposta não é definitiva, mas nos ajuda a orientar nossas decisões. Se a resposta para a pergunta for: “essa palavra é geralmente vista pela primeira vez no ensino médio”, pode ser que seja um indicativo de que ela é complexa para um aprendiz de inglês como língua adicional. Por outro lado, se a resposta para nossa pergunta [mental] for “Possivelmente, uma criança utiliza essa palavra diariamente.”, pode ser indício de acessibilidade.</p> <p>Caso não tivermos certeza da complexidade da palavra, podemos utilizar uma estratégia sugerida por Silva (2018) com linguística de <i>corpus</i>.</p> <p>Para verificar se uma palavra é mais comum que outra em determinada língua, é possível acessar um <i>corpus</i> de referência desse idioma (para o inglês, recomendamos o COCA⁷⁹) e pesquisar a frequência de palavras sinônimas. A palavra que for mais frequente é a mais utilizada naquela língua e, conseqüentemente, a mais acessível.</p> <p>Se compararmos a palavra ‘<i>bright</i>’ e a palavra ‘<i>dazzling</i>’ no COCA, por exemplo, vamos descobrir que a primeira tem uma frequência de 50.779, enquanto a segunda de 3.547. Com base nesses números, podemos concluir que ‘<i>bright</i>’ é mais utilizada em língua inglesa do que ‘<i>dazzling</i>’ e, conseqüentemente, seu potencial de complexidade é menor.⁸⁰</p>
2. Cognatos ou semelhantes	<p>a) Considerar a substituição de palavras potencialmente complexas por outras que sejam cognatas verdadeiras ou ortograficamente semelhantes à língua-materna do estudante.</p>

⁷⁹ <https://www.english-corpora.org/coca/>

⁸⁰ Como linguistas, temos a consciência de que não existem sinônimos perfeitos. Entretanto, essa estratégia pode auxiliar na verificação de frequência entre palavras que a professora considere como sinônimas para adequar o léxico do texto com vistas a potencializar a acessibilidade.

	<p>Se, por exemplo, a frase a seguir estivesse em um texto escrito para níveis iniciais, seria possível alterar o adjetivo por <i>'essential'</i>, <i>'important'</i>, <i>'necessary'</i>, <i>'crucial'</i> ou <i>'imperative'</i>, uma vez que todas são cognatas verdadeiras à língua portuguesa.</p> <p><i>"It's paramount that students read this chapter."</i></p>
--	---

Fonte: Elaboração própria

Portanto, a métrica IA apresenta melhores resultados quando o texto é redigido com palavras mais simples. De forma bastante redutiva, podemos explicar a métrica pensando que “quanto mais idade for preciso pra entender uma palavra, mais difícil ela será”. Sendo assim, se tivermos por objetivo aumentar a acessibilidade de um texto por vias lexicais, podemos adotar as estratégias sugeridas no quadro acima. É preciso lembrar, contudo, que essas orientações devem ser adotadas em casos particulares e com moderação para que as modificações não interfiram negativamente na manutenção linguística.

A seguir, apresentamos orientações sobre outra métrica relacionada a léxico que pode ser observada como complemento da IA.

10.9 COMO UTILIZAR A FAMILIARIDADE COM PALAVRAS DE CONTEÚDO (FPC) PARA ESTIMAR COMPLEXIDADE

Como sugerido pelo nome, a métrica Familiaridade com Palavras de Conteúdo (FPC) pode ser utilizada para estimar complexidade porque ela verifica a presença de substantivos, adjetivos, advérbios e verbos que sejam possivelmente familiares para o leitor e de unidades que podem não ser; havendo uma quantidade maior de palavras possivelmente conhecidas, pode-se considerar o texto potencialmente mais simples.

Se, por exemplo, um texto for rico em palavras como *'mother'*, *'food'*, *'school'*, *'eat'*, *'sleep'*, *'hungry'* e tiver poucas palavras menos conhecidas (como *'anthropology'*, *'lexicon'*, *'overthrow'* etc.), ele será possivelmente acessível mesmo para leitores com maiores limitações linguísticas. Quanto mais palavras de conteúdo são possivelmente desconhecidas, mais complexo um texto pode ser considerado.

Nosso estudo nos permitiu observar o comportamento da FPC e organizar os valores da métrica para fins de interpretação de complexidade, conforme o quadro abaixo (Quadro 119):

Quadro 119 – Referências de complexidade de FPC

VALORES DE REFERÊNCIA DE FPC	
Fácil:	Acima de 580
Médio:	564 a 579
Difícil:	Abaixo de 565

Fonte: Elaboração própria

Visto que essa métrica é diretamente relacionada ao léxico, as indicações já apresentadas para o índice IA se aplicam também para FPC. Além delas, os procedimentos descritos no quadro a seguir (Quadro 120) podem aumentar a acessibilidade dos textos:

Quadro 120 – Estratégias para aumento de acessibilidade de FPC

CARACTERÍSTICAS QUE INFLUENCIAM NA FPC	ESTRATÉGIAS PARA AUMENTO DA ACESSIBILIDADE DE FPC
1. Tipos de palavras apresentadas	<p>a) Observar especificamente os substantivos, adjetivos, advérbios e verbos. Diante de palavras dessas classes que forem consideradas potencialmente menos conhecidas, optar por um sinônimo mais frequente. A estratégia de verificação em <i>corpus</i> descrita nas orientações da métrica IA também se aplica aqui;</p> <p>b) Verificar as palavras de conteúdo que têm potencial de ser pouco conhecidas. Caso não seja possível substituí-las por sinônimos mais simples, uma das seguintes estratégias pode ser adotada:</p> <p>b.1) Apresentação da palavra em um glossário; b.2) Tradução da palavra em um glossário; b.3) Adição de uma explicação em forma de aposto (que pode se tornar um fator que aumenta a complexidade caso seja muito longo e redigido com palavras complexas); b.4) Utilização de imagem que ilustre a palavra (quando aplicável); b.5) Eliminação da informação se não for prejudicial para a globalidade do texto.</p>
2. Cognatos	a) Verificar indicações sobre cognatos no Quadro 118.

Fonte: Elaboração própria

É válido lembrar que a adoção dessas estratégias deve ser coerente com as finalidades pedagógicas e que a docente deve cuidar para não eliminar do texto palavras que façam parte do vocabulário de controle da unidade. Além disso, ao adotar estratégias como adição de aposto explicativo, precisa-se ter o cuidado com a extensão do aposto e com os potenciais de complexidade das palavras escolhidas.

Na próxima seção, apresentamos a métrica Concretude Lexical que também auxilia na verificação do potencial de complexidade do léxico.

10.10 COMO UTILIZAR A CONCRETUDE LEXICAL (CLP) PARA ESTIMAR COMPLEXIDADE

A métrica Concretude Lexical (CLP) se refere à possibilidade (ou dificuldade) de relacionar as palavras do texto com referências concretas no mundo real. Ao ler a palavra ‘cadeira’, por exemplo, uma leitora consegue “construir” uma imagem mental de algo que é concreto no mundo real e imaginar um móvel feito para sentar-se (ou seja, o significante não evoca um significado abstrato). Por outro lado, ao ler a palavra “serenidade”, essa mesma leitora não conseguirá evocar um significado concreto; ela até pode relacionar esse substantivo a um momento ou local, mas o significado da palavra não é algo necessariamente tangível.

Na relação com complexidade textual, a CLP indica que quanto mais palavras concretas estiverem presentes no texto, mais acessível ele potencialmente será. Entretanto, são raríssimos (quase inexistentes) os textos que são redigidos somente com palavras concretas, especialmente porque muitos substantivos do cotidiano são abstratos (como ‘honestidade’, ‘velhice’, ‘saudades’ etc.) e fazem parte do léxico comum.

Ainda sobre abstração, há um conceito proposto pelo linguista Hayakawa (1939) chamado “escada de abstração”, que explica que a maior parte das palavras é categorizada em um espectro entre concreto e abstrato dependendo do contexto e da perspectiva assumida. Sendo assim, o potencial de complexidade da não-concretude de algumas palavras em um texto não está somente relacionado ao léxico em si, mas à capacidade de evocação de significados das palavras abstratas. Em materiais escritos para aprendizes de alguma língua adicional, é preciso verificar não só a complexidade das palavras (por meio de sua morfologia, extensão ou frequência, por exemplo), mas se o conceito/ideia/significado que elas representam é compreensível para o leitor. Se tomarmos as palavras ‘love’, ‘tough’ e ‘wistfulness’ como exemplos, podemos ilustrar seus potenciais de complexidade de forma clara. Ainda que as três palavras denotem conceitos abstratos, o significado da primeira (‘love’) é altamente comum para falantes da maior parte das línguas, mesmo que o significante seja diferente para cada idioma (‘amor’, ‘amour’, ‘kjærlighet’, ‘אמור’) fato que facilita a

compreensão do seu sentido. Ainda que um leitor desconheça a palavra *'love'*, o conceito em si não é necessariamente estranho. Se traduzirmos a palavra como *'amor'* para um falante de português, ele não terá dificuldade de compreender o significado da palavra, mesmo ela sendo abstrata.

Por outro lado, a segunda palavra (*'tough'*) pode ser complexa não só porque o significado é um pouco menos comum que *'love'*, mas porque sua ortografia é similar – e geralmente confundida - a outras palavras do inglês (*'thought'*, *'though'*, *'through'*) e porque sua pronúncia é peculiar.

Por fim, o terceiro exemplo (*'wistfulness'*) ilustra o tipo de palavra que é complexa não somente pelo fato de não evocar um significado tão fácil (talvez uma criança, por exemplo, não entenda o significado de melancolia), mas também pela sua ortografia complexa (palavra formada por um radical e dois sufixos).

Os três exemplos mencionados acima nos ajudam a compreender que a concretude lexical pode ser adotada como característica indicativa de potencial de complexidade, mas que sua análise é muito mais profunda do que a quantidade de palavras concretas ou abstratas que existem em um texto. Ao utilizarmos o Coh-Metrix para coletar o valor de CLP de um texto, devemos utilizar o resultado como sinalizador de textos com alta presença de palavras abstratas para que possamos, manualmente e cuidadosamente, verificar o potencial de complexidade das unidades não-concretas. Nossa pesquisa nos auxiliou a desenvolver uma classificação dos valores de CLP que podem ser utilizados para estimar potenciais de complexidade textual, conforme exibimos no quadro abaixo (Quadro 121):

Quadro 121 – Referências de complexidade de CLP

VALORES DE REFERÊNCIA DE CLP	
Fácil:	80 a 100
Médio:	54 a 79
Difícil:	0 a 55

Fonte: Elaboração própria

Por meio do quadro acima, uma professora de ILA que tenha por objetivo tornar textos de livros didáticos mais acessíveis para alguns alunos pode verificar quais deles resultam em um valor de CLP interpretado como “médio” ou “difícil” e verificar se as palavras abstratas do texto são, de fato, potencialmente complexas. Dessa forma, é

possível adotar estratégias que diminuam a complexidade dos materiais, conforme ilustramos no quadro abaixo (Quadro 122):

Quadro 122 – Estratégias para aumento de acessibilidade de CLP

CARACTERÍSTICAS QUE INFLUENCIAM NA CLP	ESTRATÉGIAS PARA AUMENTO DA ACESSIBILIDADE DE CLP
1. Frequência de palavras concretas	a) Verificar a proporção entre palavras concretas e abstratas. Se um texto sugerir maior presença de palavras abstratas, é possível que uma ou mais tenham significados que sejam difíceis de serem interpretados pelo leitor. Diante dessas palavras, pode-se adotar uma das seguintes estratégias: a.1) Substituição por um sinônimo mais conhecido; a.2) Eliminação da palavra caso ela não seja essencial para o texto; a.3) Adição de um aposto explicativo sobre a palavra (desde que escrito de forma acessível).
2. Significação das palavras abstratas	a) Observar se o significado das palavras abstratas é conhecido na língua-mãe dos estudantes. Muitas palavras abstratas têm referenciais conhecidos independentemente do idioma. Caso essas palavras sejam comuns, é preciso verificar se outros traços delas não sejam complexos (como a ortografia, por exemplo). Diante de unidades que são comuns, mas que a grafia ou extensão seja potencialmente complexa, adotar o passo a.1 da orientação anterior.

Fonte: Elaboração própria

Como destacamos a respeito de todas as métricas anteriormente apresentadas, é preciso utilizar a CLP de forma coerente com as finalidades pedagógicas. Se o propósito de uma unidade é desenvolver vocabulário sobre sentimentos, por exemplo, a eliminação ou substituição de palavras abstratas que se referem a emoções pode ser prejudicial ao objetivo do texto.

Na próxima seção, apresentamos a última métrica de léxico utilizada na nossa pesquisa.

10.11 COMO UTILIZAR A RELAÇÃO TYPE-TOKEN (TT) PARA ESTIMAR COMPLEXIDADE

O índice *Relação Type-Token* (TT) é utilizado para observar a variedade lexical em um texto. *Tokens* é o número total de palavras, enquanto *Types*, quantidade de palavras únicas. Naturalmente, um texto é geralmente composto por muitas palavras que se repetem, logo, a proporção *Type-Token* raramente é 1:1 (fato que só

aconteceria caso nenhuma palavra do texto se repetisse). Se tomarmos a frase “Ele comeu pizza e depois comeu lasanha.” como exemplo, temos o total de sete *tokens* (por ser uma frase com sete palavras) e de seis *types* (visto que a palavra ‘comeu’ se repete). Para essa frase, o valor de TT é 0,857.

Variedade lexical está altamente relacionada com potencial de complexidade, uma vez que quanto mais variado for o léxico dentro de um texto, maior é o vocabulário que é necessário ter dominado para a interpretação. Entretanto, esse índice não deve ser utilizado sozinho, porque um texto pode ter alta variedade lexical, mas todas as palavras serem simples e outro pode ter baixa variedade lexical, mas com muitas unidades complexas. Sendo assim, aconselha-se a adotar o TT como métrica aliada a outras que observem complexidade das palavras (como a IA e a FPC, por exemplo).

Partindo do trabalho de Silva (2018), verificamos os valores de TT na nossa pesquisa e corroboramos sua classificação de valores em relação à interpretação de complexidade. Na nossa tese, simplificamos as categorias de interpretação em três, conforme apresentamos no quadro abaixo (Quadro 123):

Quadro 123 – Referências de complexidade de TT

VALORES DE REFERÊNCIA DE TT	
Fácil:	Abaixo de 680
Médio:	679 a 749
Difícil:	Acima de 750

Fonte: Elaboração própria

Se utilizarmos o quadro acima para interpretar a frase que apresentamos anteriormente como exemplo (“Ele comeu pizza e depois comeu lasanha.”), poderíamos dizer que ela é considerada difícil, uma vez que seu valor de TT é 0,857. Entretanto, como leitores do português, sabemos que as palavras utilizadas são comuns e simples. Esse exemplo confirma nossa observação anterior de que esse índice deve ser utilizado em equilíbrio com os demais.

Diante de valores altos de TT e especialmente quando estivermos trabalhando com textos para alunos de ILA de níveis iniciais (ou com limitações linguísticas ou neuroatípicos), a estratégia apresentada no quadro abaixo (Quadro 124) pode potencializar a acessibilidade dos textos:

Quadro 124 – Estratégias para aumento de acessibilidade de TT

CARACTERÍSTICAS QUE INFLUENCIAM NA TT	ESTRATÉGIAS PARA AUMENTO DA ACESSIBILIDADE DE TT
1. Quantidade de palavras únicas	a) Verificar a presença de muitas palavras únicas que tenham o “mesmo” significado. Caso existam muitas palavras que expressam a mesma ideia, escolher apenas uma e repeti-la.

Fonte: Elaboração própria

Voltamos a enfatizar que a adoção da estratégia supracitada deve ser coerente com os objetivos pedagógicos e que ao aplicá-la é preciso cuidar para não diminuir o potencial de manutenção linguística do texto. A repetição de palavras (para evitar apresentar um vocabulário muito expandido em um só texto) é uma estratégia altamente conhecida para fins de acessibilidade textual de materiais escritos para falantes nativos (visto que, nesses casos, o objetivo do texto é apresentar alguma informação). Já para aprendizes de línguas, essa estratégia deve ser utilizada quando não interferir no desenvolvimento de vocabulário proposto pelo texto. Obviamente, como docentes, sabemos quais palavras são essenciais para cada momento dos nossos alunos e que, em muitas vezes, algumas unidades lexicais foram escolhidas para evitar repetição sem que fosse levado em consideração que quanto maior a variedade lexical, maior é o léxico que os estudantes precisam saber para interpretar o texto. Em um material escrito para o nível básico, por exemplo, a palavra ‘*paramount*’ pode ser facilmente substituída por ‘*important*’ ou ‘*tough*’ por ‘*hard*’ sem desservir o propósito pedagógico do texto, uma vez que ‘*paramount*’ e ‘*tough*’ são mais complexas e geralmente aprendidas em níveis mais avançados. Por outro lado, evitar substituir a palavra ‘*eat*’ por ‘*have*’ ou ‘*beautiful*’ por ‘*pretty*’, por exemplo, pode reduzir o potencial de um texto de auxiliar o aluno a ampliar seu vocabulário, visto que essas palavras são comuns e que suas presenças nos textos colaboram para o aprendizado.

Neste capítulo, revisitamos as métricas utilizadas nesta tese e suas possíveis relações com análises de acessibilidade textual de forma a instruir professores e pesquisadores a como utilizá-las para observar potenciais de complexidade em textos de ILA.

Gostaríamos de enfatizar, novamente, que a adoção de índices textuais para fins de análises de complexidade deve ser coerente com as propostas pedagógicas e sempre sustentada por verificações humanas de professores e/ou linguistas; valores métricos não podem ser tomados como indicadores supremos de complexidade –

especialmente porque, conforme já citamos, acessibilidade/complexidade é sempre em relação ao leitor para o qual o texto será destinado. Além disso, ao adotar uma ou mais das estratégias sugeridas para textos de ensino de línguas adicionais, é preciso tomar o cuidado para não prejudicar a manutenção linguística (ou seja, para não remover ou substituir informações e componentes dos textos que sirvam como insumo linguístico proposto pela unidade na qual o material é utilizado).

11 RESUMO DA TESE EM LINGUAGEM SIMPLES

O QUE NÓS FIZEMOS NA PESQUISA?

Neste trabalho, nós queríamos descobrir se textos usados para aprender inglês são difíceis para os alunos brasileiros. Na verdade, nós já sabíamos que algumas características são complicadas para os alunos porque somos professores.

Então, o que nós fizemos foi escolher dois livros usados para ensinar inglês e analisar os textos que estão neles.

QUE LIVROS ESCOLHEMOS?

Nós escolhemos dois livros que são bem conhecidos entre professores de inglês. Eles são de uma coleção chamada '*English File*' da editora Oxford.

Escolhemos os livros que são usados com o nível básico e com o nível pré-intermediário. O nível básico é o primeiro (quando os alunos começam a aprender inglês). O nível pré-intermediário é depois do básico (mas ele não é um nível avançado. O pré-intermediário é um pouco antes do nível médio).

Nós escolhemos esses livros porque é normal termos dificuldade em ler textos quando a gente está começando a aprender uma língua. Então, conseguimos descobrir o que faz os textos serem difíceis para os alunos.

POR QUE ISSO É IMPORTANTE?

Analisar esses livros para ver se tem elementos que fazem a leitura ser difícil para os alunos é importante por alguns motivos. Esses motivos são:

1. Para deixar o texto mais fácil para que todos consigam entender.
2. Para tirar aquelas palavras muito difíceis e que não vão ajudar a aprender.
3. Para que os textos possam ser usados para ajudar os alunos a aprender inglês de um jeito melhor.
4. Para ajudar professores a facilitar outros textos que eles queiram usar para ensinar inglês.

COMO FIZEMOS?

Os passos que a gente seguiu para fazer a pesquisa foram:

1. Escolhemos os livros.
2. Pegamos só os textos dos livros.
3. Usamos um programa de computador para pegar umas informações sobre os textos. (Essas informações são números que nos dizem algumas coisas sobre os textos. Esses números podem ser usados para descobrir por que os textos podem ser difíceis).
4. Estudamos as informações que os números nos deram.
5. Comparamos os números dos textos dos nossos livros com os números de outros grupos de textos. Um dos grupos era de textos muito fáceis. Outro grupo era com textos muito difíceis. Se os números dos nossos textos fossem parecidos com os números do texto fácil, nós iríamos entender que eles são mais simples. Já se os números dos nossos textos fossem parecidos com os textos mais difíceis, nós iríamos entender que os textos são mais difíceis.
6. Vimos se a dificuldade dos textos aumentou de forma regular. Por quê? Porque temos que usar textos que aumentem pouco a pouco a dificuldade se nós queremos usar textos para ensinar inglês.
7. Depois de ver que alguns textos são muito difíceis, nós apresentamos umas ideias para fazer os textos ficarem mais fáceis.

O QUE DESCOBRIMOS?

Como nós pensávamos antes de fazer a pesquisa, alguns textos são difíceis para os alunos. Isso não quer dizer que todos os textos são difíceis. E não quer dizer que tudo no texto é difícil. Algumas informações que descobrimos foram:

1. Muitos textos têm palavras muito difíceis. Palavras difíceis atrapalham a leitura dos alunos.
2. Alguns textos têm informações que não precisavam aparecer.
3. Alguns tipos de palavra podem dificultar a leitura.
4. Pode ser complicado entender um texto quando algumas palavras não estão nos seus lugares mais comuns.
5. Tem algumas coisas em alguns textos que são difíceis para cada nível. Por exemplo, se um aluno está no nível básico, algumas características do texto são difíceis para ele. Se o aluno está no nível pré-intermediário, outras características do

texto podem ser complicadas para ele. Isso não significa que o texto vai ser sempre difícil. O que significa é que dá para mudar algumas coisas em alguns textos para facilitar o aprendizado.

E O QUE FAZER?

O que fazer depende do nível. Tem algumas coisas que a gente pode fazer se a gente quer deixar um texto mais fácil para os alunos do nível básico. Tem outras coisas que a gente pode fazer para os alunos do nível pré-intermediário.

Vamos apresentar alguns exemplos em português, mas os textos que a gente estudou nesta pesquisa são em inglês. As 10 informações que a gente acha mais importante são:

1. Ver se as palavras são muito difíceis e trocar por outras.

Exemplo: João é **formoso**.

Trocar por: João é **bonito**.

2. Diminuir a quantidade de certas palavras (tipo ‘adjetivos’, que são aquelas palavras que servem para descrever alguém ou alguma coisa).

Exemplo: Os médicos indicam esse **ótimo** remédio.

Trocar por: Os médicos indicam esse remédio.

3. Cuidar para ver se as palavras têm uma ligação clara.

Exemplo:

Se eu disser “**João ama seu pai. Ele é legal.**”

Quem é legal? O João? Ou o pai? Fica difícil saber, não é?!

Trocar por: “João ama seu pai. **João** é legal.”

4. É interessante colocar algumas palavras nos lugares mais comuns onde elas geralmente aparecem. Dá para trocar a posição das palavras se elas estiverem fora do lugar.

Exemplo: No verão passado, Maria viajou.

Trocar por: Maria viajou no verão passado.

5. Diminuir o tamanho do texto.

Textos muito longos são difíceis. Se o texto for muito longo, é preciso verificar se tem informações que podem ser retiradas.

6. Se uma palavra for muito grande e tiver outra menor com significado igual, é bom usar a palavra menor.

Exemplo: Os valores são **equivalentes**.

Trocar por: Os valores são **iguais**.

7. Colocar as palavras mais difíceis em um quadrinho com a tradução ao final do texto também ajuda.

8. Dividir as frases em duas ou mais se as elas forem muito grandes.

Exemplo: Maria foi ao mercado ontem e comprou café e, depois, foi ao cinema para assistir uma comédia.

Trocar por: Maria foi ao mercado ontem. Ela comprou café. Depois, Maria foi ao cinema para assistir uma comédia.

9. Usar as matérias e as palavras que estão sendo ensinadas também ajuda.

Se os alunos estão aprendendo as cores em inglês, por exemplo, é importante colocar vários exemplos das cores no texto.

10. Sempre ver quais são as facilidades e dificuldades dos alunos antes de fazer qualquer mudança nos textos.

O QUE É PRECISO LEMBRAR?

É preciso lembrar de não tirar dos textos os conteúdos e as palavras que são ensinadas na unidade onde o texto está. Também é importante lembrar de não tirar matérias e palavras que já foram estudadas. Além disso, tem que saber escolher o que vamos fazer com cada texto para não tirar do texto coisas que são importantes para o aluno aprender inglês.

12 CONSIDERAÇÕES FINAIS

No âmbito do ensino de línguas adicionais, metodologias modernas de aprendizado (como o Ensino 3.0, 4.0 e 5.0) estão cada vez mais enfáticas sobre a importância de customização e personalização de práticas e de materiais didáticos para respeitar múltiplas inteligências e habilidades individuais e para alcançar diferentes perfis de alunos e, assim, oportunizar acesso ao aprendizado de forma justa. Metodologias e crenças antigas (que acreditavam que todas as pessoas aprendiam da mesma forma e que um único material seria suficiente para promover o acesso ao conhecimento para todos os alunos de forma igualitária) estão sendo cada dia mais descreditadas, não só pela evolução da compreensão do significado da Didática, mas também graças aos avanços tecnológicos. Esses avanços têm permitido o desenvolvimento de ferramentas computacionais que assessoram a coleta de dados relativos ao funcionamento cognitivo humano (como as máquinas de ressonância magnética) e que auxiliam profissionais de diversas áreas a mapear com maior precisão como a nossa mente funciona.

Além disso, na última década, principalmente, temos percebido (e advogado a favor) de movimentos inclusivos e de práticas que respeitem a diversidade - visto que em um planeta habitado por quase oito bilhões de pessoas⁸¹, a diversidade não é exceção. Assim, professores que buscam oportunizar acesso ao conhecimento aos alunos de forma justa (e não só a grupos seletos, como tem sido feito ao longo da história da civilização, que tende a favorecer as classes sociais mais altas) devem aprender a ensinar e a adequar materiais de ensino para diferentes perfis (não só para contemplar múltiplas inteligências, mas também para oportunizar acesso ao conhecimento de forma eficiente para estudantes neuroatípicos). Se cada pessoa aprende de forma diferente (seja por habilidades, por experiências de mundo, por limitações, por questões cognitivas, por aspectos sociais, ou por qualquer outra razão), é evidente que a padronização de práticas e de materiais e que a tentativa de “tentar colocar todos os alunos na mesma caixa” são ideias que precisam ser abandonadas. Se temos informações e ferramentas que nos auxiliam a potencializar o ensino, devemos adotá-las frequentemente. É preciso sempre lembrar que escola

⁸¹ <https://www.cnnbrasil.com.br/internacional/populacao-mundial-deve-ultrapassar-marca-de-8-bilhoes-ainda-em-2022/>

não é chão de fábrica e que os alunos não estão em uma esteira de produção em massa com o objetivo de formar “produtos” exatamente iguais.

No que diz respeito ao ensino de línguas, portanto, a customização e personalização de materiais de apoios para aulas de línguas adicionais é sumária. Parte dessa customização é a adaptação de textos que servem como *input* linguístico. Visto que textos são recursos essenciais na manutenção linguística, é preciso potencializar suas eficiências e, como afirmam Silva, Moll e Perna (2021), a quantidade de estratégias de acessibilidade é diretamente proporcional à quantidade de perfis de estudantes (ou seja, toda análise de complexidade e adoção de táticas de simplificação deveriam ser orientadas pelo perfil do(s) aluno(s) com quem utilizaremos os textos).

Uma vez que a reescrita com propósitos de acessibilidade é direcionada pelo perfil de leitor para o qual adaptaremos um texto, é importante que níveis e elementos textuais sejam escolhidos tendo esse sujeito em mente. Para isso, é possível conduzir pesquisas de complexidade textual apoiadas por *software* de PLN e, a partir das indicações de complexidade, escolher estratégias de simplificação que se adequem aos propósitos. Ao trabalharmos com textos propostos para o ensino de uma língua adicional (no nosso caso, o inglês), é essencial que possamos observar mais de um nível textual. Por conseguinte, a escolha de métricas que nos apresentem informações sobre léxico, semântica e sintaxe permite estudos mais completos e interpretações mais detalhadas de potenciais de complexidade.

No nosso trabalho, tivemos como objetivo principal analisar o potencial de complexidade de textos utilizados com estudantes de ILA de uma coleção de livros conhecida e de renome e que é adotada em diversos lugares do mundo. Para a nossa pesquisa, delimitamos um perfil de estudantes brasileiros adultos de nível universitário que estejam cursando disciplinas eletivas de ILA nos níveis elementar e pré-intermediário. Contudo, acreditamos que nossa metodologia possa ser replicada para outros perfis e que nossos resultados também podem ser válidos para outros tipos de estudantes (incluindo alunos neuroatípicos), desde que utilizados com coerência ao perfil desse aluno e ao objetivo da metodologia a ser empregada.

Ao finalizarmos esse trabalho, cremos ter resultados suficientes para retomar e responder nossas questões de pesquisa apresentadas no início desta tese:

a. Os textos didáticos da nossa amostra, conforme se apresentam originalmente, apresentam características que dificultam a compreensão e o aprendizado do inglês como língua adicional para estudantes universitários brasileiros nos níveis linguísticos elementar e pré-intermediário?

Sim. A maior parte dos textos apresenta características que potencializam a complexidade em, geralmente, mais de um nível. Por vezes, o léxico e a semântica são os maiores responsáveis por altas complexidades; em outras instâncias, a sintaxe. Entretanto, na maior parte dos casos, a alta complexidade é consequência não apenas de um nível, mas de características de todos.

Além do destaque dado por nós para alguns textos do nosso *corpus* de estudo no capítulo 8.5 desta tese (*'Análise semiautomática do corpus de estudo'*), as tabelas das métricas organizadas por cores indicativas de complexidade também podem ser utilizadas para direcionar a atenção de professores de ILA para traços potencialmente complexos nos textos.

Diante de métricas destacadas em amarelo ou vermelho nos quadros das métricas, docentes de ILA podem verificar comentários específicos na seção de análise ou utilizar o capítulo 10 desta tese (*'Orientações para acessibilidade textual linguodidática'*) para interpretar os resultados e adotar estratégias de simplificação que considerem adequadas e coerentes para os seus estudantes.

b. Que características podem contribuir para a possível complexidade desses textos?

As características potencializadoras de complexidade dependem de cada texto. Como mencionado no parágrafo de cima, maiores complexidades ou acessibilidades são resultados de diferentes traços para cada um dos textos: por vezes léxico e semântica, por vezes sintaxe, por vezes os três.

c. Existem textos dentro de um mesmo nível que se destacam dos outros por apresentar características de alto potencial de complexidade textual?

Sim. Alguns textos apresentam mais traços indicativos de complexidade do que outros. No nível *Elementary*, por exemplo, o texto E12 tem quatro métricas de léxico

e semântica e quatro índices de sintaxe que indicam complexidades altas e médias. No livro *Pre-intermediate*, o texto P17 é um exemplo de destaque de alta complexidade, visto que todas as métricas de léxico e semântica e quatro dos índices de sintaxe são indicativas de complexidades médias e altas.

Esses dois textos são apenas exemplos de cada um dos níveis. Por meio dos quadros das métricas, é possível visualizar outros que também se destacam dos demais por terem maiores quantidades de índices classificativos de baixos potenciais de acessibilidade.

d. Quais indicações de estratégias de simplificação podem ser adequadas para cada um desses níveis linguísticos de aprendizado?

As indicações variam de acordo com cada texto (visto que a complexidade é percebida em traços distintos para cada um deles). Sugestões pontuais de adequação para fins de acessibilidade podem ser encontradas no capítulo 9 desta tese (*'Sugestões de estratégias de acessibilidade para os textos do corpus de estudo'*).

Contudo, é importante voltar a destacar que as estratégias apresentadas são apenas sugestões provenientes da nossa pesquisa e não normas fixas a serem seguidas. É essencial ter coerência na escolha das estratégias e lembrar que a complexidade/acessibilidade é direcionada pelo perfil e necessidade dos estudantes.

e. Existe uma progressão de complexidade gramatical e lexical (naturalmente esperada no desenvolvimento linguístico) nos textos estudados para garantir a manutenção linguística?

Não de forma constante, mas não inexistente completamente.

Os valores das métricas dispostas em gráficos-linha auxiliam a visualizar que a progressão de complexidade raramente é constante. Como comentado anteriormente, existem estratégias que poderiam ser adotadas para adequar a progressão de forma a otimizar a manutenção linguística.

Acreditamos que – futuramente - poderemos nos concentrar em estudar estratégias que sejam voltadas especificamente para fins de adequação de progressão a partir dos dados da nossa pesquisa.

Ainda que a progressão não seja constante, ela não é inexistente. Conforme explicamos no capítulo de análise de progressão, textos são manifestações linguísticas que se configuram de diferentes formas. Logo, seria impossível observar um comportamento perfeitamente regular e progressivo para todas as métricas. Entretanto, observamos que existem diferenças muito grandes (“saltos” grandes de redução ou aumento de complexidade) não só entre um texto e outro, mas também entre os níveis. Cremos, então, que seria válido estudar com mais detalhes as características que acarretam essa não-linearidade na progressão.

Além das respostas para as perguntas de pesquisa, nossa análise permitiu que desenvolvêssemos conclusões sobre nossas hipóteses, que retomamos a seguir:

Hipótese 1: Embora adaptados para o nível linguístico aos quais são propostos, os textos presentes na quarta edição da coleção *English File* da editora Oxford ainda apresentam características que dificultam a compreensão de leitura e, conseqüentemente, a manutenção linguística para estudantes brasileiros universitários de inglês como língua adicional nos níveis elementar e pré-intermediário.

Corroborada. Ainda que os textos tenham sido adaptados para os fins didáticos (como indica nas descrições dos livros e dos textos), ainda existem características potencializadoras de complexidade – conforme retomamos anteriormente.

Hipótese 2: Os principais fatores de complexidade são relacionados ao léxico e à semântica e os segundos, à sintaxe – especialmente pelo fato de que alguns aspectos desses níveis textuais são apresentados/ensinados em unidades e/ou livros distantes de onde aparecem (por exemplo, um conteúdo trabalhado no livro *Pre-intermediate* que está constantemente presente em textos do livro *Elementary*).

Parcialmente corroborada. Pela interpretação das métricas, percebemos um equilíbrio entre indicações de complexidade provenientes de léxico e semântica e vindas da sintaxe. Entretanto, nossa experiência com os textos do *corpus* de estudo em sala de aula nos indica que o léxico e a semântica são os níveis que mais causam

dificuldade para interpretação dos textos e, conseqüentemente, manutenção linguística.

Hipótese 3: Alguns textos apresentam uma ou mais características que os tornam potencialmente complexos, especialmente tendo em vista sua posição no livro.

Corroborada. Além das interpretações das métricas, a disposição dos índices em gráficos-linha e a nossa análise dos textos de forma individual nos permitem perceber que alguns textos são potencialmente mais complexos que os outros e que alguns desses materiais com altos níveis de complexidade são posicionados em partes iniciais do livro ou apresentam vocabulário e conteúdos gramaticais que são ensinados em unidades muito distantes e posteriores daquela onde estão inseridos.

Hipótese 4: A progressão de complexidade pode ser adequada dentro de cada nível e entre os diferentes níveis para garantir manutenção linguística dos estudantes.

Corroborada. Conforme explicado anteriormente, não há progressão constante na complexidade nem entre os textos de um mesmo nível, nem entre níveis. Algumas estratégias podem ser adotadas para aprimorar o desenvolvimento de complexidade e essas técnicas podem ser orientadas a partir das métricas que se destoam das demais ou que indiquem que determinados textos são mais ou menos complexos que seus anteriores e posteriores.

Hipótese 5: Existem estratégias de acessibilidade textual que se adequam para o perfil de aluno observado para cada nível linguístico e que podem ser aplicadas aos textos permitindo, simultaneamente, a compreensão leitora e a manutenção linguística.

Corroborada. Visto que o perfil-base dessa pesquisa são estudantes de ILA, algumas estratégias podem ser adequadas para promover maior acessibilidade e, ao mesmo tempo, permitir manutenção linguística. Em determinados textos, por exemplo, a substituição de uma palavra complexa por outra mais simples não tende a impedir o desenvolvimento de novos vocabulários. É possível substituir uma palavra complexa por outra mais simples que também seja nova para o aluno e,

potencialmente, garantir o aprendizado de uma nova unidade lexical. Posteriormente (e de acordo com o nível), pode-se apresentar a palavra que é considerada complexa e que foi substituída por outra mais simples.

Estratégias mais pontuais para cada texto são apresentadas no capítulo 9 desta tese (*'Sugestões de estratégias de acessibilidade para os textos do corpus de estudo'*) e metodologias de acessibilidade que também podem ser adequadas são apresentadas no capítulo 10 (*'Orientações para acessibilidade textual linguodidática'*).

Em suma, esperamos que nossa pesquisa possa conscientizar sobre a importância de observarmos a complexidade de textos voltados para o ensino e que ao adotar estratégias de simplificação, professores de ILA possam oportunizar um ensino mais justo para todos os perfis de alunos. Lembramos que a adequação de um texto com vistas à acessibilidade não tem por objetivo deixar um texto simplório, mas otimizar seu potencial de comunicação e, para fins didáticos, ser eficiente na manutenção linguística.

Acreditamos que nossa metodologia pode ser útil para orientar futuras práticas de análise de complexidade de textos utilizados para o ensino de línguas adicionais. Esperamos que nossas interpretações dos resultados e nossas sugestões de práticas de acessibilidade possam ser úteis não só para professores, mas também para pesquisadores e linguistas, de forma que as informações apresentadas nessa tese possam servir, talvez, como “pontapé inicial” para mais estudos acadêmicos de Acessibilidade Textual Linguodidática e, também, de outras tipologias textuais.

Futuramente, temos a pretensão de expandir a análise prévia que apresentamos do livro *Intermediate* e, quiçá, ampliar para os demais níveis dessa coleção.

Por fim, fica aberto nosso convite (e apelo) para professores e linguistas a investigar a complexidade textual, a estudar estratégias de simplificação e a advogar em prol de uma educação que possa ser acessível para distintos tipos de estudantes. Acessibilidade não é um diferencial, é um direito.

REFERÊNCIAS BIBLIOGRÁFICAS

- AZEVEDO, A. F.; DELGADO, H. O. K.; SILVA, A. D. C., The Use of Technology for EFL Classes in a Brazilian School: Consolidating Education 3.0. **IAFOR**, v.5, edição especial, 16 ago. 2017.
- BERBER SARDINHA, T. **Análise multidimensional**. São Paulo: Delta, v.16, no.1, 2000.
- BERBER SARDINHA, T. **Linguística de Corpus**. Barueri: Editora Manole, 2004.
- BERBER SARDINHA, T. **Linguística de corpus**. Barueri: Manole, 2000.
- BERBER SARDINHA, T. **Tamanho de corpus**. São Paulo: The ESPecialist. v.23 no. 2, 2003.
- BIBER, D. **Dimensions of Register Variation – A Cross-Linguistic Comparison**. Cambridge: Cambridge University Press, 1995.
- BIBER, D. **Variation across Speech and Writing**. 1a ed. Cambridge: Cambridge University Press, 1988.
- CÉSAR, S. V. D., PROF. PE. J. V. Abstrato ou Concreto. **Revista Letras**, v. 9, 25 nov. 2010.
- DAVIDSON, Cathy. Standardized tests for everyone? In the Internet age, that's the wrong answer. - The Washington Post. **The Washington Post**. Washington D.C., 23 set. 2011. Disponível em: <https://www.washingtonpost.com/opinions/standardized-tests-for-everyone-in-the-internet-age-thats-the-wrong-answer/2011/09/21/gIQA7SZwqK_story.html>. Acesso em: 10 jul. 2022.
- DUBAY, W. H. **Smart language: readers, readability, and the grading of text**. Costa Mesa, Calif.: Impact Information, 2007.
- DUBAY, W. H. **The principles of Readability**. Costa Mesa: Impact Information, 2004.
- DUBAY, W. H. Unlocking Language: The Classic Readability Studies. **IEEE Transactions on Professional Communication**, v. 51, n. 4, p. 416–417, dez. 2008.
- DUNKEL, B.; SOPARKAR, N. **Data organization and access for efficient data mining**. Proceedings 15th International Conference on Data Engineering (Cat. No.99CB36337). **Anais...**Sydney, NSW, Australia: IEEE, 1999. Disponível em: <<http://ieeexplore.ieee.org/document/754968/>>. Acesso em: 27 ago. 2022
- EUROPARAT (ED.). **Common European framework of reference for languages: learning, teaching, assessment ; companion volume**. Strasbourg: Council of Europe Publishing, 2020.
- FENFANG, L. Relationship between EFL Learners' Belief and Learning Strategy Use by English Majors in Vocational Colleges. **Journal of Language Teaching and Research**, v. 1, 1 nov. 2010.
- FERRER, E. et al. Longitudinal models of developmental dynamics between reading and cognition from childhood to adolescence. **Developmental Psychology**, v. 43, n. 6, p. 1460–1473, 2007.

FILMORE, C. **Corpus linguistics or computer corpus linguistics**. In: Directions in corpus linguistics. Proceedings of nobel symposium 82. Nova York: De Gruyter, 1992.

FINATTO, M.J.B. **Acessibilidade Textual e Terminológica: um Novo Tópico de Pesquisas em Terminologia no Brasil 1** - Estudos Geossociolinguísticos do Português Brasileiro. Porto Alegre: Editora Pontes, 2018.

FINATTO, M.J.B. **Complexidade Textual em Artigos Científicos: contribuições para o estudo do texto científico em português**. Porto Alegre: Organon (UFRGS), v. 50, 2011.

FLESCHE, R. **A new guide to better writing**. New York: Warner Books, 1984.

FLESCHE, R. **How to write plain English: a book for lawyers and consumers**. New York: Barnes e Noble, 1981.

FLESCHE, R. **Say what you mean**. 1st ed. ed. New York: Harper & Row, 1972.

FLESCHE, R. **The ABC of style: a guide to plain English**. Second Perennial Library edition ed. New York: Perennial Library, 1980.

FLESCHE, R. **The art of readable writing: with the Flesch readability formula**. 25th anniversary ed., rev.enl ed. New York: Harper & Row, 1974.

FLESCHE, R. **Why Johnny can't read: and what you can do about it**. New York: Harper & Row, 1955.

FLESCHE, R.; BURKHARDT, M. L. **Why Johnny still can't read: a new look at the scandal of our schools**. New York: Harper and Row, 1983.

FRIEDRICH, P. **Linguística do não-matar: Aplicações práticas**. Honolulu: Center for Global Nonkilling, 2012.

FRY, E. **Elementary Reading Instruction**. New York, NY: McGraw-Hill, 1977.

GARDNER, H. **Estruturas da mente: a teoria das inteligências múltiplas**. Porto Alegre: Editora Artes Médicas, 1983.

GARDNER, R. C.; TREMBLAY, P. F.; MASGORET, A.-M. Towards a Full Model of Second Language Learning: An Empirical Investigation. **The Modern Language Journal**, v. 81, n. 3, p. 344–362, 1997.

GILHOOLY, K. J.; LOGIE, R. H. Age-of-acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. **Behavior Research Methods e Instrumentation**, v. 12, n. 4, p. 395–427, 1 jul. 1980.

GILHOOLY, K.; GILHOOLY, M. The validity of age-of-acquisition ratings. **British Journal of Psychology**, v. 71, p. 105–110, 13 abr. 2011.

GILMORE, A. Authentic materials and authenticity in foreign language learning. **Language Teaching**, v. 40, n. 2, p. 97–118, 2007.

GOLDIM, J.R. **Índices de Legibilidade de Flesch-Kincaid e de Facilidade de Leitura de Flesch**. Porto Alegre: UFRGS, 2003.

MATOS, F. G. DE. **Comunicar para o bem: Rumo à Paz comunicativa**. 2a ed. São Paulo: Ave Maria Editora, 2006.

GRAESSER, A.C., MCNAMARA, D.S., LOUWERSE M.M. **What do Readers Need to Learn in Order to Process Coherence Relations in Narrative and Expository Text?** Memphis: Guilford Press, 2003.

GRAHAM, S.; HARRIS, K. R.; LOYNACHAN, C. The Basic Spelling Vocabulary List. **The Journal of Educational Research**, v. 86, n. 6, p. 363–368, jul. 1993.

HARVEY, J. How the Ladder of Abstraction Can Help the Public Speaker. **Presentation Guru**, 2016. Disponível em <<https://www.presentation-guru.com/the-ladder-of-abstraction-for-speakers/>> Acesso em 15 de Outubro de 2022.

HAYAKAWA, S. I. **Language in thought and action**. 4th ed ed. New York: Harcourt, Brace, Jovanovich, 1978.

HAYAKAWA, S. I.; HAYAKAWA, A. R. **Language in thought and action**. 5th [rev.] ed ed. San Diego: Harcourt Brace Jovanovich, 1990.

HODGSON, J. The Status of Metalinguistic Skills in Reading Development. **Journal of Learning Disabilities**, v. 25, n. 2, p. 96–101, 1992.

HOLMES, G.; ABINGTON-COOPER, M. Pedagogy vs. Andragogy: A False Dichotomy? **The Journal of Technology Studies**, v. 26, 1 maio 2000.

IBPAD. **O que é Linguística de Corpus? – Veja 5 aplicações**. IBPAD. 2018. Disponível em: <www.ibpad.com.br/blog/comunicacao-digital/o-que-e-linguistica-de-corpus-veja-5-aplicacoes> Acesso em 06 de maio de 2018.

JOHNSON, A. P. **Teaching reading and writing: a guidebook for tutoring and remediating students**. Lanham, Md: Rowman & Littlefield Education, 2008.

KNOWLES, M. S. **The adult learner: a neglected species**. Houston [Tex.]: Gulf Pub. Co, 1973.

KRASHEN, S. D. **Principles and practice in second language acquisition**. 8. ed ed. New York: Phoenix ELT, 1987.

LATHAM-KOENIG, C. **English file. Pre-intermediate. Student's book: with online practice**. 4th ed ed. Oxford: Oxford University Press, 2019.

LATHAM-KOENIG, C.; OXENDEN, C.; LAMBERT, J. **English file. Elementary, Student's book**. Fourth edition ed. Oxford: Oxford University Press, 2018.

LATHAM-KOENIG, C.; OXENDEN, C.; LAMBERT, J. **English file. Intermediate, Student's book**. Fourth edition ed. Oxford: Oxford University Press, 2019.

MAGAÑA, S. **Disruptive classroom technologies: a framework for innovation in education**. First Edition ed. Thousand Oaks, California: Corwin, A SAGE Company, 2017.

MANNING, C.D., SCHÜTZE, H. **Foundations of Statistical Natural Language Processing**. Massachussets: MIT Press, 1999.

- MARKSON, L.; BLOOM, P. Markson, L. e Bloom, P. Evidence against a dedicated system for word learning in children. *Nature* 27, 813-815. **Nature**, v. 385, p. 813–5, 1 mar. 1997.
- MCCANDLISS, B. D.; POSNER, M. I.; GIVÓN, T. Brain Plasticity in Learning Visual Words. **Cognitive Psychology**, v. 33, n. 1, p. 88–110, 1 jun. 1997.
- MCNAMARA, D. S. et al. **Automated Evaluation of Text and Discourse with Coh-Metrix**. Cambridge: Cambridge University Press, 2014.
- MOTTA, E. Sentenças Judiciais e Acessibilidade Textual e Terminológica. **Domínios de Linguagem**, v. 15, n. 3, p. 761–813, 13 ago. 2021.
- PERFETTI, C. A.; LANDI, N.; OAKHILL, J. The Acquisition of Reading Comprehension Skill. Em: SNOWLING, M. J.; HULME, C. (Eds.). **The Science of Reading: A Handbook**. Oxford, UK: Blackwell Publishing Ltd, 2005. p. 227–247.
- PERNA, C. L.; SILVA, A. D. C.; DELGADO, H. O. K. Recursos digitais bem-sucedidos para aprimorar as aulas de língua inglesa. **Ilha do Desterro A Journal of English Language, Literatures in English and Cultural Studies**, v. 74, n. 3, 8 set. 2021.
- SAUSSURE, F. **Curso de Linguística Geral**. 28a ed. (2012). São Paulo: Cultrix. 1975.
- SILVA, A. D. C., **Guidelines de Plain Language adaptados para o Português**. Grupo TEXTEC. UFRGS, Porto Alegre, 2016.
- SILVA, A. D. C., Textos de Divulgação para Leigos sobre o Transtorno do Estresse Pós-Traumático em Português: alternativas para a acessibilidade textual e terminológica. 2018. 427 p. *Dissertação de Mestrado. UFRGS, Porto Alegre, 2018*.
- SILVA, A. D. C.; DELGADO, H. O. K.; FINATTO, M. J. B. Acessibilidade textual e terminológica para o português brasileiro: pesquisa, estratégias e orientações de [re]escrita simplificada. **MOARA – Revista Eletrônica do Programa de Pós-Graduação em Letras ISSN: 0104-0944**, n. 58, p. 322–343, 31 jul. 2021.
- SILVA, A. D. C.; MOLL, E. DA S.; PERNA, C. B. L. Acessibilidade textual e endereçamento: contribuições bakhtinianas para estratégias de simplificação textual. **Revista Gatilho**, v. 20, n. 01, 3 set. 2021.
- STEPHENS, C. **The neuroscience behind writing for fluency**. Plain Language Association International e Journal, v. 3, n. 2, 2021.
- WIGFIELD, A.; GLADSTONE, J. R.; TURCI, L. Beyond Cognition: Reading Motivation and Reading Comprehension. **Child Development Perspectives**, v. 10, n. 3, p. 190–195, set. 2016.
- YAMAZAKI, M. et al. Two age of acquisition effects in the reading of Japanese Kanji. **British Journal of Psychology**, v. 88, p. 407–421, 1 ago. 1997.
- YANG, N.-D. The relationship between EFL learners' beliefs and learning strategy use. **System**, v. 27, n. 4, p. 515–535, dez. 1999.
- ZEVIN, J. D.; SEIDENBERG, M. S. Age of Acquisition Effects in Word Reading and Other Tasks. **Journal of Memory and Language**, v. 47, n. 1, p. 1–29, 1 jul. 2002.

APÊNDICES

QUADROS ORIGINAIS DO CEFR

Descrição do CEFR de habilidades globais de leitura

	Overall reading comprehension
C2	Can understand virtually all types of texts including abstract, structurally complex, or highly colloquial literary and non-literary writings. Can understand a wide range of long and complex texts, appreciating subtle distinctions of style and implicit as well as explicit meaning.
C1	Can understand in detail lengthy, complex texts, whether or not these relate to their own area of speciality, provided they can reread difficult sections. Can understand a wide variety of texts including literary writings, newspaper or magazine articles, and specialised academic or professional publications, provided there are opportunities for rereading and they have access to reference tools.
B2	Can read with a large degree of independence, adapting style and speed of reading to different texts and purposes, and using appropriate reference sources selectively. Has a broad active reading vocabulary, but may experience some difficulty with low-frequency idioms.
B1	Can read straightforward factual texts on subjects related to their field of interest with a satisfactory level of comprehension.
A2	Can understand short, simple texts on familiar matters of a concrete type which consist of high frequency everyday or job-related language. Can understand short, simple texts containing the highest frequency vocabulary, including a proportion of shared international vocabulary items.
A1	Can understand very short, simple texts a single phrase at a time, picking up familiar names, words and basic phrases and rereading as required.
Pre-A1	Can recognise familiar words/signs accompanied by pictures, such as a fast-food restaurant menu illustrated with photos or a picture book using familiar vocabulary.

Fonte: Europarat, 2020, p.54

Descrição do CEFR de habilidades da leitura para informação

	Reading for information and argument
C2	Can understand the finer points and implications of a complex report or article even outside their area of specialisation.
C1	Can understand in detail a wide range of lengthy, complex texts likely to be encountered in social, professional or academic life, identifying finer points of detail including attitudes and implied as well as stated opinions.

continua na próxima página

B2	<p>Can obtain information, ideas and opinions from highly specialised sources within their field.</p> <p>Can understand specialised articles outside their field, provided they can use a dictionary occasionally to confirm their interpretation of terminology.</p> <p>Can understand articles and reports concerned with contemporary problems in which particular stances or viewpoints are adopted.</p> <p>Can recognise when a text provides factual information and when it seeks to convince readers of something.</p> <p>Can recognise different structures in discursive text: contrasting arguments, problem–solution presentation and cause–effect relationships.</p>
B1	<p>Can understand straightforward, factual texts on subjects relating to their interests or studies.</p> <p>Can understand short texts on subjects that are familiar or of current interest, in which people give their points of view (e.g. critical contributions to an online discussion forum or readers' letters to the editor).</p> <p>Can identify the main conclusions in clearly signalled argumentative texts.</p> <p>Can recognise the line of argument in the treatment of the issue presented, though not necessarily in detail.</p> <p>Can recognise significant points in straightforward news articles on familiar subjects.</p> <p>Can understand most factual information that they are likely to come across on familiar subjects of interest, provided they have sufficient time for rereading.</p> <p>Can understand the main points in descriptive notes such as those on museum exhibits and explanatory boards in exhibitions.</p>
A2	<p>Can identify specific information in simpler material they encounter such as letters, brochures and short news articles describing events.</p> <p>Can follow the general outline of a news report on a familiar type of event, provided the contents are familiar and predictable.</p> <p>Can pick out the main information in short news reports or simple articles in which figures, names, illustrations and titles play a prominent role and support the meaning of the text.</p> <p>Can understand the main points of short texts dealing with everyday topics (e.g. lifestyle, hobbies, sports, weather).</p> <p>Can understand texts describing people, places, everyday life and culture, etc., provided they use simple language.</p> <p>Can understand information given in illustrated brochures and maps (e.g. the principal attractions of a city).</p> <p>Can understand the main points in short news items on subjects of personal interest (e.g. sport, celebrities).</p> <p>Can understand a short factual description or report within their own field, provided simple language is used and that it does not contain unpredictable detail.</p> <p>Can understand most of what people say about themselves in a personal ad or post and what they say they like in other people.</p>
A1	<p>Can get an idea of the content of simpler informational material and short, simple descriptions, especially if there is visual support.</p> <p>Can understand short texts on subjects of personal interest (e.g. news flashes about sports, music, travel or stories) composed in very simple language and supported by illustrations and pictures.</p>
Pre-A1	<p>Can understand the simplest informational material such as a fast-food restaurant menu illustrated with photos or an illustrated story formulated in very simple everyday words/signs.</p>

LISTAS DE PALAVRAS

Lista de palavras dos textos do livro *Elementary*

TEXTO E1 (110 types 191 tokens)	TEXTO E2 (124 types 258 tokens)	TEXTO E3 (147 types 298 tokens)	TEXTO E4 (118 types 238 tokens)	TEXTO E5 (189 types 364 tokens)
a	a	a	a	a
always	about	about	about	about
american	all	academic	after	above
an	also	after	agent	activities
and	and	all	am	air
are	are	also	an	albert
at	art	am	and	alexander
atlantic	awful	an	another	all
beach	bank	and	answer	along
before	beautiful	at	arrive	also
best	because	away	at	always
breakfast	britain	balance	back	amazing
building	british	be	because	and
bus	brother	because	breakfast	anywhere
but	but	between	brush	april
buy	call	break	but	are
by	camille	busy	calm	around
camera	canada	but	car	art
can	changes	call	club	at
car	china	can	coffee	atmosphere
center	chips	cardiac	cup	autumn
central	chorlton	daily	david	bad
cheap	cold	darius	day	be
city	come	day	dinner	beautiful
coney	cornwall	describe	dj	biggest
day	country	desk	do	birds
delicious	culture	dinner	door	botanical
dinner	darling	do	down	bowie
dog	days	does	drink	buckingham
don't	different	don't	eggs	buildings
down	do	during	emails	but
eat	does	europe	even	buy
empire	don't	every	evening	café
enough	eggs	example	excited	can
explore	especially	family	finish	cancel
floor	everybody	fantastic	first	career
foot	everything	for	for	celebrities
for	example	from	forever	ceramics
from	favourite	get	four	changing
go	fish	go	fruit	charge
good	food	good	get	children
grass	for	hairdresser	go	city
great	foreigners	hard	guetta	classics
has	france	have	gym	clocks
have	freedom	he	have	close
hot	friendly	helps	having	closes
hour	from	her	home	clothes
icon	go	him	hour	cold
if	great	home	hours	collections
important	happy	hour	house	colour
in	have	how	i	come
internet	he	husband	ibiza	daily
is	here	i	impossible	david

island it liberty listen long lunch meal more new of on online or park plan queues remember rent restaurant rockefeller salads sandwiches see serving shoes sit slow state statue steak street subway sunny swim take the them there these thing tickets tips tired to top traditional trip very visit walk with york you your	history hot i if in iraq is it khalal lian like live liverpool lives london lot love makes manchester me michelle minute my newcastle next nice of often oman one only or people potatoes primark pubs rains really says scenery scotch scotland she shops shuroog so street sweetheart take tang that the thing things think time to too topshop uk	in iran is isn't it jahangiri late learn lectures life lives london lot love lunch marjan me meetings more music my need news night nights of off often ok old on one only operations or originally papers phone piano play probably professional professors read reason relax research routine same she so son speak spend start successful surgery technical than	in is isn't job juice life like listen live manager me morning music my need never nightclub normal now of on or out outside playing restaurant say says secret security sleep spend stage starts studio sun superstar takes tea teeth thank that the then thing this three through to try two typical up usually very wake want when wonderful	day design designer different dinner do doesn't entrance europe every exhibitions expensive famous fantastic fashion favourite fee find finish for free from full furniture gardens go good great ground happily has highlights if in include including insects is it jewellery kew know london lot marathon marathons may mcqueen mean metres miles museum museums my near not of on one open
---	---	---	---	---

	<p>want we weather what when where who why women</p>	<p>that the their there they think time to too turn tv twice two up use usually values want watch way we week weekends what why with women work year you your</p>	<p>work you</p>	<p>or other out own palace park people performance picnic place places plastic prepared quite raincoats rains regent round route run runners</p> <p>see september shakespeare showers some sometimes special spend spring stadium summer sun take takes the theatre they this through tickets times to too top tourists tower trees treetop types until up use usually very victoria visit walkway watch</p>
--	--	---	---------------------	--

				weather wembley what when where which wind windy winter wonderful world year you your
TEXTO E6 (179 types 449 tokens)	TEXTO E7 (141 types 275 tokens)	TEXTO E8 (145 types 305 tokens)	TEXTO E9 (180 types 417 tokens)	TEXTO E10 (128 types 264 tokens)
a after again and animals answers are arrives asks at be beautiful bed black bottle box boy bread breakfast brother brothers but buy can careful clever clothes coffee coins cold comfortable day days decides desert don't drink empty evening every father fields find	a actually affects all also always and aretha bad bass be because below brain build builds but by called can change choice choosing coldplay comments compiled considered criteria data david day definitely download dr dragons drums early elements emotionally energetic energy favourite feel	a academy actor adele adkins album american an and anti art arts at back before best big billie billy blue book born brit britain british but c called camp caretaker child classmate dee different during eminem empire escape face fan father film films	a about after again ago all almost always amazing an and andy asked at awful back bar because before ben between big bit boots bought but button came car carefully celebrating completely could couldn't countdown dance danced day didn't dinner dj down drank	a about after again amanda an and anything arrived asleep at barbara bed bedroom between big breakfast but by came claudia closed come cup dead did didn't died different difficult dinner do don't door down finished first followed for found garden get go

finds	first	fine	dress	gordon
five	fits	first	early	got
food	for	five	electrical	granger
for	franklin	for	empty	had
from	from	from	eve	half
get	gently	gallery	everyone	hated
gets	getting	germany	exactly	he
give	good	he	expensive	hear
gives	greenberg	her	famous	heard
glass	happening	high	fantastic	him
goes	hard	his	fault	his
gold	has	holiday	felt	hour
goodbye	hate	however	few	husband
happy	have	ii	fine	i
has	he	immediate	fireworks	in
hassan	help	in	first	inspector
have	how	indianapolis	floor	into
he	i	interested	floors	is
help	imagine	into	for	it
her	in	is	forget	jeremy
here	including	isn't	found	killed
him	is	it	fourth	know
his	it	jedi	friends	last
home	its	jessie	get	library
hot	kinds	kurt	going	living
house	knowing	later	got	looked
hungry	la	laurie	guy	love
i	leave	leona	had	man
if	like	lewis	half	me
in	list	life	happened	midnight
inside	listening	lift	have	morning
into	little	london	hear	moustache
is	lot	made	help	mr
it	lyrics	major	home	mrs
just		more	hotel	my
kisses	many	most	hour	night
leave	mine	mother	i	no
leaves	moment	much	iceland	nobody
live	momentum	music	in	of
long	mood	musical	incredible	one
look	morning	national	into	opened
looks	music	new	it	please
makes	must	novel	jacket	quarter
money	necessary	now	jenny	read
morning	needs	of	just	room
mother	news	online	knew	rooms
mountains	of	operator	last	
my	on	painter	later	said
near	people	painting	lift	sat
need	perfect	performing	looked	saw
next	physiologically	piano	lovely	sleep
night	play	player		slept
no	positive	politics	me	somebody
now	prayer	portrait	midnight	spoke
old	prefer	prison	mina	talk
one	psychologist	project	minutes	tall
only	really	published	morning	tea
open	rest	return	music	tell
or	rhythm		my	that
palace	right	school	needed	the

part perhaps plan poor prince quietly re remember return rich ring room rooms same says see sell she shines silver sits sky sleeps small so some something strange sun surprised sweet takes thank that the their them then there they thing thinks this through tired to towards town two under until up valuable very village valid walks want wants	say says science seems shows slowly song songs spotify start strong that the them then there thing things think three to top up us using vida viva wake ways website what when which with world you your	schoolwork self seller she singer singers slaughterhouse soldier songwriter star strikes student success successful tastes teenagers than the theatre this to two very vonnegut war wars was was washington wasnt were when where williams work world writer writing years york	new next night not of on one only or out outside park part party people person phone planned play pressed programme put realize realized really remember restaurant said saw screens see slow so some start stopped strange streets stupid suddenly taxi that the then there they third this thought till time to took top tracks tv tvs two until	them then they this thought time tired to took travers twelve up walk want was we went were what when with yes you your
--	--	--	--	--

watch water way we welcomes what where why windows with wonderful work works you your			wait warm was wasnt watch watched we went were when where which with wore year years	
TEXTO E11 (158 types 292 tokens)	TEXTO E12 (154 types 274 tokens)	TEXTO E13 (114 types 202 tokens)	TEXTO E14 (114 types 255 tokens)	TEXTO E15 (187 types 314 tokens)
a again also and another appeared appears are arthur asleep at bathroom bbc bed being between bond britain built but central clothes conan couldn't cricket daily did doctor door doyle duke during early english especially evening eye famous film floating floor	a about affair agatha allan also an and appeared are arthur as asked at author authors baskervilles because best between bit book books born brown but by can centuries characters chesterton child christie come conan created crime day decades detective different	a about after albert all also an and avocado bacon because bit black bread breakfast but buying by can change cheese choosing clothes coffee common copy cream cup day decided diet different dinner don't eat eating eggs einstein every food for	a about actor afternoon an and another are ask asked at attractive because behind boyfriend but can card cards come concentrate couldn't country dark do don't door fall for fortune fourth from future going good he heart here him his house	a above adventure afford all also although amazing an and arctic are area as at balconies banderas bay beautiful bedroom below bikes bit boat book both breakfast bring buildings but by c cabins campsite can candles candlesticks canton carts caves circle

for	dog	get	i	classes
fourth	dorothy	good	idea	cold
german	doyle	green	in	comes
ghost	early	have	into	conferences
ghosts	edgar	having	is	corridors
golden	enormous	he	it	couples
got	famous	healthy	jane	dark
guests	fans	herbs	jim	day
happened	fascinating	i	last	deepest
has	father	if	let	different
haunted	fiction	in	look	dog
having	first	is	looked	dogs
he	flynn	it	love	don't
him	for	jobs	madame	down
himself	four	kinds	man	during
his	ghost	know	me	emergencies
holmes	gillian	less	met	enjoy
hot	give	lunch	mm	evening
hotel	going	mark	month	excellent
hotels	guess	me	move	explore
houses	he	meals	my	fires
hyde	her	money	new	fish
i	hercule	more	next	foot
iii	holmes	my	no	for
in	hound	need	not	four
including	how	never	oh	from
irish	important	not	opened	get
is	in	now	over	go
it	including	of	part	guesthouses
james	influence	on	party	guests
journalist	inspire	one	questions	guide
jumped	interesting	or	re	has
killed	introduced	piece	room	have
know	invented	plan	ross	heated
known	know	probably	said	honeymoon
langham	later	promise	says	hotel
lights	laura	re	second	hotels
like	learn	salmon	see	huskies
london	lee	same	sitting	husky
mail	like	smoked	small	if
man	lippman	so	soon	in
many	lived	some	sure	infinity
more	love	spend	table	is
morning	many	spices	tall	it
moves	me	steve	tell	join
murdered	millions	sure	teller	kennels
napoleon	modern	that	that	lakes
newspaper	more	the	the	light
night	most	them	there	like
of	murderer	they	third	lodge
off	mysterious	thing	to	love
on	mystery	thinking	today	lovers
one	name	this	turned	magical
opposite	nearly	time	very	maybe
oscar	new	tired	voice	metres
other	novel	to	want	mexico
out	novels	two	was	mine
park	of	using		mobile
past	only			morning
people				most

<p> players prince ran recently regent reported reputation room rooms saw say several shape sherlock sir situated sleep some stay stayed stories strange suddenly taps team that the then there things third through to told towards turned up very victorian walks was wearing well when who wife wilde window windsor with woke would writer wrote you </p>	<p> over past people perhaps plot poe poirot popular present probably read readers reading relationships rendell rowling ruth sayers sherlock short solves something sometimes still stories story styles such than the these they this thought to type us very view was we were when where who writer writers wrote </p>	<p> vegetables waste wear well what why with wore worry year you zuckerberg </p>	<p> well went what who with works worry written yolanda york you your </p>	<p> nature need nervous new next no not of offers on only or other people perfect phones place pool puerto pull radio remember restaurant romantic rooms runs sala schwyz see seminars siberian sides silver silvermine sledges sleep small so spa spend suite summer surface sweden swiss ten the their then there they this though three time to too training travel </p>
---	--	---	---	---

				treatments trees tuna underground unusual vallarta variety verana views warm watching wedding well whale where who winter with wood work world would yoga you your
TEXTO E16 (191 types 371 tokens)	TEXTO E17 (211 types 420 tokens)	TEXTO E18 (164 types 305 tokens)	TEXTO E19 (216 types 471 tokens)	TEXTO E20 (150 types 287 tokens)
a about all also always america american and are arrived as at be because belarus believe bills boss but buy by call came can canada chef cold could couldn't country cut dangerous	a about abroad after again ago all amsterdam an and are argentina as assistant at australia away back balloon beach because beginning best book bucket but buy can canary car chile christmas	a address after ago all alphabetical always amazingly an and any appointment are argument asked at basic because before big book books but can cassette cd chinese conversation couldn't days did didn't	a about actor always am american an and andrew answering any are as at audio away beautiful because bed before being best between biography book books bridges but by cambridge can character	a about again alan always an and are aren't as asked at atmosphere bar be because been best big birthday bit book booked bridget busy but cake can daughter definitely delicious didn't

day didn't do doctor doctors doesn't don't drunk earn earned eat ecuador enough even everything example explain felt first food for friends frightened from give go going good got group hair hairdresser half have he here his home hospitals hotels hours huge husband i in india insurance is isn't it julie just know late lena life like live lot loud	come coming company concert cook cookbook couldn't country day decided didn't different dive do don't dream dreams during end ever every experiences explore festival field finally finished first flat flowers fly follow food for friend from germany get go got guitar happy hardly have having here hi holiday holidays hometown horse i imagine important in is islands istanbul it job	do easy email encyclopaedias encyclopedia enormous exists fat fax for francesca genuinely glasses go had hairdresser have heavier home how idea if in information intern internet interview introduced is it job know library life like listen listened local looked lot machines magic maps meeting mentioned miracle morning mp music names needed new no number of office on online or order	choose city classical concerts cook countryside day desk did discovering do dogs earned easily eat edinburgh emails encyclopedia english enjoy especially etc ever family favourite february filming first for foreign four french friends from gandalf get go going hammerstein hammersteins hardly have historical holiday home hour hours house how i ian if in india information instrument internet interview is it	different dinner don't early eat enjoy excellent exciting expensive fantastic feel fine flat food for free friday friendly fun george get go going good got great have her here home however i if in is isn't it jane less lot made menu mind minutes more most my near night noisy not obviously of often ok old online only or other
--	---	--	---	---

luxuries me medical money more my name natasha need new next night of officers often old on or our out outside pakistan patients pay people person places police polite portions president public quite quito rahim really rent richest safe salary same saw see seem share shirt shouting smart social some speak stop stopped street subway surprise surprised system take	jobs karaoke kilda language learn least left let life like list lot make me melbourne money months more morocco my need needed new next now of office on one or outdoor over passed passion pay photo photography photos play portugal prague probably project pub put rain read realized really recipes rest ride routine run save scares secret semesters	our over pages part people phone player post print read reading remember restaurant seconds send sending sent she slow small so some somebody something sometimes song sort started still street studies take takeaway technological teletext ten than that the them these they this thought to too town tv twelve up urgently very walkman walkmans wanted was we went were	its jazz job kind languages late learn learned leave letters like listen little living london lord love many mckellen me moment money more morning most movie much music musical my new news next night no not of often on one only or oscar papers parts people perhaps pets piano place play played plays pop prague prefer preparing professional read	our pay people perfect pizza pizzas place places popular produced quite really recommend romantic salads same sarah saturday service sit slow so starters steve table tasted the then there they thing three to two until us usually value very wait waited waiters waiting want was we well went were west when with year yesterday you
--	--	--	---	--

<p>tara teachers tell that the their them there they things think tip tipping to too transport two ukraine until usually very want was we wearing weekend well were what when who why winter with worked xbox york you</p>	<p>shout sitting six so some something sometimes spend spent</p> <p>stef studied studies summer switzerland sydney take taxi than that the them then these things three through time to travel travelled traveller travelling true two uk underwater very visa volunteer walk want wanted was week went what windsurf with work worked working would year years you your zealand</p>	<p>what when who win with years yellow you young your</p>	<p>reading recordings relax remember rings river room school see shakespeare sing sir small something speak spend student sudoku thames the theatre there thing things think three time to too traditional trilogy university up usually venice very views visited waiting was well went what when where which who william with wonderful working worst would written year yes you your zealand</p>	
--	--	---	---	--

Lista 2: Lista de palavras dos textos do livro *Pre-intermediate*

TEXTO P1 (178 types 324 tokens)	TEXTO P2 (145 types 295 tokens)	TEXTO P3 (164 types 342 tokens)	TEXTO P4 (158 types 300 tokens)	TEXTO P5 (83 types 164 tokens)
a	a	a	a	a
after	about	advertising	advantages	again
ago	afternoon	again	all	aim
agrees	again	agency	also	airport
alone	all	also	although	all
always	alps	although	am	alphabet
an	always	an	and	and
and	amazing	and	angles	beginning
are	an	animals	are	begins
ask	and	as	aren't	by
asking	another	at	around	can
before	anything	attacking	as	catch
believe	anywhere	autumn	at	complete
both	at	beach	back	contestant
bouchez	back	because	be	continues
boyfriend	bag	before	beautiful	correct
bruce	because	birds	beijing	definition
businessman	but	botanical	belgium	definitions
businesswoman	can	boyfriend	best	each
but	car	bread	better	example
buy	climb	bresson	birds	five
can	climbed	broke	breakfast	for
century	colder	but	bruges	from
changed	could	by	but	game
charlotte	couldn't	called	buy	gets
child	day	camera	camera	gives
children	decided	camping	can	goes
clever	didn't	cartier	city	got
clint	distance	coming	crowded	has
close	do	copy	day	he
dad	down	describe	details	if
date	email	early	didn't	in
dates	end	end	different	including
dating	felt	escaped	do	is
daughter	find	famous	enthusiastic	letter
destiny	for	favourite	evening	letters
different	forget	feature	exactly	minutes
difficult	forgotten	fell	example	next
divorced	friends	following	feeding	not
doesn't	get	for	for	of
don't	go	found	forbidden	on
drink	got	france	friends	or
end	group	french	from	pass
evening	had	frenchman	gardens	passed
everything	half	friendly	gave	passes
face	happens	friends	going	place
fans	haven't	from	good	plane
feeling	holiday	gardens	got	presenter
find	hour	give	had	prize
five	hours	got	happy	quickly
for	i	guardian	haven't	reads

friday	in	had	having	repeats
from	is	happy	here	right
fun	it	has	holiday	round
funny	jacket	he	home	say
generous	just	him	hotel	saying
gentleman	last	holiday	how	says
get	late	home	i	she
go	leave	i	in	start
going	let	immediately	is	starting
got	long	in	it	starts
had	looking	it	less	that
has	lunch	jane	light	the
he	make	job	little	then
helping	mantra	know	local	there
her	me	later	look	think
him	missing	little	looking	this
his	money	live	lot	to
hopeful	mountain	lived	lots	until
how	my	living	louvre	up
i	never	london	low	wheel
in	nice	looking	market	when
independent	no	lot	maybe	where
india	now	love	me	which
internet	of	magazine	memories	wins
is	oh	man	millions	with
it	ok	maybe	more	word
just	on	me	most	words
knows	one	medicine	much	wrong
later	only	met	my	you
like	or	my	need	
look	ourselves	newspaper	needed	
looking	out	night	new	
lot	own	nightclub	nice	
love	passport	noise	no	
man	phone	noticed	not	
many	photo	now	of	
meet	photos	number	often	
met	pocket	of	on	
mind	probably	often	one	
mum	really	on	only	
my	recording	one	other	
needs	said	opportunity	our	
new	sat	our	outside	
nights	say	owl	paintings	
none	sea	parents	paris	
now	see	paris	patiently	
of	send	parks	people	
often	so	people	perfect	
old	spent	perfect	photo	
on	started	phone	photographer	
one	stories	phoned	photographers	
online	stupid	photo	photos	
only	suddenly	photographs	picture	
or	sun	photos	pictures	
out	sure	picture	place	

over parent parents partner pays people perfect picture planning please preferably probably profile question re really reigate remarried right romantic saturday says several she single so someone sometimes sooner spark spending springsteen still surrey ten that the their there they think thinks this to together too treat trip up very walk want warm	take terrible the then there they thing things those thought tickets to took top two up us very view walk want wanted was wasnt we well went were what when where which while with year you your	places pocket press put rangeley re returned said saw secretary sent shy signed six some south started still stopped studying summer sunday sundays taking that the there they this time to together touch trees twenties unesco university up us very waited walk walking was we weekly went were when where why with working year	places postcard prefer quickly re record right same saw seen shining sky so someone spent strange sun surroundings take taking than that the there these thing think this three time times tips to took try two up useful very visit wait walk walking want warm was we went were when with	
--	--	--	---	--

was website were when who with woman women works would wrote year years you young		years young zoo		
TEXTO P6 (155 types 300 tokens)	TEXTO P7 (169 types 309 tokens)	TEXTO P8 (136 types 236 tokens)	TEXTO P9 (189 types 356 tokens)	TEXTO P10 (153 types 271 tokens)
a about aged always and another anybody appear are ask at away because becoming begin best bored boring boyfriend break british by colleagues details dinner don't even exciting facebook fake feel feeling for friend friends give had	a about according adopting all already also always and another appeared are as barista bars became bird blog brands breakfast brunch but by café can century change changing coffee combining come common companies contained created describe dictionary	a activities after ago and anymore apps are average bar be being better british broadband bus but can card changed connected credit delayed delivery depends doesn't don't download downloads drink eat episode even every everything expect facebook	a about according acid actually add adults advice aged alcohol alert all also amounts and are arthritis as at attack awake bad believe black blood bones both but butter by caffeine calcium can cancer cause cheese children	a about according advance all always americans an and anxiety at avoid be because bed better bottle break breakfast but by called can cent cheer college control course defensive don't each early eat even everybody everything exactly

has	difference	fast	coffee	exam
have	different	faster	connected	example
having	do	five	contains	fact
her	each	flight	contradicts	fear
home	english	food	cut	feel
important	enter	for	cuts	find
in	every	get	day	fine
increasingly	everyone	getting	delicious	for
instagram	example	google	diabetes	future
interesting	examples	growth	died	go
interviewed	existed	happen	disappear	goes
invent	existing	happened	disease	good
inventing	expressions	has	do	happy
invents	fi	hated	doctors	has
is	find	have	don't	have
it	first	hour	drink	help
james	for	how	drinking	home
jealous	foreign	i	drinks	how
judi	french	if	drug	hungry
just	from	immediate	effects	i
lie	gadgets	immediately	enough	if
lies	gastronomy	impatient	especially	imagining
life	gastropub	in	every	important
like	general	instant	extra	in
lives	google	is	fat	is
living	how	it	few	it
main	imported	leave	food	julie
make	in	less	foods	just
me	information	like	for	kind
media	into	live	from	late
men	invention	load	fruit	lead
monday	is		full	leave
more	it	made	get	lives
morning	italian	meet	give	
most	its	minute	glasses	look
much	knows	minutes	good	lots
new	language	mobile	has	manage
night	languages	more	have	massachusetts
of	late	most	heart	means
often	latte	need	help	my
on	like	netflix	high	negative
one	lot	new	how	night
ones	lunch	next	ideal	norem
online	make	now	if	of
only	makes	of	immune	often
or	many	on	improves	oh
organizers	meaning	only	in	optimist
other	meanings	order	includes	or
our	message	out	increases	out
own	months	page	increasing	pen
partner	more	patience	into	pencil
party	names	people	is	pencils
people	need	person	it	pens
person	networking	phone	jeremy	people
photos	new	picture	juice	per

popular	news	posting	keeps	pessimism
posts	newspaper	prepared	know	pessimist
psychologist	noise	quickly	last	pessimists
put	not	replacement	latest	place
re	noun	restaurants	laurance	plan
read	nouns	return	levels	positive
reality	now	searches	like	possible
really	oed	seconds	liquid	prepare
reasons	of	see	litres	prepared
romantic	often	series	live	problem
said	on	shopping	liver	professor
saturday	one	site	lot	psychology
school	online	so	low	puts
see	or	stuck	makes	re
seems	other	talking	may	results
she	oxford	ten	medical	run
shown	paper	than	memory	say
shows	people	that	men	says
sitting	phenomenon	the	middle	show
so	popular	they	milk	situation
social	pub	things	millions	situations
some	publishes	three	more	sleep
something	put	time	much	smile
stop	quickly	tinder	musclesstronger	so
stories	rage	to	natural	sometimes
sun	really	traffic	need	specific
sunday	recent	tv	no	strategy
survey	ringtone	users	not	studies
talk	road	wait	of	successful
tan	selfie	waiting	older	take
tell	site	walking	one	than
ten	skype	want	orange	that
than	smartphone	wasting	other	the
that	social	way	people	their
the	some	we	prefer	them
their	somebody	web	pressure	then
them	sometimes	week	prevent	there
there	speakers	whatever	problems	they
these	start	why	really	things
they	started	with	recent	think
think	stick	years	red	thinking
this	take	you	reduces	thinks
to	technology		research	thirsty
told	text		reviews	this
travel	that		risk	tired
truth	the			to
twitter	them		salt	up
venice	they		say	use
want	three		scientific	used
we	to		seems	very
website	together		short	water
weekend	tweet		small	wellesley
went	twitter		so	where
when	two		some	who
who	uk		sometimes	why

with women work you your yourself	understand update updates use verb verbs very video vlog was way ways we when where which with word words		stimulant stress studies study sugar suggest system tastes tea teeth term that the their they to too two types typical us very water we week weight what when which who wine with without woman women worse you young your	will worry wrong
TEXTO P11 (166 types 315 tokens)	TEXTO P12 (163 types 301 tokens)	TEXTO P13 (217 types 465 tokens)	TEXTO P14 (124 types 269 tokens)	TEXTO P15 (127 types 222 tokens)
a about adviser advisers african against also american americans an and aretha as	a abroad after all america an and any are around as asking at	a about address advantages after ago ah all am and angrily answer anything	a according all always and andrews another answer are at bags be because	a actor actress actresses after again age alfred all also and around at

at	bad	apartment	best	award
audience	be	are	book	baron
audiences	because	arlington	busy	be
august	been	arrived	by	became
award	brit	as	can	been
away	british	ask	card	best
be	brits	asked	cash	bestselling
became	but	at	chance	books
because	by	be	checkout	called
before	called	blue	checkouts	children
began	can	born	choose	christopher
best	carry	broadway	course	comedies
bring	cent	building	david	comedy
but	children	but	do	course
by	city	by	does	curtis
character	communication	calmly	don't	daughter
children	companies	can	express	death
church	company	city	fast	different
civil	continue	come	faster	director
clarence	could	completely	fastest	divorced
close	country	cost	few	episode
colour	couple	could	fewest	experiences
content	course	dark	give	famous
continued	decided	day	handed	father
day	did	dear	has	favourite
decided	different	detective	how	film
demonstrations	difficult	didn't	if	films
detroit	discover	dinner	in	finds
dream	do	distance	into	fish
end	don't	do	is	for
event	each	dollars	it	four
events	effort	don't	items	friends
famous	else	dreamily	join	from
felt	employees	east	just	globe
first	english	else	know	godchildren
five	everyone	enjoy	lane	golden
for	everything	evening	leave	got
four	example	ever	left	guest
franklin	expects	everything	less	gyllenhaal
friend	final	exclaimed	line	had
from	find	expenses	lots	haden
gave	finished	eye	men	has
give	for	eyes	middle	her
going	foreign	feel	more	his
gospel	future	find	most	hitchcock
grammy	gave	first	move	horror
greatest	go	follow	moves	husband
have	had	for	moving	in
he	have	forget	much	including
helped	he	found	new	inspiration
her	hear	friends	notice	is
his	him	from	of	jake
history	his	get	often	jamie
hostile	i	girl	on	janet
i	if	give	one	john

important	in	gold	only	kenneth
in	intensive	hair	other	kinds
including	is	happy	part	lady
inspired	it	hartlev	patient	lasted
into	its	hartley	paying	later
it	journalists	have	people	lee
itself	knows	haven't	probability	leigh
jackson	language	he	putting	lies
jones	languages	hear	queue	life
judged	latin	heard	queues	like
june	learn	henry	quicker	many
king	learning	her	right	marriage
kings	like	here	seem	married
knew	little	him	shorter	months
large	live	his	should	more
legend	local	home	simple	mother
little	london	how	slower	musical
live	long	h�loise	slowly	of
looked	madrid	i	so	old
loved	making	if	sometimes	one
luther	many	in	supermarket	or
mahalia	mark	interrupted	than	own
march	max	into	that	places
marches	menu	is	the	pretending
martin	met	it	then	psycho
member	month	kindest	there	queens
mentor	multinational	knew	things	recreated
minute	never	know	think	rest
more	newspaper	last	this	
most	nowadays	left	three	says
movement	of	less	to	scream
music	official	letter	too	screenwriter
musical	often	life	trolley	series
my	on	lives	two	several
nation	one	living	unloading	she
national	only	looked	up	starred
not	other	looking	use	successful
of	out	lucky	using	than
often	partly	man	usually	the
on	paula	masterfully	very	then
one	people	me	what	they
parts	per	men	when	third
people	phrases	mine	whereas	this
preachers	really	moment	which	times
presenter	restaurant	montgomerys	why	to
previous	say	moved	will	tony
public	school	mr	winner	trading
racism	see	much	with	true
radio	sending	must	women	tv
rights	situations	my	won	two
said	so	never	you	wanda
sang	some	new	your	was
say	sometimes	no	yours	when
she	soon	nobody		which
shouted	spain	not		won

shouting singer singers skin so soundtrack south speaking speech speeches suddenly talked tell than that the their them then there they threw to towards usa voice was washington well went what where who will win with woman words world written year	speak speaks spend stop survive teacher ten test tests that the their then there they think thirty this time to too tourists true try typical until visit waiter want weekend went when who with world you young	of offer office often on one only other out paid paper part person piece promise promised quiet quietly red remember said saw say sea see she shopping short shouldn't skirt slowly smiling so someone sometimes stairs standing street suburbs such suddenlv sure suspicious tell ten thank that the theatre then there think to took top	writing written year years
---	--	--	-------------------------------------

		tram trust twenty understand up very visit vivienne you waiting walking want wanted was wearing week well went were what when where while white who why will window with words work would you your		
TEXTO P16 (155 types 296 tokens)	TEXTO P17 (222 types 363 tokens)	TEXTO P18 (215 types 417 tokens)	TEXTO P19 (162 types 295 tokens)	TEXTO P20 (238 types 559 tokens)
a about act actually again all always and are aren't as at athletes athletic babies ball basketball be	a after again ahn all also an and angel announced apologized are arguing as at athletics attacked back	a about academy acting actor adapted alex all already and anne answering anybody arctic are arm arrogant as	a about adopted after age almost an and another are as at away born both boys bremen brother	a about after again alarm alive all an and angry apartment argentinian around arrived asked asleep at australia

because	bad	asked	brothers	away
best	banned	at	but	back
between	bauer	author	by	baggott
bit	be	baker	café	because
both	beat	band	called	been
brandi	because	based	chance	behind
but	becoming	basketball	ciara	ben
can	began	be	coincidence	big
compare	beijing	because	colour	bought
could	blaming	become	conversation	box
difference	brazil	beginning	cordelia	but
differences	bronze	best	couple	buy
do	but	bit	dark	called
don't	car	black	dayton	callers
each	celebrated	books	dead	calls
enjoy	championship	bought	decided	came
equally	club	boys	did	car
especially	coach	broke	dublin	cat
events	competitions	building	english	city
ever	contract	bullied	erasmus	claudia
example	crashed	but	except	climbed
exciting	cry	by	fact	closed
fall	crying	came	few	collected
faster	cuba	cd	find	colwell
female	cup	changed	first	come
football	de	characters	for	coming
for	decided	child	found	confirmed
get	defeat	classmates	friends	contacted
gymnastics	dehydrated	clever	from	cookies
hand	delighted	coincidence	geaney	could
hard	didn't	coleman	genoa	couldn't
hardly	disappointed	collected	germany	couple
highest	disqualified	come	girls	cupcake
highlight	do	concerts	guizzardi	customer
hit	down	confidence	had	days
how	drink	could	hair	dead
hurt	drummond	days	he	details
i	eliseo	did	him	didn't
ice	error	didn't	his	do
identical	event	different	identical	don't
if	ever	don't	in	done
in	false	english	interested	door
individual	famous	especially	irish	dr
inferior	fans	everyone	is	driven
interesting	final	experiences	it	driving
it	fingers	fame	italy	drove
just	first	famous	jim	dvds
less	foot	fantastic	jims	earlier
level	football	first	just	ebay
like	footballer	for	knew	eight
made	for	get	later	eventually
male	from	girls	length	except
mean	furious	good	lewis	fallen
men	gaucci	got	like	false
more	gentleman	grew	lima	family

most	german	hair	lived	feel
never	germany	happy	look	filled
none	go	harry	lookalike	food
of	goal	has	looking	for
olympics	gold	have	luisa	found
on	grand	he	many	frightened
open	gucci	head	media	from
opponent	hard	her	met	front
other	hardest	him	miles	furniture
over	has	his	murphy	get
people	have	home	named	girlfriend
perform	he	honest	new	going
performs	head	hospital	niamh	got
person	here	housemaster	not	had
physical	him	i	nothing	hate
player	his	ill	ohio	have
players	hit	in	on	he
popular	host	incredible	one	heard
prefer	hours	inspired	online	help
pretend	how	interested	pale	her
probably	hwan	invited	parents	him
re	i	is	people	his
read	immediately	it	perhaps	hours
really	in	jealous	photo	how
research	intention	jk	probably	hurt
return	into	joanna	project	husband
right	is	joanne	put	i
run	italian	john	quickly	immediately
says	italy	jones	re	in
seconds	janeiro	jude	recently	into
send	jon	just	remarkably	it
serve	journalists	keeping	robert	journey
should	judge	know		julie
skating	jumped	later	same	jumped
ski	jung	law	scholarships	kick
slower	just	lead	see	kill
so	kick	lesson	seems	kilometres
sports	kicked	like	set	know
strengths	korean	liked	similar	large
stronger	kwon	long	six	last
summer	last	lyrics	size	later
swim	later	made	skin	left
team	lay	magazine	so	local
teams	learn	mark	social	looked
tell	left	me	something	losing
tennis	lesson	mike	soon	loud
than	let	money	springer	luckily
that	life	monkeys	stranger	
the	like	more	strangers	made
their	local	mr	surprising	man
then	lose	much	that	manager
they	losers	mum	the	me
think	losing	music	their	meanwhile
those	lost	my	them	microchip
thoughts	luciano	nettleship	then	mile

time to top true up us usually want watch watching we well when why will winter women you your	making match matos medal men metres minutes moments much nation nelson never no not of olympic olympics on one others out pass past paying perugia piquet played player president prix race re referee remind renewed rio robert ruined runner said salary salazar scored semi sent set several show so some somebody south sport sportsperson start	no obviously of on other outside pe perhaps person plays poetry popular possibly potter prison probably problem professor proud question quickly quiet real really remember reviews rowing said same school schools science secondary seriously she shows singer snape so some sometimes song songwriter spark sports started stories sure taught teacher teachers technician than that	there they thing think this time to together told truth try trying twin twins two typically uncommon university unrelated unusual up usa very was weeks were when who with women would years you young	miracle mistake mother neck neighbours no nobody noises notice of off officer old on one only opened or out outside owner packing petrol phone phoned pick playing police post posters put re realized really received replied returned rspca sad said scream screaming seat send several she shop shouting signal so some somebody son sorry spider
--	--	--	---	--

	<p> started still success swedish tae team technical that the their them then they time to told too track true trying two very was water we went were when who will win winning without won world would your </p>	<p> the their them then they think thought to together turner up us use used usually very wait walk was wasnt we week were weren't what when which with worked worried </p>		<p> spiders station stopped straight suburb suddenly sunday surprise survived sussex sydney terrible that the there thirsty throwing to toilet told took treated tried trying two unfortunate up very vet walter was wasnt water went were west what when where while who wife with without woke woman worthins year yes yesterday you your </p>
TEXTO P21 (128 types 278 tokens)	TEXTO P22 (174 types 323 tokens)			

a	a
about	airport
all	allowed
and	an
answered	and
arrived	apollonio
as	arrested
asked	arrived
back	as
before	at
believe	back
blowing	barman
blown	be
but	been
can	birthday
carried	board
come	but
cut	buy
damage	by
day	california
did	cards
directions	celebrate
do	clementi
don't	closed
done	continue
eastern	could
empty	couple
europe	court
every	decided
everywhere	defending
far	didn't
feather	different
feathers	dollars
flew	door
friends	engines
from	enrica
gave	erica
go	evening
good	family
gossip	find
gossiping	fined
had	flight
have	for
he	forced
her	found
hill	francisco
him	friend
his	friends
how	from
i	gate
impossible	get
in	got
is	greg
it	had
just	he

know	heard
like	her
loved	him
man	how
neighbours	however
never	hubert
next	i
not	if
now	in
of	instead
old	international
on	it
once	italian
one	italy
open	jet
other	just
out	know
people	last
pillow	lawyer
refill	leave
release	leaving
replied	let
round	like
said	lottery
say	lucky
see	malta
she	maltese
sit	man
small	matteo
so	might
some	million
speak	minutes
stories	monday
take	money
talk	my
that	near
the	next
their	not
them	notes
then	of
there	on
they	one
think	onto
time	open
to	other
told	out
tomorrow	outside
top	parashak
town	people
travels	pilots
true	places
up	plane
upon	planned
used	play
want	played

wanted	prison
was	prize
wearing	ran
went	real
were	removed
what	running
when	runway
who	ryanair
why	said
wind	san
wise	scratch
with	scratched
wives	security
woman	she
you	signalling
your	so
	spend
	spent
	spokesman
	staff
	stairs
	stop
	store
	street
	stuck
	tang
	ten
	terrible
	that
	the
	their
	them
	they
	thought
	thursday
	ticket
	to
	told
	top
	towards
	traffic
	tried
	two
	us
	used
	want
	wanted
	was
	way
	we
	wednesday
	went
	were
	when
	who

	with without won working works years
--	---

Fonte: Elaboração própria

Lista 3: Lista de palavras dos textos do livro *Intermediate*

TEXTO I1 (222 types 496 tokens)	TEXTO I2 (255 types 521 tokens)	TEXTO I3 (274 types 727 tokens)	TEXTO I4 (220 types 491 tokens)	TEXTO I5 (224 types 459 tokens)
a according acid adults after ago ahead alcohol alert also an and another are as at ate awake be because bed before begins best between blood body brain breakfast busy but by caffeine can capacity carried cause century coffee cortisol cup cycle day daylight dealing delicious diabetes digests dinner disease disrupting	a about abroad absolutely accommodation adventure afford after again all allowed also and anglia any anything appreciates are as at ate avoid back bag be beauty because before between bills bolognese boring bought brighton broadband budget buses but buy buying camped can challenge cheap chips cleaning clothes coffees confidence continued cooked	a absolutely admit afraid africa after afternoon again almeirim almost also amazon america and antarctica are arrived arriving away awful because become been before biggest bike bikes blog bottom brazil but camp can challenge changed charity checkpoint clean clothes complete completed completely could country cried cross cycling dad dARTH day daylight	a about across ahead airport all along an and approximately are as asked at back bbc be because bike birds boat bridge bus buses busy but by cameraman car careful cars challenge check chosen city clarkson classic close club colleague coming confusing could course crash cross cycle cycling decided delighted desk	a able about academic affected aggressive all along and answer are aren't argued as asked assembling at attract attracted australian author baby be believe better bosses but by california called can captions carl cartoons certain change cindy claim cliché comes common competitive completing conditions destination development did difficult direction discover discovered

do	cosmetics	days	direction	dr
doing	couldn't	decided	driving	drawing
don't	course	destination	ducks	driving
done	curry	disadvantaged	easier	easily
down	cycled	do	east	effective
drink	cyclng	done	far	eight
drinkers	day	down	fast	emotional
earlier	decided	dropped	fastest	every
early	didn't	during	felt	example
eat	difference	eight	few	faster
efficient	dinner	ended	find	feel
endurance	discovered	ever	finish	feelings
energy	don't	everything	finished	female
evening	drank	exhausted	first	field
experiment	drink	face	flying	find
experts	drinks	fast	football	finding
feel	east	few	for	five
few	electricity	filthy	forms	for
food	else	first	four	found
for	end	food	from	four
found	ended	for	frustrating	foxpz
from	essentials	four	fulham	from
general	even	freezing	gear	furniture
get	evening	from	get	gave
getting	every	full	getting	gender
give	example	get	go	general
glass	experimented	go	going	good
glucose	feel	goes	good	got
go	feels	going	got	group
good	few	gone	great	hand
greatest	finances	good	had	has
had	first	had	hammond	have
happier	fitter	hair	hated	here
have	flights	halfway	he	herefordshire
having	following	has	himal	how
healthier	food	have	his	humans
heart	for	heat	horrible	humour
help	freer	helen	hour	if
high	friend	help	how	ikea
higher	friends	helps	idea	in
hormone	gained	here	in	insecure
hours	garden	hit	into	is
however	gas	hot	it	it
ideal	go	hour	jam	jokes
ideally	gone	hours	james	judges
if	habits	however	jeremy	just
important	had	huge	journey	key
in	happier	i	kew	kind
is	hard	ice	lane	knowing
it	has	if	lanes	knowledge
its	have	in	large	language
less	healthy	incredible	lights	laura
levels	her	is	like	likely
light	here	it	liked	list
likely	herself	january	limit	lost

liver	holiday	journey	london	made
lives	home	kayaked	longer	magazine
lunch	household	kilometres	lose	make
lung	how	kite	lost	male
main	i	kited	man	many
make	ibiza	kiting	many	map
making	if	last	may	maryland
meal	in	late	meant	mathematics
melatonin	insists	leave	mercedes	maths
mentally	interest	like	miles	maze
midday	is	long	minutes	men
more	isolated	look	monday	mervin
morning	it	lot	morning	micks
most	join	lots	motorboat	minutes
much	journalist	make	moving	more
muscles	just	managed	news	mothers
natural	key	marathon	nine	moving
need	know	mask	not	necessary
not	large	may	now	new
now	last	me	of	no
nutritionists	lasted	mile	off	noise
of	learned	miles	ok	norway
one	less	minimum	on	norwegian
or	life	month	one	not
our	like	morale	only	of
out	long	morning	onto	often
people	looked	my	or	old
produces	lot	namibia	organize	on
producing	lunch	nauta	other	once
ready	made	need	others	other
recent	maybe	never	part	others
regular	mcgagh	night	past	out
research	meals	nightmare	permit	panic
researchers	meet	nine	phoned	partner
right	michelle	not	piccadilly	perform
rises	miles	now	point	pintzka
risk	missed	of	police	plan
run	moisturiser	off	presenter	preferred
same	money	on	public	problem
say	more	one	quickest	problems
science	mortgage	only	race	professor
see	most	our	realized	protein
sense	much	out	really	proved
sessions	need	outside	red	psychologists
sheffield	needing	over	remembered	questions
should	new	part	richard	rats
shows	no	past	right	really
skip	not	people	river	recent
sleep	november	per	round	relationships
small	now	peru	route	research
smell	occasionally	phone	rush	researchers
so	of	plan	see	restaurants
soon	often	polar	series	results
sports	oil	pole	set	said
start	old	poor	seven	same

starts	olive	practise	show	say
stomach	on	presenter	slowly	science
stop	one	progress	so	scientists
stops	only	promised	soon	seems
stress	open	proud	south	sense
strongest	or	pulling	speed	separated
study	out	pushing	square	showed
sugar	over	put	start	shows
suggested	past	reach	started	simply
suits	pasta	relief	steady	so
sun	pay	rest	stig	solve
super	paying	routine	stop	some
surrey	pays	run	stopped	speech
sweden	perfume	saying	stuck	stereotype
tasks	plan	scenery	terrible	stereotypes
taste	products	second	thames	stressful
team	pub	set	than	students
teenagers	quantities	she	that	studies
temperature	radical	should	the	study
than	really	since	them	tasks
that	reduce	six	then	team
the	restaurant	skelton	there	technology
their	restaurants	ski	they	than
they	rice	skiing	things	that
things	rode	skis	thought	the
this	sauce	sledge	three	their
time	save	slow	through	then
times	saved	smell	to	there
tired	say	snow	took	these
to	says	snowstorm	top	they
try	see	so	topgear	things
university	sense	soft	tower	this
up	she	some	trafalgar	three
us	simple	south	traffic	through
usually	simply	sport	train	time
versa	slimmer	starting	transport	to
vice	so	stay	tube	true
wake	socialize	staying	turned	turn
walk	some	still	tv	turned
want	something	stop	until	unemotional
we	spend	straight	use	university
weightlifting	spending	summer	valuable	use
were	spent	sun	wandsworth	various
when	stop	supplies	wanted	verified
whenever	survived	take	was	very
which	taken	talked	wasn't	virtual
who	tap	temperature	way	was
will	taxi	ten	went	way
wine	that	tent	were	we
with	the	terrible	west	were
working	their	that	what	when
worst	them	the	when	which
years	they	then	where	while
yoga	things	there	which	who
you	think	thing	worse	with

your	this those time tip to treatments tried trip unnecessarily using vinegar wallet want wanted wanting was water way we wedding week weekend went what when whenever where whether which who with woman work year you your	this thought three time times to today tonight too train training travelled travelling trek try trying tv ultra until up use using vader very visibility walking warm was wash washed water way we wearing weeks weird well went were when where which will wind winds with work world writing yesterday	would wrong	without women worked worse write yorker
TEXTO I6 (174 types 299 tokens)	TEXTO I7 (223 types 439 tokens)	TEXTO I8 (194 types 408 tokens)	TEXTO I9 (200 types 320 tokens)	TEXTO I10 (114 types 193 tokens)
a add	a about	a about	a abrahams	a act

after	acceptable	add	according	african
against	according	agnes	actions	alastair
alarm	advice	all	adrenaline	and
all	advised	along	agree	anyone
also	aeroplanes	already	allowed	arm
an	aggressively	also	also	as
and	air	an	always	at
another	all	and	an	because
anton	also	any	and	been
any	always	anyone	another	before
anyone	among	anything	any	began
are	an	app	as	behind
aria	and	articles	at	both
at	answer	australian	athlete	brother
audience	answers	be	athletes	brotherly
auditorium	any	because	bad	brownlee
available	appear	best	ball	but
backpacks	applying	better	be	by
bags	appropriate	books	been	conditions
ban	are	boyfriend	before	cozumel
be	armrest	brazil	best	cycling
beautiful	around	bus	better	didn't
because	as	but	both	done
before	ask	can	bounces	drunk
begins	asked	categories	but	enough
behaviour	at	change	can	entered
block	avoid	cinema	carry	fast
blocking	back	class	cause	feel
by	be	colombia	changing	few
can	been	comments	charm	final
cassie	before	copy	chemicals	finishing
caught	behave	could	club	for
check	best	couple	cologne	gold
christie	books	day	concluded	gone
cinema	both	describe	confidence	had
cloakroom	british	didn't	confident	hadn't
close	bryant	do	control	have
comments	but	does	couldn't	he
companies	by	doing	cristiano	help
company	changed	done	dan	helped
concert	clearly	easier	days	henri
concerts	closed	easy	distance	him
conductor	common	english	do	himself
could	consider	every	down	his
danny	contact	example	effect	hot
december	daytime	family	either	human
despised	debrett	fantastic	even	hundred
device	debretts	films	exam	i
doesn't	decline	find	experiment	ill
during	demonstrates	first	farah	in
ever	dilemmas	for	feel	instead
fine	diners	found	first	it
finishes	disorganized	fun	five	jonny
for	do	get	focus	kilometre
forgotten	don't	girlfriend	foot	km

from	during	go	football	last
give	eat	good	footballer	line
go	eating	got	for	long
going	editor	happening	forgotten	love
good	effects	has	from	medal
got	elderly	have	full	metres
great	enquiries	having	game	mexico
hall	environment	he	gift	my
halls	etiquette	head	given	natural
handel	everyone	helped	golf	nobody
has	experiment	helpful	good	of
have	final	holiday	grandmother	past
he	first	hope	had	race
him	flights	hungarian	happens	ran
hours	food	hungary	harris	reached
i	for	i	harrison	reaction
idea	frequently	if	has	remember
if	friends	important	have	road
immediately	front	improve	head	run
in	give	improved	help	running
interrupted	given	in	her	sacrifice
interval	gives	interested	his	sacrifices
it	giving	is	how	said
just	good	it	however	same
kasia	greeting	italy	i	saw
kind	guide	jazz	if	schoeman
leave	handbook	job	in	seconds
left	has	just	increase	side
live	have	karaoke	isn't	south
lose	help	language	it	stages
lot	host	laptop	judo	stopped
madrid	hostess	learn	just	swimming
marcel	how	learning	kayla	that
member	however	lets	last	the
messiah	hugely	level	leave	them
mobile	if	like	legend	then
mobiles	impersonal	lot	like	they
most	important	magazines	long	thing
museums	in	marc	lucky	to
music	inconsiderate	marta	make	took
must	individual	me	makes	triathlon
mustn't	is	medellin	may	until
national	it	mexico	memory	very
need	its	minutes	might	was
new	jo	moment	more	way
of	key	my	most	when
off	kicking	myself	negative	who
often	kiss	need	normal	will
on	kissing	nelson	not	win
one	latest	news	of	winning
or	london	noticing	on	won
other	lot	of	one	wonderful
out	make	on	onto	world
owner	makes	one	other	would
passages	manners	online	pen	

pay people performance phone phones pointed policies problem protested quickly reception right row ruined rule same should shouldn't shouted similar solve stage stopped tell that the theatres their them there they think this throws time to turned two until up was way we wednesday went when which who whose william with without work works would	many may meet meeting modern money monopolize more most nearly need needed no not number of offer offered on one only or other ourselves over own passengers people permission plane pregnant present preview problem professional public publisher putting queen question questions receive receives recent recline reclining remember restaurant rude said says seat seem selfish served	or outside own phone photography phrases plans practise practising pronunciation quickly quiet read reading really russia sandra sara saying screen see she sing singer sings so song songs speak speaking spoke sport studying subtitles switzerland tablet takes talk test than that the them then thing things think this time tip tips to too try understand	people perform performed person physical pitch played players positive problem professional psychologist psychologists race real realize really release researchers results right rituals ronaldo room runner same say says second seems sense serena serve shaves showed sit socks solve sport sporting sports star steps story study superstition superstitions superstitious take team tells tennis tests than that	
--	--	--	--	--

wouldn't written you your	short should simple situations slight smelly so social some sound space specializes stand start stay still subject table than that the their this those to today transport travelling trouble tube unclear unless up us used using very was we weeding what when which who whole within without woman year years yes you your	useful vasily version very vocabulary was watch way we week well what when whenever with without words you your yourself youtube	the their them they this those times to told twice two university used was wears well were what where which who williams with without wore work world you your	
TEXT0 I11	TEXT0 I12	TEXT0 I13	TEXT0 I14	TEXT0 I15

(186 types 361 tokens)	(275 types 598 tokens)	(213 types 495 tokens)	(102 types 166 tokens)	(80 types 132 tokens)
a actually acupuncture acupuncturist after ago all already an and another as ask asked at back bad be because beside but button came car celebrate checking cherokee chinese clinic connected continued crazy creative definitely degree didn't different do doctorate doctors door each eyes face final first following food for from get going got	a able about actors actually added addition again air alive all allowed alone also although amy an and any appear are aren't around as at be because been being ben big breath britain british budget busy but camera can cardboard cards cheer clap clapped cold coming conditioning conversation cost countries course create crowd	a about accomplished act adorable adult all also an and animal annoying any are as at baby be better book boring bosses boy boyfriend business busy but cartoon cat celebrity change character childhood children choice choices choosing clear close clothes common cool cover culture cute day death depends did does dog don't dress	a all also an and anyone are bbc be behaved body bored business but can century charges charisma charismatic cheerfully clooney coach conversation danish days don't enough everywhere example fields fix followed for from george going good hairdresser has have he his hour how i impressions in into is it just language learn	a according after age aged ahead and are asian better british can china chinese come countries days discipline education experiment five for four from gets groups half have highly hour in is latest learn mandarin maths most normal of produce rest results same schock school schools science see so some strict students studies

guy	crowds	else	life	style
had	cut	enjoying	like	subjects
he	dance	entertainment	make	successful
hi	days	entirely	man	system
him	deflated	etc	me	systems
his	degrees	even	meetings	take
home	dialogue	everybody	met	teach
how	digital	everyone	microsoft	teachers
i	digitized	everything	most	teaching
if	dimensional	exes	name	tests
in	doing	exist	negative	than
into	done	fabulous	not	the
is	dozen	facebook	of	their
isn't	during	family	paying	this
it	eastern	fan	people	three
jeep	employed	feel	plenty	to
keys	empty	fi	politics	two
knew	epics	find	quality	unique
kristina	especially	fired	says	vs
last	europe	first	seem	was
laughed	example	for	sheikh	week
like	expensive	friends	shops	weeks
little	extra	from	smile	western
lock	extras	fun	so	which
looked	extremely	get	student	will
lot	fake	girlfriend	talking	year
lots	film	go	that	years
marathons	filmed	good	the	
master	films	grown	thinks	
matt	fit	have	this	
maybe	for	having	to	
me	forbidden	here	turn	
medicine	funeral	icon	two	
met	gandhi	identity	watched	
mine	get	if	we	
moment	girl	image	when	
my	gladiator	in	with	
nearer	gloves	incompetent	work	
need	good	individual	worked	
nervously	had	is	worry	
new	hands	it	would	
no	hard	jealous	yahoo	
none	has	job	yesterday	
noticed	hats	jpeg	you	
of	have	kind		
old	hear	know		
on	homeland	least		
one	hours	life		
open	how	listen		
operation	hur	literary		
or	i	looking		
other	if	lot		
out	in	love		
pain	including	main		
passenger	individually	man		

personal	inflatable	married		
pressed	inside	media		
professional	insurgent	metal		
really	interfere	might		
rear	into	more		
replied	is	most		
resort	it	movie		
rob	jackets	music		
run	jersey	musical		
running	job	nappies		
said	just	never		
saw	keep	no		
scared	kinds	not		
session	know	now		
shopping	large	of		
show	largely	often		
side	last	old		
so	lasting	on		
some	later	one		
standing	latest	online		
started	less	or		
steal	life	other		
stopped	like	out		
store	live	own		
story	long	party		
study	look	people		
studying	looking	person		
suggested	lot	personality		
supermarket	made	pet		
surprise	make	photo		
that	making	photos		
the	means	picture		
them	miming	popular		
then	more	portrait		
there	most	posted		
they	mouths	poster		
things	moves	potential		
this	movie	probably		
thought	much	profile		
three	music	quite		
tiffany	need	reference		
time	needed	regret		
to	needs	replace		
trainer	never	same		
try	noise	say		
trying	not	says		
turned	noticed	school		
two	now	sci		
um	nowadays	search		
university	obsessed	see		
unload	of	selfie		
unloading	off	serious		
unlock	often	shirts		
up	on	slogans		
used	once	snake		

<p> very walked was wasn't way we week went what when while white work writing wrong years yoga york you </p>	<p> one ones online or ordinary other out outs paid particularly past patient people person phones photo photos playing possible posted pretend production put reading real really reasons regular remaining repeating reused right rogers said same says scene scenes secrets see set she shot shows silence silent since single sitting snow so some sometimes spend spent </p>	<p> so social something still stupid suit talk technologically teenage television than that the them then they thing things think thinks this time to too tv two type types up used usual want wear wedding what when who will with without woman you young your yours yourself </p>		
---	---	---	--	--

	stars stored strictly such summer tell than that the their then there they thick thing this time times to took totally touched truck turned tv two unrealistic up used usually very wait waiting was watch we wearing weather well what when where whether which while who why winter with without world would wrong you			
TEXT0 I16	TEXT0 I17	TEXT0 I18	TEXT0 I19	TEXT0 I20

(120 types 212 tokens)	(173 types 332 tokens)	(184 types 428 tokens)	(167 types 330 tokens)	(306 types 599 tokens)
a	a	a	a	a
accountancy	about	about	about	about
all	agreed	acknowledge	and	actor
alternatives	also	actually	any	actress
always	an	after	anything	addition
amazes	and	an	are	after
an	any	and	as	age
and	are	another	asked	aged
apprenticeship	ask	answer	at	ajax
are	asked	any	away	album
ask	at	anything	awful	albums
at	available	anywhere	back	all
benefits	be	are	because	almost
better	been	as	been	alphabet
but	believe	ask	before	also
choose	birmingham	assistant	bit	am
chose	bought	assistants	box	an
clubs	brightly	at	but	and
comments	brushes	away	by	april
company	business	be	came	aquatic
debt	businesses	becomes	can	architect
degree	by	being	connection	architects
degrees	can	best	continued	are
dentistry	clare	between	could	artists
depends	coloured	big	countryside	as
do	colours	break	course	at
earning	company	but	crying	barcelona
engineer	contestants	buy	dad	be
essential	convinced	call	dark	became
everyone	cost	came	days	before
experience	countries	certain	decided	best
find	create	colleague	desperate	black
finished	david	conversation	didn't	blackstar
first	demonstrated	corner	do	blunt
for	den	cross	don't	born
friends	designed	day	door	both
future	designers	didn't	down	boxer
gives	designs	difference	england	boxing
global	different	dinner	eyes	breakthrough
going	dragons	do	feel	british
got	dreaded	doing	few	bruce
has	england	don't	finally	buddhist
have	entrepreneurs	end	finished	buildings
how	especially	ending	flat	but
i	every	england	for	by
if	exchange	enter	found	california
in	faces	er	from	came
is	far	existence	get	campaigned
isn't	finally	expected	getting	canadian
it	five	far	go	cassius
job	for	five	got	centre
just	from	follows	hadn't	child
learning	gill		he	civil

life	good	football	hear	classic
like	hair	for	help	clay
lot	hairbrushes	friend	here	close
made	hairdresser	go	him	coach
many	has	goes	hitchhike	comments
me	have	going	home	considered
media	he	gone	hope	copies
medicine	his	happy	hospital	cry
money	i	having	i	curiosity
my	idea	he	if	darker
no	ideas	heard	ill	dated
not	if	hello	in	daughter
now	in	help	inspector	days
of	interchangeable	helpful	is	death
often	intimidating	helps	it	deaths
on	invest	here	knew	debbie
one	investment	hi	knocked	december
only	is	him	know	described
opinion	it	his	last	devastated
or	its	hope	later	die
other	jewellery	how	leeds	died
out	leeds	i	like	different
people	london	if	living	doves
place	make	improvement	local	dutch
preferred	many	in	london	early
really	market	is	look	eddie
recruitment	married	it	looked	educated
say	meet	just	make	entertained
self	men	king	maybe	equality
site	might	know	me	european
so	millionaires	later	minutes	ever
social	minutes	least	miss	every
software	money	like	missed	express
solution	more	london	mobile	fa
some	multi	look	months	famous
sort	new	looking	more	february
space	nicknamed	lost	mother	fight
still	north	lunch	mum	figure
student	not	major	my	film
studies	of	mall	nearest	films
taught	on	match	night	final
team	or	me	no	first
tech	original	minutes	north	fisher
that	out	my	now	five
the	own	near	of	football
their	pair	need	ok	footballer
there	people	needed	on	for
they	pitch	never	one	forward
thing	plastic	new	only	friends
things	possible	nice	open	from
think	prepared	no	out	futuristic
thinks	present	not	paid	go
time	presentation	now	peterborough	gold
to	produce	nowadays	phone	grammys
trust	product	of	phoned	greatest

two uk want waste what with worth you	profits programme questions range return ring room same say separately series service set share shaun show shows silver similar since single slappie so straps success successful take tangle teezer than that the their them then these they three tingatang to tv uk under untangling up usually versions very watch watches way wedding were where which	office oh ok part particular perhaps pretend problem question questions remember repeats rest round run sale sales say seems shop shopper shopping should shouts simple so sometimes station stop strange sure team tell thank thanks that the there think this to today train trip try trying uk um understand usa use very walk wallet wallets	phones please police problem ran rent rudely said sat saw see should so sol some something son soon sorry station story student talking tell ten that the then there thing think this thought ticket tickets to told train trains up upset very was wasn't well went when where who window with wondered would you	greek had hallelujah hard harry has hated he hear heavyweight her heroes him himself his history hits hour how hydra i icons in including influential inspired institute instrumentalist iraq is island it its james january jensen july june kill known later legend leia life lines literature lived london long loved made magnificent making man many
--	---	--	---	---

	<p> who with women words year you young your </p>	<p> want watch what when where who why will with without work yes you your </p>		<p> march marianne matches medal media memorable met million minneapolis minute mockingbird modern monk more most moved much multi muses musician name nicknamed no norwegian not novel novelist novels november object of olympics on one only opposite other our paul people performance played playing poet posted potter princess probably prodigy prominent public publication published purple rain </p>
--	--	--	--	--

				receive recorded refused released retreated reynolds rights role rowling royal self sensuous set severus shakespeare she shock shocked simon singer skills snape so sold song songs songwriter star states studio successful taught than that the them there they think three time times to tweeted tweets two uk united until us usa very vietnam voice want
--	--	--	--	---

				war wars was watchman we went were when where which who whom whose widely willis win with woman wonderful words work world worldwide worst written wrote year years you
TEXTO I21 (257 types 640 tokens)				
a	die	lawyer	pretty	tried
about	died	learnt	read	twenty
accept	discover	left	really	two
accepted	do	letter	reconciled	unfortunate
admirers	during	letters	refused	unknown
again	engaged	life	rendell	until
ago	engagement	like	replied	up
all	even	little	retired	very
almost	evening	live	return	violent
alone	extremely	looked	right	visit
always	fallen	looking	ring	walter
and	father	lose	ruined	want
another	fell	love	ruth	wanted
any	find	loved	said	warm
anything	finger	many	saw	was
as	first	march	say	wasn't
asked	fitted	marriage	shall	way
at	for	married	share	we
attack	forget	marry	she	wealthy
away	friends	may	sister	wearing
be	frightened	me	six	wedding
beautiful	from	met	so	went
became	funeral	money	some	were
because	furious	month	something	what

been	furniture	months	sometimes	when
before	girl	mother	sorry	where
big	give	mr	speak	which
born	given	mrs	start	who
both	good	much	still	with
brought	got	my	stole	would
but	great	neither	studying	wrote
by	had	never	summer	years
called	have	newspaper	sunny	you
came	haven't	nor	talk	young
can	he	not	teacher	younger
changeable	her	now	terribly	your
children	him	of	that	
church	his	oldest	the	
clever	holidays	on	their	
cold	home	one	them	
collected	house	only	these	
come	however	or	they	
comforted	husband	out	thing	
continued	i	over	this	
could	if	pain	those	
couldn't	in	parents	thought	
course	inside	part	thrace	
darling	intensity	passion	time	
daughter	it	passionately	to	
daughters	jewellery	past	together	
days	june	perhaps	told	
death	knew	personality	took	
decided	knife	pictures		
didn't	know	please		
	later	presents		

Fonte: Elaboração própria

TEXTOS SIMPLES DO CORPUS COMPARATIVO

LIS1

STOP WASTING FOOD!

The campaign to reduce food waste is in its third week, and we're already seeing some great results.

We'd like every student and teacher to join in the campaign at school and at home, so if you haven't started yet, read our blog and join in!

Stop wasting food at school

According to school lunch director Marge Kaplan, 'the children are making a big effort to think in a sensible way about the food they put on their plates. We now have about 15% less waste and we think this percentage will grow as more children join in the campaign.'

The school canteen now offers two portion sizes for all main meals: small and large. This week they'll replace single bread rolls with slices of bread. You can take as many slices as you want, but one at a time to save waste.

Latest news!

From next week you'll be able to put any waste from fresh fruit and vegetables into the special green bin at the back of the canteen. Miss Norris, the school gardener, will empty the bin each day into a giant compost container in the school's vegetable garden.

Stop wasting food at home

Every year, we waste about a third of our food. A lot of this waste is cereals, fruit and vegetables. In the UK, more than 97% of food waste ends up in a landfill site. That's a lot!

Come to a special meeting with supermarket owner Carlos Rodriguez to learn how we can save food and money at home!

(Disponível em: learnenglishkids.britishcouncil.org/reading-practice/food-waste)

LIS2

DIGITAL CITIZENSHIP

Everyone's talking about 'digital citizenship', but what is it?

We use the word 'digital' to talk about computer technology. Being a 'citizen' means that you belong to a group of people.

Digital citizens can use technology to communicate with others online and know how to use the internet in a safe and responsible way.

Are you a good digital citizen? Here are some ways to be safe and responsible online.

1. *Stay safe*

To stay safe, never give your personal information online, such as your address or the name of your school.

Make sure that you create strong, secure passwords and keep them secret. Don't tell your friends your passwords.

2. *Show respect*

Just like in real life, it's important to behave well online. For example, if you want to post a photo of someone, make sure you ask for their permission first.

Remember that everybody can have their own opinion. It's OK to disagree, but always be polite!

3. *Be kind*

Before you post a comment online, stop and think. Is it positive? How will your words make someone feel? Don't say something online that you wouldn't say in real life.

Never forward nasty messages. If you see that someone's being bullied online, tell a parent, teacher or another adult.

4. *Be careful*

Don't trust everyone or everything online. Remember that people online are not always who they say they are.

Never open emails from strangers, click on strange links or download apps without asking an adult. They could contain viruses. If you're not sure, ask your parents or teacher!

Spending time and learning online can be great if we're good digital citizens. So, be safe, be responsible and have fun!

(Disponível em: learnenglishkids.britishcouncil.org/reading-practice/digital-citizenship)

LIS3

THE CHICKENS TAKE A HOLIDAY

The sun was about to rise on Farmer Tim's farm. Chester Chicken woke up the cows with his important news.

"The chickens are taking a holiday today," Chester Chicken said.

"Is that so?" said Daisy the cow. "What is the special occasion?"

"We worked too hard this week," Chester said.

"You did?" asked Daisy.

"Yes! We laid ten eggs this week," Chester said, "and there are only five of us."

Daisy smiled and nodded her head. Ten was a lot of eggs for five chickens.

"Enjoy your day off," she said.

"But what about us?" the other cows said to Daisy. "We gave Farmer Tim 100 pails of milk this week. There are only ten of us!"

Daisy agreed with the cows too. 100 pails of milk would make a lot of cheese.

"But we can't take a holiday on the same day as the chickens," Daisy said. "What would Farmer Tim say?"

Daisy and the cows moved over to a patch of grass to have their breakfast.

"The chickens are taking a holiday," Daisy told the trees. "And we don't think it's fair."

The trees were not happy with this news.

"I've dropped over 1,000 apples this season," one said.

"And I've had a million cherries picked!" said another.

The wind blew and the trees put on their angry faces.

"We deserve a holiday more than the chickens!" the trees shouted together. "We worked too hard all season."

This woke up the rake that was sitting on the grass underneath the trees.

"Have you heard the news?" the apple tree asked the rake. "The chickens are taking a holiday. They think they worked too hard this week."

The rake stood up and announced its disapproval. "I raked over one million leaves this year. And there's only one of me! If anyone deserves a holiday it is a poor tired rake."

Just then Rowdy Rooster hopped on the fence. He looked up into the sky and began to crow. It was time for the farmer to wake up.

The chickens and cows and trees waited for Farmer Tim to come out and pick up the rake.

But a minute passed and Farmer Tim did not appear.

Rowdy called two more times.

"Call him again," the chickens yelled to the rooster. "He must be having a dream."

Rowdy made one last call and this time Farmer Tim woke up. But he didn't come out and pick up the rake, farmor milk the cows, or check on the eggs underneath the chickens. Instead he opened the window and shouted loud enough for everyone to hear:

"I worked too hard this week, I say. It's time I took a holiday!"

(Disponível em: englishclub.com/kids/stories/chickens-take-a-holiday.htm)

LIS4

LINDA'S PICTURE

This is a picture by my friend Linda. It shows a lively scene outdoors. I think she used bright pastels on white paper.

Linda's picture is so colourful. In the background there's a big blue sky. It's a sunny day. There are some tall buildings which are all different colours and shapes. In the foreground the grass is bright green and there's a path going up to the buildings. Maybe it's the park near her house? There are so many details in Linda's picture too. On the left there's a multicoloured cat with pointy ears. In the middle, there's a small blue cat with zigzag stripes on its tail. Above the cats there's a butterfly, flying around in the sunshine. Can you see the coloured spots on its wings? I like the yellow cat most of all. It looks a bit like a castle!

I love this picture because every time I look at it, I see different things. At first I thought there were only buildings in the background, but then I noticed the butterfly! Linda has so much imagination. Look at the right of the picture. What do you think the dark red outline shows? It isn't very clear, but sometimes I see a large bird with light blue wings, and other times I think it could be a streetlight. And I always see a different number of cats! This happens a lot in Linda's picture. It's really fun! It makes me think of a story I read once about a magic town, where everyone has adventures. I love it because it's bright and cheerful and it makes me feel happy.

I'd give this picture the title 'A Summer's Day'. I'd like to have it in my kitchen because it would make everyone smile in the mornings while we're eating breakfast.

(Disponível em: learnenglishkids.britishcouncil.org/reading-practice/art-description)

LIS5

FRIDA KAHLO

Early life

Frida Kahlo was a Mexican artist. Her full name was Magdalena Carmen Frida Kahlo y Calderón. She was born in Coyoacán, near Mexico City, in 1907. Frida grew up in a bright blue house called La Casa Azul with her parents and sisters. When she was six years old she became ill with polio and her leg was permanently damaged.

Starting to paint

Frida spent a lot of her life in and out of hospitals. She had many health problems and she suffered a lot of pain. When she was eighteen, she was in a bus accident. She broke lots of bones and she spent many months recovering. It was at this time that Frida started to paint. Many of her paintings from this time are self-portraits that show her pain and suffering.

Married life and career

In 1929 Frida married Diego Rivera, a Mexican artist who was famous for painting huge murals on walls. Both artists continued their work and became successful at home and abroad. After they got married, Frida started to wear traditional Mexican clothes and became interested in Mexican folk art. This influenced her paintings, which are very bright and colourful, with an original style that is very personal to Frida. However, they continued to show the pain and sadness which she experienced during her life.

Death and legacy

Frida Kahlo died on 13 July 1954, after suffering more and more health problems. Her husband died three years later. Today Frida and Diego's home, La Casa Azul, is a very popular museum, dedicated to Frida's life and work. It displays paintings by Frida and Diego, as well as many objects from their life, to help tell the story of one of the most important artists in the twentieth century.

(Disponível em: learnenglishkids.britishcouncil.org/reading-practice/frida-kahlo-biography)

LIS6

TOO TINY FOR TEA

Marty Mckay was already five years old, but he was still the baby of the family.

"Can I have some tea too?" Marty asked his mother. She drank her tea from a beautiful cup and stirred it with a silver spoon.

"No, Marty. You're too young to drink tea."

"But, why?" Marty asked.

"Because your fingers are too tiny to hold the cup. And tea is too hot for you, baby."

"I'm not a baby," Marty said. "I'm five-and-a-half."

Marty went out to the yard. His brother Ralph was playing basketball.

"Can I play too?" Marty asked. Ralph bounced the ball up and down under Marty's nose and then threw it into the basket.

"No Marty, you're too young to play basketball."

"But, why?" Marty asked.

"Because the basket is too high for you to reach. And the ball is too big for your tiny baby hands," Ralph said.

"I'm not a baby," Marty said. "I'm five and three quarters."

Marty went into the kitchen. His sister Jane was getting ready to ride her bicycle to the candy store.

"Can I go to the store to buy candy?" Marty asked Jane. He could feel the wind in his hair and the candy on his tongue.

"No, you're too young to go to the store," Jane said.

"But why?" Marty asked.

"Because the store is too far for you to ride to. And your baby bike is too slow."

"I'm not a baby," Marty said. "I'm nearly six."

"Six?" Jane laughed. "You just turned five!"

Marty sat on the grass and watched his sister ride away on her bike. He started to cry.

Marty's father was washing the car. He heard a tiny cry and went to find out what was wrong.

"Why are you crying?" Marty's father asked.

"Because I'm too tiny to do anything. I wish I weren't the youngest one."

"Be careful what you wish for," his father said.

Just then, Marty's mother came out to bring Marty's dad his tea. She patted her belly and smiled.

"We're going to have another baby," his mother said.

"And that means you're going to be a big brother," his father said.

"But, I'm too tiny to be a big brother," Marty said. "I'm just a baby!"

(Disponível em: englishclub.com/kids/stories/too-tiny-for-tea.htm)

LIS7

HOLI'S DIARY

I'm really excited! This week my friend Varinder's helping me make a video for my school. It's about Holi, a special Hindu festival we celebrate every year, usually in March. We're going to film some of the activities and next week we're going to show the class our video. I hope they like it!

11 a.m.

Last night we went to the big bonfire! Today the big Holi celebrations start. We went to the park this morning, where people were playing the drums and other musical instruments. People were dancing and singing everywhere - in the streets, outside temples and other buildings too. Everyone looked so happy! During Holi we celebrate the good things in life.

5 p.m.

Colours are important at Holi because it's the 'festival of colours'. That's why we put coloured powder into water and then throw the coloured water at each other. Some colours have meanings; for example, red means love, while green means a new start. The powder's been on sale in the markets for many weeks. I filmed it last weekend.

This afternoon we went to see my aunt, uncle and cousins after lunch. It was great fun. We used water pistols and we filled balloons with water and powder. Then we spent hours outside and threw the coloured water at each other. (I put a plastic bag over my camera, so that I didn't get water inside.) My aunt doesn't let us throw water inside the house, so we put coloured powder on each other's faces instead.

7 p.m.

It's the evening now. We're clean and in our best clothes. Now it's time to enjoy special Holi food and drinks. Our grandmother's prepared gujiya, a sweet made with dried fruit and nuts. She's also made my favourite drink: a kind of yoghurt drink called lassi. Later tonight we're going to dress up and visit the rest of the family. I love Holi! It's my favourite festival in the world!

(Disponível em: learnenglishkids.britishcouncil.org/reading-practice/holi-diary)

LIS 8

BIRTHDAY ON THE FARM

That's strange!

Mrs Brown's a farmer. Every day she gets up at six o'clock and Mr Brown usually makes her toast and eggs for breakfast.

But today Mr Brown isn't in the kitchen.

'Tom!' she calls. No answer.

'That's strange,' she thinks.

Mrs Brown goes outside and calls Max, her dog.

'Max!' she calls. Max doesn't come.

'That's strange,' she thinks.

Mrs Brown goes to feed the chickens.

She looks in the backyard. No chickens.

'That's strange,' she thinks.

Mrs Brown goes to see the cows.

She looks in the barn. No cows.

'That's strange,' she thinks.

Mrs Brown looks in the field. No animals.

'Where are the sheep? Where are the pigs? Where's the horse? That's *very* strange!' she thinks.

Mrs Brown hears a noise behind the house.

She goes to look.

Everyone's in the garden: Mr Brown, Max, the chickens, the cows, the pigs and the horse.

'SURPRISE!'

'Where's my breakfast, Tom?'

'Today's breakfast is cake, Maude. Birthday cake! Happy birthday, my love!'

(Disponível em: learnenglishkids.britishcouncil.org/reading-practice/birthday-the-farm)

*LIS 9***ECHO AND HERA**

Echo was a magical nymph. She lived in the forest and everyone knew two things about her. She had a beautiful voice and she was very kind.

One day Zeus ran into the forest. He wanted to hide from his wife, Hera. Hera was angry with him because he did not love her enough. Zeus saw Echo. 'Please help me,' he said. Echo was afraid of Zeus and because she was kind she agreed to help him.

A few minutes later, Hera came into the forest, looking for Zeus. She was unhappy. Echo stopped her and started to chat. 'Hello, Hera, how are you today?' she asked. Hera chatted with Echo and soon she forgot about Zeus.

After that day, every time Zeus wanted to hide from Hera, he went to the forest and looked for Echo. Every time Hera went to look for Zeus she found Echo. They chatted for a while and Hera forgot about her husband.

Finally, Hera understood that Echo was helping Zeus. She was very angry. She decided to punish Echo. She went to the forest and took away Echo's voice. From that day Echo could only repeat the last few sounds she heard. She never spoke again.

Today we call a sound that is repeated again and again 'an echo'.

(Disponível em: learnenglishkids.britishcouncil.org/reading-practice/echo-and-hera)

*LIS 10***MY FATHER MEETS THE CAT**

One cold rainy day when my father was a little boy, he met an old alley cat on his street. The cat was very drippy and uncomfortable so my father said, "Wouldn't you like to come home with me?"

This surprised the cat—she had never before met anyone who cared about old alley cats—but she said, "I'd be very much obliged if I could sit by a warm furnace, and perhaps have a saucer of milk."

"We have a very nice furnace to sit by," said my father, "and I'm sure my mother has an extra saucer of milk."

My father and the cat became good friends but my father's mother was very upset about the cat. She hated cats, particularly ugly old alley cats. "Elmer Elevator," she said to my father, "if you think I'm going [11]to give that cat a saucer of milk, you're very wrong. Once you start feeding stray alley cats you might as well expect to feed every stray in town, and I am *not* going to do it!"

This made my father very sad, and he apologized to the cat because his mother had been so rude. He told the cat to stay anyway, and that somehow he would bring her a saucer of milk each day. My father fed the cat for three weeks, but one day his mother found the cat's saucer in the cellar and she was extremely angry. She whipped my father and threw the cat out the door, but later on my father sneaked out and found the cat. Together they went for a walk in the park and tried to think of nice things to talk about. My father said, "When I grow up I'm going to have an airplane. Wouldn't it be wonderful to fly just anywhere you might think of!"

"Would you like to fly very, very much?" asked the cat.

"I certainly would. I'd do anything if I could fly."

"Well," said the cat, "If you'd really like to fly that much, I think I know of a sort of a way you might get[13] to fly while you're still a little boy."

"You mean you know where I could get an airplane?"

"Well, not exactly an airplane, but something even better. As you can see, I'm an old cat now, but in my younger days I was quite a traveler. My traveling days are over but last spring I took just one more trip and sailed to the Island of Tangerina, stopping at the port of Cranberry. Well, it just so happened that I missed the boat, and while waiting for the next I thought I'd look around a bit. I was particularly interested in a place called Wild Island, which we had passed

on our way to Tangerina. Wild Island and Tangerina are joined together by a long string of rocks, but people never go to Wild Island because it's mostly jungle and inhabited by very wild animals. So, I decided to go across the rocks and explore it for myself. It certainly is an interesting place, but I saw something there that made me want to weep."

(Disponível em: freekidsbooks.org/fathers-dragon-read-online-version/)

LIS11

These are Maria, Jeremy, and Peter. They are friends. They love their school because it's very big and cool. Maria has a short brown hair. Jeremy has small eyes and a big head. Peter has a curly black hair. Maria's purse is purple, and her shoes are pink. She loves pink color. Jeremy's bag is blue, and his pants are blue too, but his t-shirt is yellow and red. Peter has got a basketball. He likes to play basketball. His shorts are green, and his shoes are grey. He has got a red bag too. His favorite color is red. Maria is 8 years old; Jeremy is 9 years old, and Peter is 9 years old too. Maria is from Mexico. Jeremy is from Canada, and Peter is from Ecuador. They have different nationalities.

(Disponível em: en.islcollective.com/english-esl-worksheets/grammar/apostrophe-s/reading-comprehension-kids/97755)

LIS12

My name is Steve. I am eight years old. I love my green cap and my blue shoes. I was born in Pisa, Italy but at the moment I am living in Rome. My parents have a restaurant there. I haven't got any brothers or sisters, but I have a few good friends. I can speak English, German and of course Italian. My favorite dish is spaghetti.

Hi! My name is Yolanda. My hair is long and blond.

I like wearing skirts. I have one brother, Peter, who is ten years old. He is two years older than me. I have a sister who always wears trousers. Sonia is her name. Tina is my best friend. Her hair is red and is wearing a dotted blue dress. She never wears trousers because she hates them.

Hello friends, my name is Tom. I was born Manchester, but I am living in Barcelona now. I love going to school because I learn new things and also, I meet my friends. Roger is one of my best friends. Sylvia is his sister and think I am in love with her. She is very sweet and clever.

(Disponível em: momlovesbest.com/short-moral-stories-kids)

LIS13

THE BOY WHO CRIED WOLF

Once, there was a boy who became bored when he watched over the village sheep grazing on the hillside. To entertain himself, he sang out, "Wolf! Wolf! The wolf is chasing the sheep!" When the villagers heard the cry, they came running up the hill to drive the wolf away. But, when they arrived, they saw no wolf. The boy was amused when seeing their angry faces.

"Don't scream wolf, boy," warned the villagers, "when there is no wolf!" They angrily went back down the hill.

Later, the shepherd boy cried out once again, "Wolf! Wolf! The wolf is chasing the sheep!" To his amusement, he looked on as the villagers came running up the hill to scare the wolf away. As they saw there was no wolf, they said strictly, "Save your frightened cry for when there really is a wolf! Don't cry 'wolf' when there is no wolf!" But the boy grinned at their words while they walked grumbling down the hill once more.

Later, the boy saw a real wolf sneaking around his flock. Alarmed, he jumped on his feet and cried out as loud as he could, "Wolf! Wolf!" But the villagers thought he was fooling them again, and so they didn't come to help.

At sunset, the villagers went looking for the boy who hadn't returned with their sheep. When they went up the hill, they found him weeping.

"There really was a wolf here! The flock is gone! I cried out, 'Wolf!' but you didn't come," he wailed.

An old man went to comfort the boy. As he put his arm around him, he said, "Nobody believes a liar, even when he is telling the truth!"

(Disponível em: momlovesbest.com/short-moral-stories-kids)

LIS14

THE GOLDEN TOUCH

There once was a king named Midas who did a good deed for a Satyr. And he was then granted a wish by Dionysus, the god of wine.

For his wish, Midas asked that whatever he touched would turn to gold. Despite Dionysus' efforts to prevent it, Midas pleaded that this was a fantastic wish, and so, it was bestowed.

Excited about his newly-earned powers, Midas started touching all kinds of things, turning each item into pure gold.

But soon, Midas became hungry. As he picked up a piece of food, he found he couldn't eat it. It had turned to gold in his hand.

Hungry, Midas groaned, "I'll starve! Perhaps this was not such an excellent wish after all!"

Seeing his dismay, Midas' beloved daughter threw her arms around him to comfort him, and she, too, turned to gold. "The golden touch is no blessing," Midas cried.

(Disponível em: momlovesbest.com/short-moral-stories-kids)

LIS15

THE FOX AND THE GRAPES

One day, a fox became very hungry as he went to search for some food. He searched high and low, but couldn't find something that he could eat.

Finally, as his stomach rumbled, he stumbled upon a farmer's wall. At the top of the wall, he saw the biggest, juiciest grapes he'd ever seen. They had a rich, purple color, telling the fox they were ready to be eaten.

To reach the grapes, the fox had to jump high in the air. As he jumped, he opened his mouth to catch the grapes, but he missed. The fox tried again but missed yet again.

He tried a few more times but kept failing.

Finally, the fox decided it was time to give up and go home. While he walked away, he muttered, "I'm sure the grapes were sour anyway."

(Disponível em: momlovesbest.com/short-moral-stories-kids)

LIS16

THE PROUD ROSE

Once upon a time, in a desert far away, there was a rose who was so proud of her beautiful looks. Her only complaint was growing next to an ugly cactus.

Every day, the beautiful rose would insult and mock the cactus on his looks, all while the cactus remained quiet. All the other plants nearby tried to make the rose see sense, but she was too swayed by her own looks.

One scorching summer, the desert became dry, and there was no water left for the plants. The rose quickly began to wilt. Her beautiful petals dried up, losing their lush color.

Looking to the cactus, she saw a sparrow dip his beak into the cactus to drink some water. Though ashamed, the rose asked the cactus if she could have some water. The kind cactus readily agreed, helping them both through the tough summer, as friends.

(Disponível em: momlovesbest.com/short-moral-stories-kids)

LIS17

THE MILKMAID AND HER PAIL

One day, Molly the milkmaid had filled her pails with milk. Her job was to milk the cows, and then bring the milk to the market to sell. Molly loved to think about what to spend her money on.

As she filled the pails with milk and went to market, she again thought of all the things she wanted to buy. As she walked along the road, she thought of buying a cake and a basket full of fresh strawberries.

A little further down the road, she spotted a chicken. She thought, "With the money I get from today, I'm going to buy a chicken of my own. That chicken will lay eggs, then I will be able to sell milk and eggs and get more money!"

She continued, "With more money, I will be able to buy a fancy dress and make all the other milkmaids jealous." Out of excitement, Molly started skipping, forgetting about the milk in her pails. Soon, the milk started spilling over the edges, covering Molly.

Drenched, Molly said to herself, "Oh no! I will never have enough money to buy a chicken now." She went home with her empty pails.

"Oh, my goodness! What happened to you?" Molly's mother asked.

"I was too busy dreaming about all the things I wanted to buy that I forgot about the pails," she answered.

"Oh, Molly, my dear. How many times do I need to say, 'Don't count your chickens until they hatch?'"

(Disponível em: momlovesbest.com/short-moral-stories-kids)

LIS18

A WISE OLD OWL

There was an old owl who lived in an oak tree. Every day, he observed incidents that occurred around him.

Yesterday, he watched as a young boy helped an old man carry a heavy basket. Today, he saw a young girl shouting at her mother. The more he saw, the less he spoke.

As the days went on, he spoke less but heard more. The old owl heard people talking and telling stories.

He heard a woman saying an elephant jumped over a fence. He heard a man saying that he had never made a mistake.

The old owl had seen and heard what happened to people. There were some who became better, some who became worse. But the old owl in the tree had become wiser, each and every day.

(Disponível em: momlovesbest.com/short-moral-stories-kids)

LIS19

THE GOLDEN EGG

Once upon a time, a farmer had a goose that laid one golden egg every day. The egg provided enough money for the farmer and his wife to support their daily needs. The farmer and his wife continued to be happy for a long time.

But, one day, the farmer thought to himself, "Why should we take just one egg a day? Why can't we take them all at once and make a lot of money?" The farmer told his wife his idea, and she foolishly agreed.

Then, the next day, as the goose laid its golden egg, the farmer was quick with a sharp knife. He killed the goose and cut its stomach open, in the hopes of finding all its golden eggs. But, as he opened the stomach, the only thing he found was guts and blood.

The farmer quickly realized his foolish mistake and proceeded to cry over his lost resource. As the days went on, the farmer and his wife became poorer and poorer. How jinxed and how foolish they were.

(Disponível em: momlovesbest.com/short-moral-stories-kids)

LIS20

THE FARMER AND THE WELL

One day, a farmer was looking for a water source for his farm, when he bought a well from his neighbor. The neighbor, however, was cunning. The next day, as the farmer came to draw water from his well, the neighbor refused to let him take any water.

When the farmer asked why, the neighbor replied, "I sold you the well, not the water," and walked away. Distraught, the farmer went to the emperor to ask for justice. He explained what had happened.

The emperor called on Birbal, one of his nine, and wisest, courtiers. Birbal proceeded to question the neighbor, "Why don't you let the farmer take water from the well? You did sell the well to the farmer?"

The neighbor replied, "Birbal, I did sell the well to the farmer but not the water within it. He has no right to draw water from the well."

Birbal said, "Look, since you sold the well, you have no right to keep the water in the farmer's well. Either you pay rent to the farmer, or take it out immediately." Realizing that his scheme had failed, the neighbor apologized and went home.

(Disponível em: momlovesbest.com/short-moral-stories-kids)

TEXTOS COMPLEXOS DO CORPUS COMPARATIVO

LIC1

DERIVATION BY PHASE

Chomsky, 1995, Chomsky, 2000, Chomsky, 2001, Chomsky, 2004 Minimalist Program aims to formally specify the human faculty of language as “an optimal solution to minimal design specifications” (Chomsky, 2001, p. 1). For language to be usable, these specifications, he argues, must be determined in respect of two systems that interface with the language system, i.e., the conceptual–intentional (C–I) system for thought and the sensorimotor (SM) system for action. Specifically, in order for the specifications to be fulfilled the information involved in the syntactic structures generated by the computational component must be converted, on the one hand, into the semantic representation that is perfectly legible, i.e., accessible, to the C–I system and, on the other, into the phonetic form representation that is perfectly legible to the SM system.

One of the recent theoretical developments seeking to embed the language system into a human cognitive system concerns the manner in which the syntactic structures constructed in the computational component are spelled out, i.e., transferred, to the semantic component and the PF (Phonetic Form) component. In the preceding model (Chomsky, 1995) the information involved in the syntactic structure is spelled out only once when the structure has been built up completely, while in the recent model (Chomsky, 2001) it may be spelled out by phase on the way of derivation. Specifically, Spell-Out of syntactic structure is made by phase once a specific phase has been reached. Part of the syntactic structure constructed until then – the complement of the head of the phase – is spelled out to the two components and thereby becomes impenetrable to further syntactic operations in the computational component. The phases include a complementiser phrase (CP) and a light verb phrase (VP) involving a causative sense. The derivation-by-phase theory assumes an important role in reducing the “computational burden” (Chomsky, 2001, p. 11), thus firmly embedding the language system well into the human cognitive system and making the former maximally usable by the latter.

(Disponível em: [sciencedirect.com/science/article/abs/pii/S0388000109000138](https://www.sciencedirect.com/science/article/abs/pii/S0388000109000138))

LIC2

Life is improving for managers at the 2,700 stores of Sainsbury, one of the world's largest supermarket groups. A program from PA, a big software company, will make a boring job much simpler: collecting data about each shop's energy consumption, whether from refrigeration, lights or air conditioning. The automated data collection is part of Sainsbury's plan to reduce by 50% emissions of greenhouse gases from existing shops by 2019.

Sainsbury and PA may well be pioneers, but they are not alone. While governments discuss levels of carbon emissions, many companies have already started to make reductions, or are at least preparing to – leading to more and more software firms offering products to help. If predictions are correct the market for carbon-management software could soon become as large as those for other important business applications such as enterprise application software (EAS) programs, a \$7 billion market last year.

Many companies have measured energy consumption for some time in an attempt to reduce running costs. Other firms have tracked emissions of different types in order to comply with pollution regulations. In recent times, public pressure has led to more companies publishing emissions data in their annual reports or to organizations like the Carbon Monitoring Project. However, most firms will need to upgrade from the basic tools, such as spreadsheets, they have been using.

Things are changing, in spite of the recession, says Jim Scarfe, CEO of CarbonReduct, a consultancy. Increased energy costs and new regulations are all pushing companies to monitor their emissions and do so with appropriate software, he states. In the USA, for example, the Carbon Reduction Plan will come into force next year. Among other things, it requires firms that use more than 8,000 megawatt-hours of electricity per annum to evaluate and report the energy they consume.

Expecting an increase in demand, many software-publishers have moved into the market, mostly with internet-based services. In a recent survey SRP Research, another consultancy, listed no fewer than 183 suppliers. Some emphasise reporting, others compliance and still others improving business processes. There are well-established companies, such as EnergySoft and LMG. Many start-ups, such as CarbonModel and GreenData, have appeared. Even Large software firms like Oracle and IBM have also moved into the market.

For the time being, the needs of most firms are simple: making sure that energy data is collected and can be audited. But in the years ahead, this will change, predicts Susanna Sierra of SRP. Companies will need software that collects energy data automatically, while helping them to find the best ways to reduce emissions and allowing them to manage other resources, such as water.

Scarfe and Sierra both expect that Oracle and SAP, which already dominate most types of business software, will control the market in this area, too, because it is a good match for their other products. These giants also have the resources to buy the best technology. In June SAP purchased Green Standards, a start-up. Oracle is thought to be planning a similar purchase soon. But they have other rivals. LMG has been buying companies selling environmental software. Some expect great things from X8, a start-up founded by Jana Novic, who pioneered EAS software.

All this interest gives an idea of how important the business of monitoring environmental performance is likely to become. Scarfe recently suggested that in time it could even be as big a market as financial accounting.

(Disponível em: examenglish.com/cpe/cpe_reading_part7.htm)

LIC3

SECONDARY INFECTION WITH STREPTOCOCCUS PNEUMONIAE DECREASES INFLUENZA VIRUS REPLICATION AND IS LINKED TO SEVERE DISEASE

Acute pneumonia occurs in 30%-40% of influenza A virus (IAV) hospitalizations and can be caused by the virus alone or in conjunction with secondary bacterial infection, most commonly *Staphylococcus aureus* and *Streptococcus pneumoniae*. During the 2009 and 1918 H1N1 pandemics, secondary infection with Spn contributed to severity of disease, especially in young adults. Animal models to study the relationship between IAV and Spn have largely focused on immune responses, disease outcome, and the influence of IAV on Spn colonization. Given the epidemiological importance of viral–bacterial coinfections, there is a critical need for studies that examine how bacterial pathogens influence viral fitness within the host and its transmission to new hosts.

The prevailing model describes a synergism between IAV and Spn with regards to morbidity and mortality. Exacerbated disease during coinfection, compared to a single infection, is supported by epidemiological studies as well as studies in murine and ferret models of respiratory infection. Studies in the murine model suggest that heightened disease results from a combination of dysregulation in immune responses, extensive epithelial damage, release of planktonic bacteria from biofilms, as well as increased bacterial colonization and dissemination. Mechanistic studies indicate that increased bacterial colonization is a consequence of upregulation of bacterial receptors and IAV-induced increases in sialic acid, which serves as a nutrient source for Spn (. Further, enhanced systemic dissemination of bacteria is facilitated by virus-induced cytokine release, oxidative stress, and pneumolysin release. Many of these published studies focused primarily on the colonization, transmission, and pathogenesis of Spn with limited studies on the impact of viral replication and airborne dissemination.

In addition to colonization and disease, a few studies have also investigated the influence of coinfection on contact transmission. These studies, done primarily in infant mice, have shown that coinfection with IAV dramatically enhances Spn transmission with the efficiency of IAV being highly variable. One study further demonstrated that Spn infection preceding IAV inoculation can decrease IAV contact transmission. However, mice in general are suboptimal models for airborne viral transmission.

(Disponível em:

academic.oup.com/femsmicrobes/article/doi/10.1093/femsmc/xtac007/6542440)

LIC4

FLEX YOUR FUNNY BONE

Looking for something to smile about at work? How about this: Flexing your funny bone can significantly enhance your professional prospects. Ninety-one percent of executives surveyed by Robert Half International consider a sense of humor important to career advancement.

Consider the following real-life office pranks from a Robert Half survey that asked executives to describe the funniest April Fools' jokes they have observed in an office environment:

"Someone frosted a cookie with children's toothpaste and offered it to a co-worker." "A couple of employees cleaned out everything from another person's office." "One person's colleagues loaded his office with sand and beach toys." "We put tape underneath a colleague's mouse so he couldn't move it." "We turned everything backward in one person's office -- the computer, desk and pictures." "Someone changed the keys on my keyboard." "We filled a colleague's office with balloons."

Displaying levity on the job can help you build rapport with those around you, facilitate open communication, and contribute to a positive work environment. And, perhaps most importantly, a comic touch can work to relieve tension on even the most stressful days.

But keep in mind that not all high jinks are well received. It's crucial to take into consideration your organization's and co-workers' perspectives when it comes to comic relief. Humor should be work appropriate and never mean spirited or at the expense of others.

Here are some tips to ensure you're not an April Fool:

Say no to sarcasm

People often use humor as an indirect way of berating others. Here's an example: "I can't believe you're here on time -- what's the occasion?" Sarcasm is rarely a good idea, so keep these types of comments to yourself.

Be the butt of your own joke

Go ahead, poke fun at your foibles. Doing so can put others at ease in your presence, and you don't risk offending someone else by making him or her the target of your joke. For example, if you trip while giving a presentation, a comment like, "I hope you're as head over heels about this idea as I am" can help ease any awkwardness. Just be sure to keep your comments light; you don't want your co-workers to think your attempt at humor is a cry for help.

Laugh with others

You can be perceived as having a great sense of humor without ever telling a joke. Just tune in to the humor styles of those around you and share in the fun.

Create a 'funny file'

You have files for various projects or committees you're involved in, so how about developing a "funny file," as well? Create a folder filled with appropriate workplace cartoons, amusing newspaper articles, humorous letters or e-mails from friends, or anything else that tickles your funny bone. The next time one of your co-workers feels overwhelmed or under the weather, you can surprise him or her with a snippet from your file. Just be mindful to avoid items that are offensive or otherwise in poor taste.

Convene a fun committee

Invite co-workers to join in your quest to "up the office fun factor." Together, brainstorm ways to add excitement to the workweek -- surprising co-workers with breakfast or treating the team to an afternoon at the park, for instance. Just be sure to consult your department head before executing any plans; you'll want to obtain his or her approval and determine other details, such as budget and scheduling.

Issue trivia quizzes

What did Tom Cruise and Katie Holmes name their baby? From which state did the last "American Idol" hail? Most people enjoy keeping up with pop culture, so why not create a friendly competition around it? You also can grill people on sports, geography or food -- whatever topics the team finds interesting. Even if the only prize for answering the most questions correctly is posting the name of the winner in a prominent spot, the joy of conjuring up random information and discussing the "stumpers" can increase the general playfulness of any work area.

Capture Kodak moments

Keep a disposable camera on hand for all to use to capture those moments when you and your colleagues are at your best -- or worst. Then, post the pictures on a community bulletin board. A candid snapshot from the day everyone unintentionally wore lime-green shirts, for example, is bound to brighten the mood.

A culture of fun at work can improve communication, reduce stress and increase productivity. So don't be afraid to flex your funny bone -- just be sure to do so in a business-appropriate way.

(Disponível em: englishuniversity.eu/?L=-8&E=1&S=70)

LIC5

CHARACTERIZATION OF SYNTHESIZED NANOPARTICLES

UV-1800 Shimadzu spectrophotometer was used to obtain UV-visible spectra of AgNPs at different time intervals using seed extract diluted with water (1:9) as a blank. Further shape and size of formed nanoparticles were determined by SEM, FESEM equipped with EDS, and TEM. A lyophilized sample of AgNPs was subjected to Zeiss EVO-18 scanning electron microscope at 20 kV to study the morphological features of silver nanoparticles. Further, the shape, morphology, and elemental mapping of AgNPs were studied using field emission scanning electron microscopy (NOVA NanoSEM 450). For this purpose, the lyophilized sample was sonicated for a sufficient amount of time, the smear was made on a platinum grid, and allowed to dry overnight under vacuum. The grid was then coated with a thin film of palladium and finally subjected to FESEM. Transmission electron microscopy was done to exactly determine the size of nanoparticles. Sonicated sample was loaded on a carbon-coated copper grid and was allowed to dry overnight in a vacuum and subjected to transmission electron microscopy (FEI-TECNAI G²20 TWIN). The crystalline nature of AgNPs was confirmed by XRD pattern obtained from Rigaku-MiniFlex 600 X-ray diffractometer at 2θ range from 0 to 100°. The sample for XRD measurement was prepared by casting the powder of silver nanoparticles on a glass slide and subsequently air-drying it under ambient conditions. The pattern was recorded by CuK α radiation with λ of 1.5406 Å at a voltage of 40 kV and current of 15 mA with a scan rate of 10°/min. Presence of functional groups present in AgNPs was determined by using FTIR (ALPHA BRUKER Eco-ATR). Four per centimeter resolution was taken in all spectra with 500–4000 cm⁻¹ IR range.

(Disponível em: jast-journal.springeropen.com/articles/10.1186/s40543-018-0163-z)

LIC6

MEDICATION USE AND RISK OF AMYOTROPHIC LATERAL SCLEROSIS—A SYSTEMATIC REVIEW

This systematic review summarized all available studies assessing the link between the use of seven classes of medications, including cholesterol lowering drugs, anti-inflammatory drugs, immunosuppressants, and antibiotics, oral contraceptives (OCs) or hormone replacement therapy (HRT), antihypertensive drugs, antidiabetics, proton pump inhibitors, and drugs for psychiatric and neurological disorders and risk of ALS. In brief, we found that there was generally a lack of agreement on the associations of these medications with the subsequent risk of ALS between studies. However, it appeared that statins, aspirin, OCs/HRT, antihypertensives, and antidiabetics were unlikely related to a higher risk of ALS. The positive associations suggested for the use of antibiotics, antidepressants and skeletal muscle relaxants might be attributable to prodromal symptoms of ALS.

An inverse association of cholesterol lowering drugs with ALS risk was found in four studies, while a null or marginal association was observed in four. The marginal association was likely due to the higher risk of ALS during the first year after statin use. The differing results between studies might be attributable to diverse study designs, sample sizes, variable adjustments, or biological difference between populations. Although still controversial, some studies have suggested that hyperlipidemia may be a risk factor ALS. If this is the case, the null or inverse association noted between cholesterol lowering drugs, especially statins, and ALS risk might indicate a potentially protective role of statins on ALS. After considering indications for statin use, Freedman et al. indeed found a lower ALS risk in relation to statin use, but not other cholesterol-lowering drugs. The potentially protective role of statins may be attributed to their antioxidative and anti-inflammatory properties.

Steroids were not associated with the risk of ALS in the present study. Some experimental studies have shown that steroids may exhibit a neuroprotective role on neurons, glial cells, and blood vessels via steroid receptor signaling, whereas others found a detrimental effect of steroids on motor neurons. In terms of NSAIDs, most previous studies failed to show an association with ALS risk, except for one. The discrepancy between studies might be partly due to sample size, recall bias with self-reported data, not considering reverse causation, and lack of correction for multiple testing. Real population differences might also exist. For instance, two US studies did not find an association between aspirin use and ALS, whereas two Taiwanese studies found aspirin use to be associated with a lower risk of ALS. A potentially protective role of aspirin on ALS has been identified in an experimental study which suggested that a derivative of aspirin “AAD-2004” might decrease motor neuron degeneration and

improve motor function and life span by blocking free radical production, prostaglandin E formation, and microglial activation in the spinal cord of ALS mice.

The findings on immunosuppressants are inconsistent. Further, immunosuppressive treatment was not shown to be related to disease progression and survival of ALS patients. Emerging evidence suggests that gut microbiota dysbiosis might influence the development and disease progression of ALS through the microbiota-gut-brain axis. The relationship between antibiotics use and ALS risk, however, is less studied in human. One study showed a positive association, which was mainly attributable to the use of antibiotics during the years before ALS diagnosis.

(Disponível em: [bmcmedicine.biomedcentral.com/articles/10.1186/s12916-022-02442-2](https://www.bmcmedicine.biomedcentral.com/articles/10.1186/s12916-022-02442-2))

LIC7

RAMAN MICROSPECTROSCOPY FINGERPRINTING OF ORGANOID DIFFERENTIATION STATE

Raman imaging of organoids is complicated by a pervasive Raman signal from culture surfaces and ECM or other scaffolding materials. Here, a method employing confocal Raman microspectroscopy was developed to capture only Raman photons originating from the cell-dense, primarily epithelial, regions of the organoids and not from cells or proteins in the surrounding matrix. These studies were first performed in hydrated, intact, fixed organoids to determine the appropriate amount of sampling required to classify the organoids. To eliminate signals from plastic culture dishes, the unlabeled organoids were transferred to a quartz-based chamber, and the organoids remained submerged in 1× PBS for imaging. The average area of the fixed organoid was determined while in the culture dish, and it was shown to be larger for the growth factor-treated organoids compared with untreated organoids. Upon transfer to the quartz-based chamber, the cell-dense regions, representing regions composed primarily of epithelial cells, were identifiable and distinct from the Matrigel-dense regions, representing stromal cell-rich regions. After the X–Y coordinates of the cell-dense regions were identified using brightfield imaging, a series of Raman confocal measurements along the Z-direction was performed to focus on the cell-dense volume and exclude the spectra from the highly fluorescent Matrigel. The laser and acquisition settings were selected to facilitate the easy separation of Matrigel-rich regions from cell-dense regions.

(Disponível em: cmbl.biomedcentral.com/articles/10.1186/s11658-022-00347-3)

LIC8

EXTRACELLULAR MEMBRANE VESICLES AND NANOTUBES IN ARCHAEA

Archaea have been recognized as a separate domain of life, besides Bacteria and Eukarya, only in 1977 by Carl Woese and colleagues who compared the ribosomal RNA (rRNA) gene sequences from diverse organisms. The name Archaea, instead of archaeobacteria, was proposed later on by Carl Woese to emphasize the fact that Archaea and Bacteria form two distinct lineages in the universal tree of life. Although archaeal cells are of the prokaryotic type, their informational systems (DNA replication, transcription, translation), as well as several membrane-associated machineries, such as the ATP synthase complex, the Sec secretion system and the signal recognition particles, are much more similar to those of eukaryotes. Similar to bacteria and eukaryotes, archaeal cells commonly secrete extracellular vesicles (EVs) and some produce tubular structures resembling bacterial nanopods and/or nanotubes. The production of various types of EVs (apoptotic bodies, exosomes, microvesicles, etc.) and nanotube-like structures (tunnelling nanotubes, etc.) has been extensively studied in Eukarya. In Bacteria, EVs were observed for the first time by electron microscopy in *E. coli* in 1966, but their biological importance was first dismissed and, hence, microbial EVs have been considered as artefacts of cell growth or lysis for many years. However, bacterial EVs are increasingly recognized to play important roles in many processes from pathogenesis, bacterial communication and biofilm formation, to horizontal gene transfer and protection against viral infections. Nanotubes have been described in Bacteria more recently, but their physiological role remains controversial. In Archaea, EVs have been first described over two decades ago when they were found to carry protein toxins. Although until recently, archaeal EVs received relatively little attention, perhaps due to the fact that their discovery was from the very beginning linked to a defined function (i.e. toxin transfer), archaeal EVs were not dismissed by the archaeal community as cellular 'junk,' as in some other branches of microbiology. The research on archaeal EVs has primarily focused on two orders of hyperthermophilic species, Sulfolobales and Thermococcales, whereas nanotubes have been primarily described in Thermococcales and Haloarchaea.

(Disponível em: academic.oup.com/microlife/article/doi/10.1093/femsmi/uqab007/6296418)

LIC9

A BRAIN MECHANISM UNDERLYING THE EVOLUTION OF ANXIETY

New research using genome editing technology has allowed scientists to create a model and assess a gene mutation associated with neuropsychiatric disorders in humans. The study has revealed how the mutation functions in the brain and affects anxiety and sociality.

Monoamine neurotransmitters such as serotonin and dopamine play important roles in our cognitive and emotional functions. Their evolutionary origins date back to metazoans, and while the function of related genes is strongly evolutionarily conserved, genetic variation within and between species has been reported to have a significant impact on animal mental characteristics such as sociality, aggression, anxiety, and depression.

A research group led by Dr Daiki Sato and Professor Masakado Kawata has previously reported that the vesicular monoamine transporter 1 (VMAT1) gene, which transports neurotransmitters to secretory vesicles in neurons and secretory cells, has evolved through natural selection during human evolution. In particular, the 136th amino acid locus of this gene has evolved in the human lineage from asparagine (Asn) to threonine (Thr), and moreover, a new allele (isoleucine, Ile) has emerged and increased in its frequencies around the world. Previous reports suggested that people with the Ile genotype are less prone to depression and anxiety than those with the Thr genotype, but it was unclear how these human-specific mutations function in the brain and lead to changes in neuropsychiatric behavior.

(Disponível em: [sciencedaily.com/releases/2022/08/220805091](https://www.sciencedaily.com/releases/2022/08/220805091))

LIC10

AIRCRAFT MAINTENANCE CHECK SCHEDULING USING REINFORCEMENT LEARNING

This paper proposed a reinforcement learning solution to optimize the long-term scheduling of A and C-checks for an aircraft fleet. The goal was to maximize the FH usage between checks, which allows for an increase in aircraft availability in the long run. The proposed approach uses a Deep Q-learning algorithm with experience replay and the Double Q-learning variant. At each decision step, the RL agent selects an aircraft to have its next check scheduled on the best available day, prioritizing the scheduling of C-checks.

This approach is evaluated in a real scenario using data from a fleet of 45 aircraft from three sub-fleets (A319, A320, and A321). A comparison with airline estimations for the same period shows that the RL solution produces more efficient maintenance plans. The higher fleet usage for both A and C-checks results in a lower amount of checks being scheduled during the entire time horizon, which has significant financial benefits for the airline. The RL results are also promising when compared to the DP approach, achieving better results in the C-check scheduling.

A second test case is defined to evaluate the ability of RL in dealing with different initial conditions of the fleet. Multiple scenarios are simulated by introducing disturbances in the initial usage metrics of each aircraft. These scenarios are used to train the RL agent with different fleet conditions. The results obtained with several testing scenarios show that this approach can generate quality maintenance plans in a short amount of time.

There are not many studies regarding long-term maintenance problems in aviation despite its great relevance, which encourages future research on this topic. A potential opportunity is to extend the check scheduling algorithm by defining the set of maintenance tasks to be performed in each one of them. This task allocation problem would produce an optimal maintenance plan with a combination of maintenance checks and tasks.

(Disponível em: mdpi.com/2226-4310/8/4/113)

LIC11

MY MOST INSPIRING TEACHER

Hadley Freeman

I met Charlie Ritchie when my mother and I travelled up to Cambridge together in 1998, looking for a boarding school for me. He was enthusiastic, laidback, inspiring, friendly, and, as a bonus, he had a Scottish accent that I thought exotic and my mother found impenetrable.

I was an obsessive student, channeling my anxieties into my studies. But whereas some teachers would encourage that, Charlie wasn't having any of it. He encouraged me to think of my studies as fun, a heretofore unconsidered concept.

But more than that, he urged me to do more than just study. He would casually let me know when a movie or play he thought I'd like was on in town, and he suggested I start writing for myself, not just to do well in my A levels. In short, he told me to stop taking everything so seriously while also telling me to have more confidence in myself. 20 years on, I still remember his lessons. Only the best kind of teacher can get you to do that.

Lola Okolosie

Most of us will remember one teacher in particular.

We carry vivid images of their gait and can recall the way in which they held a room. At their very best, they can be the most captivating of professionals. For me, it was Mrs. Jones, my English literature teacher, who had a huge impact on my school life.

I was all set for studying sociology, but Mrs. Jones' lessons made me reconsider. There was her glamour and sense of style. With the red lipstick and chunky jewellery, she stood out, and made learning English cool. Mrs. Jones inspired me because she resisted treating us like children.

At times it felt a little bruising, but then the knowledge that she expected the absolute best from you made you believe you were, in some way, capable of it. By the time I began my undergraduate degree it felt like an anti-climax, we seemed to be covering old ground. She collapsed the sense of distance that shrouds literary giants, such as Shakespeare and Keats, and made them seem as much ours as anyone else's.

Decca Aitkenhead

Andy Pickering taught me a skill I would use every single day of my life - how to approach an essay not as an opportunity to demonstrate how many facts I'd managed to memorize, but as an exercise in presenting competing arguments. He also taught me to identify salient facts and deploy them to maximum effect to justify an argument. It was a total revelation.

At university I discovered that this was a secret very few had been taught. When I landed an internship at my first newspaper, the editor asked me to create a new daily feature - The Big Debate - presenting two contradictory arguments about issues of the day, such as euthanasia or foxhunting and so on. 'But this is easy!' I remember thinking. 'It's just what Andy taught me!' It was on the strength of The Big Debate's success that my career in journalism began.

I have been a journalist for more than 20 years now, and can still detect Andy's teaching in every single column, feature, and interview I have ever written.

Paula Coccozza

He would enter the room swinging a briefcase, which he often slammed on the desk with a moody thud. The class bristled with expectation and dread. You never knew what you were going to get with Mr. Hartley. He demanded concentration. Once, he stood on his desk and watched us file in, remaining statuesque long after we took our seats. Another time, he sang Elvis.

In his classes we hovered on the edge of adulthood.

One lesson was spent investigating how we knew we existed. The question seemed to us idiotic. But Mr. Hartley rebuffed all our proofs and when the bell rang, we trudged out wearily, as if we had fought for our lives. That was our introduction to Descartes.

Most of all, he gave me a love of books. I carried his reading list for years from the age of 12. I would stop at the library on my way home. I was spending more time alone with a book, but I felt so much less alone. My family had always veered towards work rather than education, and as my GCSEs approached, I began to consider which apprenticeship to apply for. But books had taken hold and I enrolled instead for A levels at the local sixth form. Mr. Hartley unlocked worlds for me. I sidestepped the future.

(Fonte: *English File Advanced Plus* – Página 15)

LIC12

MEASURES TO PREVENT MOSQUITO BITES

As malaria prevention in travellers becomes reliant on bite avoidance, much greater attention to the correct use of insect repellents is required. It is concerning to note that in a study published online recently, only 2.5% of travellers applied the recommended dose of repellent. DEET is a common insect repellent and is available in a variety of concentrations. A 50% concentration is the most effective. Lower concentrations are available, but they may not be as effective as 50%. A variety of studies has concluded that there is a low risk of adverse effects when DEET is applied according to product directions. DEET can damage some plastic watch straps, watches, and jewellery; these items should not be allowed to come into contact with DEET. The user should ensure that repellents are not ingested or inhaled and do not come into contact with their eyes or mouth.

If sleeping outdoors or in unscreened accommodation, insecticide-treated mosquito nets should be used. Protective efficacy against malaria for travellers has been estimated at 50%. Mosquito bed nets must be free of holes and should be tucked in under the mattress. Insecticide-impregnated bed nets improve protection because they help to prevent (a) mosquitoes biting through the net on parts of the body touching the net, and (b) mosquitoes surviving long enough near a net to find any holes in the net which may exist. Most of the nets now available are long-lasting impregnated nets. In these products the insecticide is incorporated into the material of the net itself. They have an expected useful life of at least 3 years. If using standard (i.e. not long-lasting) nets, these will need to be re-impregnated every 6 to 12 months (depending on how frequently the net is washed) to remain effective. If a traveller purchases a standard impregnated net, the 6 months starts from the date when it starts to be used and washed, as washing and handling are the main factors removing the insecticide.

Within the limits of practicality, cover up with loose-fitting clothing, long sleeves, long trousers, and socks if out of doors after sunset, to minimize accessibility to skin for biting mosquitoes. There is no evidence that the colour of clothing is relevant to mosquitoes.

Clothing may be sprayed or impregnated with an insecticide, or purchased pre-treated to reduce biting through the clothing. As an alternative, cotton clothing (for example socks) can be sprayed with DEET. DEET is useful as a clothing repellent. DEET preparations for use only on clothing are commercially available.

Air conditioning reduces the likelihood of mosquito bites because of substantial reduction in nighttime temperature. Ceiling fans reduce mosquito nuisance. Doors, windows, and other possible mosquito entry routes to sleeping accommodation should be screened with fine mesh netting which must be close-fitting. The room should be sprayed before sunset with an insecticide to kill any mosquitoes which may have entered the accommodation during the day. During the night, where electricity is available, use an electrically-heated device to vaporize a tablet containing an insecticide in the room. A new tablet is needed each night.

We strongly advise against relying on any herbal remedies for the prevention of malaria.

Herbal remedies have not been tested for their ability to prevent or treat malaria.

We also strongly advise against relying on any homeopathic remedies for the prevention of malaria. There is no scientific proof that homeopathic remedies are effective in either preventing or treating malaria. In addition, the Faculty of Homeopathy does not promote the use of homeopathic remedies for malaria prevention.

Electronic buzzers (emitting high frequency sound waves) are completely ineffective as mosquito repellents. Companies selling them have been prosecuted and fined under the UK Trades Descriptions Act and our advice is that they should not be used. There is no evidence that vitamin B1 or vitamin B12 taken orally repel mosquitoes. There is also no evidence that garlic taken orally repels mosquitoes, nor that tea tree oil is an effective mosquito repellent.

(Fonte: *English File Advanced Plus* – Página 25)

LIC 13

There are some things we may not be looking forward to as we get older, but research is increasingly showing that many of our mental and physical peaks occur well after our youth. There's a best age to do everything, from resolving conflicts to running marathons...

Be happy

Studies have proved happiness isn't necessarily for the young, and definitely not for the middle-aged, who always come out of the data the most miserable. A study at the London School of Economics using German data found that people aged 69 were at the peak of well-being, and there was also a smaller peak at 23. Those aged 45-54 were at the lowest point. We usually see happiness develop over life as a U-curve, says Meik Wiking, the chief executive of the Happiness Research Institute. 'One theory is that we become better at prioritizing what matters most, choosing what will improve our happiness.

Have a baby

Biologically speaking, our grandmothers were right: the time to have babies is from the late teens to early twenties. The decline in fertility starts from about 32, and this accelerates rapidly from 37. However, John Mirowsky, at the University of Texas, spent years studying the data, and found that women who felt fittest and most energetic in middle and old age had their first babies at 29 or 30, and those with the lowest rates of chronic disease and pain in later life had them at 34, as did those who lived the longest. That gives rise to a best age of 31, he concluded.

Run a marathon

You may expect it to be younger because aerobic fitness as measured by oxygen uptake peaks between 23 and 25. That's the best age range for shorter races, but marathon-running requires more skills. 'It's a combination of strength, speed, endurance and the actual experience of running marathons, says John Brewer, professor of applied sports science and veteran of 19 London marathons. 'You rarely run your best marathon first - there's a learning effect of running 26.2 miles - the build-up, diet, hydration, pacing and mental attitude. A Spanish study of New York marathon times shows that 18-year-olds take as long to finish as 55-year-olds, while the best times for men were achieved at 27, and for women, 29.

Do your best work

Einstein claimed that if you hadn't made a great contribution to science before 30, you never would, but research shows that 40 is now the age at which most Nobel prize winning work is

conducted. That also holds true for lesser mortals with more mundane jobs, concluded the authors at the US National Bureau of Economic Research. Another study seems to back that up: researchers in Canada conclude that the best-quality academic papers are written 12 years into a person's career, and after that, quality and quantity tended to decline.

Learn to drive

You're more likely to pass your test if you take it aged 17, according to government figures. That's when pass rates are the highest: 58 per cent for boys and 53 per cent for girls. Pass rates then fall steadily throughout life to about 35 per cent in your fifties. 'It's probably because 17-year-olds are used to learning and their brains are open to new skills - I've watched 26-year-olds learning and they definitely find it more difficult, possibly because they are more scared,' says Edmund King, the president of the AA. However, whether 17 is the safest age is another question. Young drivers (aged 17 to 24) make up 18 per cent of drivers involved in road accidents, despite doing only 5 per cent of miles.

Do maths in your head

You may expect younger brains with their superior processing speeds to be best at arithmetic, but a study of 48,000 people's cognitive powers across the ages in 2015 revealed the peak for mental maths was actually 50. Why? Possibly because older adults have more experience, so they use a better mix of strategies to work out sums, to compensate for their slower speed. However, Professor Stuart-Hamilton, at the University of South Wales, suspects it might be a generational thing. 'I'm 58 and I was drilled in mental arithmetic at school, and I've had students look at me like I'm a master of the dark arts because I can multiply 56 by 24 in my head, he says.

(Fonte: *English File Advanced Plus* – Página 26)

*LIC 14***HOW PEOPLE IN HIGH-PRESSURE JOBS COPE WITH STRESS...***Clara Amfo*

I feel stressed when I feel unprepared. I like to research things meticulously, so if someone is coming on my radio show, or I'm hosting something on TV, I need to have all the material as soon as possible. If I don't, I feel very anxious. When I'm stressed for longer periods, I stutter a lot, my skin goes to hell and my short-term memory becomes laughable.

When it comes to countering stress, I've realized the necessity of physical activity. I really feel the difference in my mood when I take the time to go boxing, do strength training, or just grab my trainers for a run. I started running in 2015 because I had FOMO from my mates who were in a running club. I hated it at first, then after training for my first half-marathon, I began depending on it as my stress relief. I've also become much stricter with my bedtime. Getting eight hours definitely makes me less anxious throughout the day. Corny as it sounds, the best advice I'd give is the same advice friends have given to me - 'this too shall pass. Feeling stressed and anxious can be frustrating and crippling, but it can't last for ever.

Henry Marsh

As a neurosurgeon, I was in a state of mild chronic anxiety all the time. Surgery is intensely stressful - not so much the actual operating as having to deal with patients and their families afterwards, when things have gone badly. I never resented the anxiety, as I never doubted the value and importance of my work, and it seemed (and still seems) an appropriate price to pay for the privilege of being a surgeon.

It helped to remind myself that I had been through similar crises before and survived, even if the poor patient hadn't. Physical exercise has always been extremely important for me; there are few problems that are not helped by a long walk or run, and I still try to run every day. I also find that writing, in particular, keeping a daily diary, can be very therapeutic.

Talking to colleagues can be helpful as well; only they can properly understand what it feels like when a patient comes to harm. Over time, I got better at being open about my problems. It is only human to make mistakes. My advice? Ask for help. Have good colleagues. Be a good colleague in return.

James Haskell

I try to manage pre-match performance anxiety with visualizations. If I'm worried about a specific tackle, I imagine it going amazingly, and really feel every sensation of that great tackle. It makes a big difference.

I've learned through working with a sports psychologist that music is also a really powerful tool. I use carefully chosen playlists to regulate my mood - whether that's to try and pump myself up and feel like a million dollars before a match, or to try and wind down when I'm experiencing stress relating to injury. Being injured is a very specific kind of stress. One of the most anxious periods of my career was in 2017, when I had toe surgery and wasn't sure I'd ever be able to play rugby again.

I run a couple of businesses outside of rugby, and time management can be another source of stress. When I have a million things going through my head, writing notes of the most important things can help stop my mind from racing. An important lesson has been remembering to only worry about what you can control. You can't control other people, or situations that are out of your hands; you can only control your response.

Angela Hartnett

When I'm feeling stressed in the kitchen, I can get quite quiet and irritable. If food isn't going out quickly enough, or something goes wrong, I've learned to stop and take five minutes. It's important to address smaller issues as they arise; dealing with the situation as calmly as possible means that it doesn't spiral out of control.

When it comes to bigger problems, you shouldn't jump in straight away. If the situation allows, then sleep on it, or think about what's bothering you. Things can look easier the next day. Emails really up my anxiety levels, so I always pick up the phone or speak to someone face to face if I can.

Years ago, when I was at the Connaught, stressed, and quite new to being a head chef, Richard Corrigan told me one of the most important things is to get to know your team. They might travel for two hours each day, or have other things going on outside of work, and as an employer, being approachable and understanding the team is key. At the same time, I've learned to try and guide people, but also to let them deal with things themselves - I don't want to be like a kindergarten teacher. That balance is important.

(Fonte: *English File Advanced Plus* – Páginas 48 e 49)

LIC 15

THE DESCENT OF MAN, GRAYSON PERRY

Grayson Perry has been thinking about masculinity - what it is, how it operates, why little boys are thought to be made of slugs and snails - since he was a boy. Now, in this funny and necessary book, he turns round to look at men with a clear eye and ask, what sort of men would make the world a better place, for everyone?

What would happen if we rethought the old, macho, outdated version of manhood, and embraced a different idea of what makes a man? Apart from giving up the coronary-inducing stress of always being 'right' and the vast new wardrobe options, the real benefit might be that a newly fitted masculinity will allow men to have better relationships and that's happiness, right?

Grayson Perry admits he's not immune from the stereotypes himself... and his thoughts on everything from power to physical appearance, from emotions to a brand new Manifesto for Men, are shot through with honesty, tenderness and the belief that, for everyone to benefit, upgrading masculinity has to be something men decide to do themselves. They have nothing to lose but their hang-ups.

How to Win Friends and Influence People, Dale Carnegie

The most famous confidence-boosting book ever published; with sales of over 16 million copies worldwide... In *How to Win Friends and Influence People* Carnegie offers practical advice and techniques, in his exuberant and conversational style, for how to get out of a mental rut and make life more rewarding. His advice has stood the test of time and will teach you how to: make friends quickly and easily; increase your popularity; win people to your way of thinking; enable you to win new clients and customers; become a better speaker and a more entertaining conversationalist; arouse enthusiasm among your colleagues. This book will turn your relationships around and improve your dealings with all the people in your life. Dale Carnegie, known as 'the arch-priest of the art of making friends, pioneered the development of personal business skills, self-confidence and motivational techniques. His books have sold tens of millions worldwide and, even in today's changing climate, they remain as popular as ever.

Surrounded by Idiots: The Four Types of Human Behaviour, Thomas Erikson

Do you ever think you're the only one making any sense? Or tried to reason with your partner with disastrous results? Do long, rambling answers drive you crazy? Or does your colleague's abrasive manner get your back up? You are not alone. After a disastrous meeting with a highly successful entrepreneur, who was genuinely convinced he was 'surrounded by idiots',

communication expert and bestselling author, Thomas Erikson dedicated himself to understanding how people function and why we often struggle to connect with certain types of people. Erikson's *Surrounded by Idiots* is already an international phenomenon... It offers a simple, yet ground-breaking method for assessing the personalities of people we communicate with... Erikson will help you hone... communication and social skills, handle conflict with confidence, improve dynamics with your boss and team, and get the best out of the people you deal with and manage. He also shares simple tricks on body language, improving written communication and advice on when to back away or when to push on, and when to speak up or indeed shut up. *Surrounded by Idiots* will help you understand and influence those around you, even people you currently think are beyond all comprehension.

The Power of Habit: Why We Do What We Do, and How to Change, Charles Duhigg

...Why do we do develop habits? And how can we change them? ...In *The Power of Habit*, award-winning New York Times business reporter Charles Duhigg translates cutting-edge behavioural science into practical self-improvement action, distilling advanced neuroscience into fascinating narratives of transformation. Why can some people and companies change overnight, and some stay stuck in their old ruts? The answer lies deep in the human brain, and *The Power of Habit* reveals the secret pressure points that can change a life. From Olympic swimmer Michael Phelps to Martin Luther King Jr., from the CEO of Starbucks to the locker rooms of the NFL, Duhigg explores the incredible results of keystone habits, and how they can make all the difference between billions and millions, failure and success - or even life and death. *The Power of Habit* makes an exhilarating case: the key to almost any door in life is instilling the right habit. From exercise to weight loss, child-rearing to productivity, market disruption to social revolution, and above all success, the right habits can change everything. Habits aren't destiny. They're science, one which can transform our businesses, our communities, and our lives.

Why We Sleep: The New Science of Sleep and Dreams, Matthew Walker

Sleep is one of the most important aspects of our life, health and longevity and yet it is increasingly neglected in twenty-first-century society, with devastating consequences: every major disease in the developed world - Alzheimer's, cancer, obesity, diabetes - has very strong causal links to deficient sleep. Until very recently, science had no answer to the question of why we sleep, or what good it served, or why its absence is so damaging to our health. Compared to the other basic drives in life - eating, drinking, and reproducing - the purpose of sleep remained elusive. Now, in this book, the first of its kind written by a scientific expert, Professor Matthew Walker explores twenty years of cutting-edge research to solve the mystery of why sleep matters. Looking at creatures from across the animal kingdom as well as major

human studies, *Why We Sleep* delves into everything from what really happens during REM sleep to how caffeine and alcohol affect sleep and why our sleep patterns change across a lifetime, transforming our appreciation of the extraordinary phenomenon that safeguards our existence.

(Fonte: *English File Advanced Plus* – Página 55)

LIC 16

A MILLENNIAL GOES BACK TO THE EIGHTIES

Born in 1983, I just scrape into the category “digital native”. By around 14, I was using email and instant messengers; by 18, a mobile phone; by 23, Facebook. Within a year or two of starting work in London, I had apps in my pocket that presented me with a door-to-door route to wherever I needed to go. Innovations have flowed steadily into my life, making everyday tasks simpler, quicker. The problem is that they also make my life manic. Dead time used to be built into our lives: waiting for a letter to reach somebody and their response to reach you, for example. With modern tech, if you haven't completed five tasks in the next five minutes, well, what the hell have you been doing? It's probably why I've been a bit overwrought lately - my eyelid twitching, my brain waking me up an hour after bedtime. So I asked myself this: would I have been a more relaxed woman in the eighties? I decided to find out.

'You can only use cash!', 'No more Uber!', 'Get a Filofax!' they chorused. I walked to a stationer's near the office where there was a large and dusty Filofax display. All the organizers were firmly attached to the display. 'Do people steal them?' I asked. 'You wouldn't believe how often,' replied the assistant, grimly. Back at the office, I scrawled down the names and numbers of everyone I might need to contact, and copied out the next week of my calendar from my phone. That night I met my Luddite cousin Sarah for a drink. As it happens, she has barely progressed past the technology of 1988 anyway, and had arrived at the pub by writing its name on a scrap of paper and asking strangers where it might be. She owed me £60, and rather than transfer it to me, had brought a cheque. This would be my new life.

I delayed leaving. I double-checked I had all the right numbers written down. I packed my A-Z and checked my email as many times as I could before actually having to walk out of the door. I circulated the number of my desk phone to my parents, my sister and my flatmate, in case of emergency. I set up an email auto-reply notifying anyone who wrote to me that I wouldn't be online, and so if they needed something urgently, they should call. Then I set off, feeling both weighed down (map, Filofax, printed newspaper - The Times app not being an option) and strangely light.

So I had to consult the A-Z. This is part of the price we pay for being a digital native: because we know that we can google it, our brains don't bother to remember it. At work, a colleague asked me to write a paragraph for a newsletter, so I dictated it to her. We both enjoyed the novelty, but there's no doubt it was an utter waste of time. Then I persuaded a friend at work to go out for lunch with me, but I didn't have the phone number to book a table at the restaurant, and was not allowed to google it. I tried ringing a series of numbers, half-remembered by colleagues, for directory inquiries. After 25 fruitless minutes, I asked around the office to see if anyone had an old-fashioned phone book, and in return I received nothing helpful whatsoever.

I snapped, and googled the phone number. Lunch itself was a delight. Once I accepted that I couldn't check my phone to see if any disasters were unfolding at the office, it was blissfully relaxing.

It was a sorely needed reminder that it isn't a world-shattering catastrophe to waste a bit of your time. But more surprisingly, I realized that my ever-vigilant email habits make me feel like I'm being chased. Maybe that accounts for the fight-or-flight adrenaline rushes that had been waking me up at antisocial hours. Ben Carter, a senior lecturer at King's, told me an anecdote; a colleague had sent him an email, and on not receiving a response in the first half hour, pointedly walked past his office door several times. '20 years ago, he wouldn't have expected a reply in writing within two weeks,' said Carter. 'As a society, we've got ourselves into the situation where we think we should respond to something immediately.'

While I was off-radar, I had received more than 100 emails each day, which meant that a lot of people got my out-of-office reply to 'Call me if it's urgent'. Yet I didn't get a single phone call. So much shouting in my inbox, and not one urgent matter. I have now reverted to Google, and contactless payments, not cash. But I am logging out of my email when I leave the office, and creating more dead minutes in which I can quietly zone out. It's left me with more time to ponder the big question, too: why on earth would someone want to steal a Filofax?

(Fonte: *English File Advanced Plus* – Páginas 56 e 57)

*LIC 17***OPINIONS ON RENZO PIANO'S SHARD, LONDON**

Amanda Levette, Architect

In many ways the Shard is a one-liner - but what a line! Its height and relationship to the city is daring and awe-inspiring, respectful of both historical neighbours and contemporaries without yielding its confidence and command of the cityscape. One of my favourite views is from the top of Parliament Hill, where the Shard stands like a proud grandfather overlooking an infant St Paul's Cathedral nestling below. In what other city, with a history dating back to Roman times, could one of the world's most advanced skyscrapers be built on top of one of the world's oldest railway stations? It's this juxtaposition and tension between tradition and progress that gives London its edge and life. The prowess and technical feats of Victorian infrastructure and super-modern tower are beautifully expressed side by side. It would have been more perfect if the planners had not reduced the height by 300 feet. Having bravely agreed to a tower so high, why take off the top? Whose views would have been compromised 1,300 feet up in the air? London has never really been comfortable with tall buildings, but hopefully acclaim for the Shard will change that.

Simon Jenkins, Writer and Journalist

This tower is anarchy. It conforms to no planning policy. It marks no architectural focus. It offers no civic forum or function, just luxury flats and hotels. It stands apart from the City cluster and pays no heed to its surrounding context in scale, materials or ground presence. The Shard was furiously opposed by local people, by Southwark council and by historic buildings and conservation authorities, but it was pushed as a symbol of Britain's love affair with financial bling at the turn of the 21st century. There is no case for buildings like this on grounds of urban density. Their space ratios make them costly and inefficient to service. Any Londoner knows there are thousands of acres of unused and underused land within the M25 awaiting the high-density, low-rise building preferred by the property market. Some people find the Shard beautiful. I am sure I would in the Gulf, as I admire the Burj Khalifa. But London is not Dubai. Nor is this just a matter of one person's opinion against another's. It is the destruction of one for the other's gain. There are plenty of places for architects to play their games. Why must they tip paint over my Canaletto?

Charles Jencks, Architectural Historian

Renzo Piano says he was inspired by 'London spires' but this is a reminder that, as built, the Shard is a very secular structure, and as for his 'sailing masts', they also underline the missing reality, further boat imagery and curves of the first design. The 'shards' are neither the

potsherds of the name, nor the slivers of ice, stone and glass of the main metaphor. Instead they are large, flat, industrial window-walls that taper - long, thin "sky-wedges of see-through rectangles' - that may create a pleasing silhouette, depending on the view, and weather. The 'pyramid and obelisk', also invoked by critics, are traditionally backed up by solar symbolism at a smaller scale - also missing. And the six-degree slope does not quite 'reflect the sky' on most days, because the glass is so transparent. Indeed, if visual metaphors are the issue then the major architectural fault is not only their lack of resolution in details, but the repetition of an inadvertent meaning. The result is the typical Late-Modern Malapropism - 'Monothematatis'. As if to confirm this interpretation, the social and cultural differentiation of a vertical town of 10,000 people is given no articulation, no symbolic expression. But still I like the Shard and am glad to see it built, especially when contemplating what commercial architects might have done. As for positive visual metaphors, the Shard provides a 'navigation point' for lost Londoners, popping helpfully into view at many points, yet disappearing into clouds when they are low. Indeed, as a naturalistic and cosmic metaphor it is something of the ultimate 'cloudscraper'. I can even like the building for its non-precious detailing, the way its remorseless windows disappear into nothingness. But maybe I'm not the only person conflicted between opposite tastes and thoughts Europe's biggest building demands to be judged at the highest level.

Richard Rogers, architect

The Shard is the most beautiful addition to the London skyline. Its beauty is in part due to the ever-changing play of light across the facades of the building. Even when the sky is dark it captures the light around it and stands like a blade cutting through the clouds. The Shard represents the ideals behind the 'vertical village with live-work and leisure facilities stacked one on top of another. It is also located in absolutely the right place, standing above one of the most important transport hubs in Europe, limiting the need to use a car. It is, after all, more sustainable to take a lift than drive a car. The Shard anchors the best piece of urban regeneration in the world - the walk from Westminster Bridge to Tower Bridge and beyond - and will help regenerate Southwark: one of the most run-down areas of London. It is a welcome new London landmark.

(Fonte: *English File Advanced Plus* – Página 65)

LIC 18

WHY DO WE CARE WHODUNNIT?

From Sherlock Holmes to Hercule Poirot, Miss Marple to Columbo, Sarah Lund to Saga Noren, we have long been gripped by detective fiction. An opium addiction, elaborate moustaches, a knitting habit, a shabby overcoat, lovely Nordic jumpers, and a green Porsche: the only thing these idiosyncratic and frequently frustrating sleuths seem to have in common is their ability to solve crimes. So what is it that makes the whodunnit such compelling entertainment? A crime is committed, the detective arrives; we watch in suspense as he or she tries to piece together the clues to apprehend the baddie in time. A killer is on the loose, and they need to be caught. But this misses the complexity of detective fiction. The trick in fact lies in its manipulation of time.

Tzvetan Todorov, the Bulgarian-French critic, said that there were three types of narrative drive in crime fiction: the drive of the whodunnit, that of the thriller, and that of the suspense plot. The whodunnit effectively moves backwards to piece together past events (a la Poirot); the thriller moves forwards in step with the characters (a la James Bond); the suspense plot combines both (a la Hitchcock). The whodunnit poses the question: what happened? The thriller makes us ask: what will happen next?

This formula has had us riveted for at least 300 years. Crime writing was popular in the 18th century, with the added tension of the death penalty imposed even for minor misdemeanours. It was, however, only really in the 19th century that the idea of the whodunnit was born. The poet T.S. Eliot believed that Wilkie Collins's *The Moonstone* (1868) invented the genre. The great crime-writer Dorothy L Sayers called it the finest example ever written, and its tale of theft, murder and intrigue continues to grip 21st century readers. More recently, the final season of the Nordic noir series *The Bridge* garnered huge contemporary audiences at home and abroad. *The Moonstone* is set in an upper-class English country home in 1848. *The Bridge* takes place between Sweden and Denmark in 2018. *The Moonstone* follows the disappearance of a priceless diamond stolen by the British from India. *The Bridge* investigates the gruesome murder of a civil servant. At first glance, these two works seem totally different. Most classic detective stories start with a crime, and then pit an ordinary policeman against the brilliant mind of the private detective or amateur sleuth. There is one plot, and the whole story is told in retrospect by one narrator, typically by a loyal companion (e.g. Dr Watson for Sherlock Holmes), who replicates the experience of the reader, witnessing events and desperately trying to keep up. *The Moonstone* and *The Bridge*, on the other hand, contain not one but two interwoven plots.

Wilkie Collins structures *The Moonstone* in a highly innovative way, using multiple narrators. The story oscillates between the drive backwards on the quest for knowledge, and the drive

forwards as a kind of thriller. Rachel Verinder, a young heiress, is bequeathed the priceless but cursed moonstone. On the night of her 18th birthday, the jewel goes missing. The investigation into the crime by Sergeant Cuff dominates the beginning of the story. The first narrator is the head servant, Gabriel Betteredge, who writes down not only what he remembers of the theft, but also of the original tragic history of the moonstone, moving us backwards in time. But at the same time, the investigation moves forward; we focus on relationships in the present, while the puzzle of the past lingers in the background.

The Bridge plays with perspective and time in a similar way to *The Moonstone*. Season four starts with one murder which is quickly followed by two more bodies. When detective Saga Norén discerns a pattern in the murders - execution methods used to carry out death sentences - suspense about the future rises. We discover that there are seven methods of execution used in the world today, so four more to come. Who will the victims be? And when?

But also, by season four, the audience already knows some of the characters' histories. Now, we find Saga in prison, falsely convicted of murdering her mother, and conducting an unofficial investigation into the historical disappearance of her colleague Henrik's children. We continually look back to the characters' pasts - Saga's childhood and Henrik's hallucinations of his daughters - alongside the forward-moving investigation into serial killings in the present. As with *The Moonstone*, the development of personal relationships and the investigation provide forward momentum.

The puzzle of detective fiction is enthralling. The characters may know something that we desperately want to discover, or we may have witnessed something and worry that the detectives won't catch up in time. This timing of knowledge is crucial to the creation of suspense. But the really thrilling stuff combines multiple plots and multiple time perspectives. In expert hands, this combination creates constant momentum. It is not simply the concealing of the significance of certain clues or emphasizing red herrings, but a careful oscillation between piecing together past events and foreshadowing future ones.

(Fonte: *English File Advanced Plus* – Páginas 68 e 69)

*LIC 19***FOOD MEMORIES**

I am not the only one for whom many of life's most intimate details come flooding back at the sight, smell and taste of particular foods. Everyone I speak to seems to have a favourite or, in some cases, a most hated dish with which they can recall particular moments of their lives.

In many cases the taste or smell of a sweet, a cake or an entire meal is capable of painting a picture with richer, deeper brush strokes than any snapshot in their photograph album. It is curious that, while I struggle to remember my mobile phone number or grapple helplessly to recall the closest of friends' names when I am required to introduce them to someone, the merest sniff of chocolate ice cream has been known to bring back memories from 20, 30, 40 years ago with frightening clarity. Put that same ice cream on a little wooden spoon and I can recall the cinema I was in when I ate it. the feel of the (red) velvet seats on the back of my bare knees, the colour of the ice cream attendant's overall (lemon, with green piping). Details more glowing than if I had eaten that ice cream two hours ago.

Curiously, although the flavour of beetroot in vinegar might still instantly transport me to the back lawn of 67 Sandringham Road, Wolverhampton on a summer's day in 1965, it is actually what hits the nose - sweet. earthy, acidic - rather than the tongue that really hits the mark. In terms of total immersion in a time long gone the nostrils seem to win over either the eyes or the taste buds. When people say tomatoes don't taste like they used to, they may be commenting on the changes in commercial production of the fruit (the variety, the soil, the modern chemicals) or simply on their own taste buds' inability to experience a flavour the way they used to. Yet I would argue that it is the aroma from the freshly snapped stalk of a home-grown tomato that is more likely to whisk them off to Dad's old greenhouse than putting one in their mouth.

Of course, food linked with memorable events is more likely to kick off a few reminiscences than something we eat every week, which is why we probably remember birthdays more clearly than the average Sunday lunch. Even so it was nothing more than a mundane pickled herring that whisked me back to St Ives on a magic-carpet ride of onion, dill and vinegar the other day. And not only just to St Ives, but to the car park where my young niece slammed her fingers in the car door. Sharp flavours that brought back every tearful moment of a painful event. Having had many a herring since that dark day, it was only this one, with its distinctive piercing whiff of dill and white vinegar, that opened that particular box. And how one particular mango out of the hundreds whose juices dribble down my chin each year, was the one that reminded me of a woman I once met in a hotel in Sri Lanka. A crazy woman, who played opera at full blast from the bedroom of her hotel in the early hours of the morning - and yet someone I had completely forgotten about till every line in her face came back to me with one bite of that

absurdly juicy fruit. And what exactly was it about that mouthful of boiled potato that reminded me of my school history teacher?

Of all the food triggers that can pop up in our lives there are a few that have done it for me more than others. Treacle tart makes me think of a mysterious uncle, gammon and parsley sauce is unmistakably linked to a row with my stepmother, grilled kippers to an Edinburgh bed and breakfast where I tossed my bacon into a plant pot in the dining room, eggs mayonnaise of being sacked from a job in a hotel, cucumber soup conjures up one of the happiest days of my life. Recipes that all spark deeply detailed recall of not just where I was and whom I was with, but of details so bright and sharp they could have happened yesterday; to tell the truth, probably a damn sight clearer.

(Fonte: *English File Advanced Plus* – Página 76)

LIC 20

A PLATE OF PEAS

My grandfather died when I was a small boy, and my grandmother started staying with us for about six months out of every year. She lived in a room that doubled as my father's office, and which we invariably referred to as "the back room". She carried with her a powerful aroma. I don't know what kind of perfume she used, but it was the double-barrel, 90 proof, knock down, render the victim unconscious moose-killing variety.

She kept it in a huge atomizer and applied it frequently and liberally. I was almost impossible to do into her room and remain breathing for any length of time. When she would leave the house to go spend six months with my Aunt Lillian, my mother and sisters would throw open all the windows, strip the bedding and the curtains and the rugs, and spend several days washing and airing things out, trying frantically to make the pungent odor go away.

This, then, was my grandmother at the time of the pea incident.

It took place at the Biltmore Hotel, which to my eight-year-old mind was just about the fanciest place to eat in all of Providence. My grandmother, my mother and I were having lunch after a morning spent shopping. I grandly ordered a Salisbury steak, confident in the knowledge that beneath that fancy name was a good old hamburger with gravy. When brought to the table it was accompanied by a plate of peas.

I do not like peas now. I did not like peas then. I have always hated peas. It is a complete mystery to me why anyone would voluntarily eat peas. I did not eat them at home. I did not eat them at restaurants. And I certainly was not about to eat them now.

'Eat your peas, my grandmother said.'

'Mother,' said my mom in her warning voice. 'He doesn't like peas. Leave him alone.'

My grandmother did not reply, but she got that certain glint to her eye and grim set to her jaw that signaled she was not going to be thwarted. She leaned in my direction, caught my eye, and uttered the fateful words that changed my life.

'I'll pay you five dollars if you eat those peas.'

I had absolutely no idea of the impending doom that was headed my way like a giant wrecking ball. I only knew that five dollars was an ENORMOUS, nearly UNIMAGINABLE amount of money, and as awful as peas were, only one plate of them stood between me and the possession of that five dollars. I began to force the wretched things down my throat.

My mother was livid. My grandmother had that self-satisfied look of someone who has thrown down an unbeatable trump card. I can do what I want, Ellen, and you can't stop me.' My mother glared at her mother. She glared at me. No one can glare like my mother. If there were a glaring Olympics, she would undoubtedly win the gold medal.

I, of course, kept shoving peas down my throat. The glares made me nervous, and every single pea made me want to throw up, but the magical image of that five dollars floated before me, and I finally gagged down the very last of them. My grandmother handed me the five dollars with a flourish. My mother continued to glare in silence. And the episode ended. Or so I thought. My grandmother left for Aunt Lillian's a few weeks later. That night, at dinner, my mother served two of my all-time favorite foods, meatloaf and mashed potatoes. Along with them came a big, steaming bowl of peas. She offered me some peas, and I, in the very last moments of my innocent youth, declined. My mother fixed me with a cold eye as she heaped a huge pile of peas on my plate. Then came the words that were to haunt me for years

'You ate them for money, she said.' 'You can eat them for love.'

Oh, despair! Oh, devastation! Now, too late, came the dawning realization that I had unwittingly damned myself in a hell from which there was no escape.

'You ate them for money. You can eat them for love'

What possible argument could I muster against that? There was none. Did I eat the peas? You bet I did. I ate them that day, and every time they were served thereafter. The five dollars was quickly spent. My grandmother herself passed away a few years later. But the legacy of the peas lived on, as it lives to this day. If I so much as curl my lip when they are served (because after all, I still hate the horrid little things), my mother repeats the dreaded words one more time

'You ate them for money,' she says. 'You can eat them for love.'

(Fonte: *English File Advanced Plus* – Página 85)