

Níveis e descritores de complexidade textual para adultos de baixa literacia: um referencial do projeto iRead4Skills

Levels and descriptors of textual complexity for adults with low literacy: a benchmark from the iRead4Skills project

Ricardo Monteiro¹, Susana Correia¹, Raquel Amaro¹, Michell Moutinho¹, Sílvia Barbosa¹, Maria Leonor Reis^{1,2}

¹ Centro de Linguística da Universidade Nova de Lisboa

² Academia das Ciências de Lisboa

Abstract

This paper presents the iRead4Skills benchmark of textual complexity, a novel approach to text complexity difficulty tailored to low-literacy adult native speakers. Despite the existence of various instruments and guidelines related to reading proficiency (CEFR, PISA, etc.), little attention is given to adult reading difficulties and how these manifest linguistically in texts. Recognizing the pivotal role of literacy in personal and professional development, this tool addresses the challenges of identifying and adapting texts for this demographic. Based on existent benchmarks (ANQEP, CEFR, PIAAC, ALTE), this article builds upon them and proposes the delineation of levels of textual complexity for European Portuguese, pertaining to low-literacy adults. Integrating lexical, syntactic, and cohesion features, the framework defines three levels of textual complexity—Very Easy, Easy, and Plain—roughly aligned with the CEFR levels A1, A2, and B1. Descriptors of complexity are also advanced to detail each level. Alongside this proposal, the levels and descriptors were tested against real texts, resulting in the compilation of a *corpus*, graded by levels of complexity. With the aim of testing and validating this proposal, both the levels and descriptors were evaluated and discussed by experts in Adult Learning and Vocational Education and Training. The iRead4Skills benchmark of textual complexity aims at bridging theoretical research and practical application by combining quantitative and qualitative measures, ensuring a robust and contextually relevant analysis of textual complexity. Key contributions include the framework's alignment with authentic texts and its validation by experts in Adult Learning and Vocational Education and Training. Practical applications of this framework encompass innovative tools such as a text complexity evaluation system and a writing assistant for text adaptation, addressing gaps in suitable reading materials for low-literacy adults. This research offers valuable insights for cross-linguistic studies on text complexity, paving the way for future exploration in literacy and text complexity frameworks.

Keywords: literacy, textual complexity, descriptors, benchmark, low-literacy adults.

Resumo

Este trabalho pretende descrever o processo de desenvolvimento de um referencial de complexidade textual inserido no projeto iRead4Skills. O nosso objetivo é fazer uma descrição e análise dos traços linguísticos que conferem complexidade aos textos, em particular àqueles que se destinam a adultos de baixa literacia. Apesar da existência de orientações sobre a proficiência em leitura (QECR, PISA, etc.), pouca atenção é dada às dificuldades de leitura sentidas por adultos em programas de ensino e formação de adultos (EFA), e a como estas dificuldades se manifestam linguisticamente nos textos. Baseado em referenciais usados atualmente (ALTE, ANQEP, QECR,

PIAAC), este trabalho propõe a definição de níveis de complexidade do português europeu para adultos de baixa literacia, com base na análise crítica desses referenciais, na observação e análise de textos, e na validação por profissionais da área de EFA. Composta por três níveis – Muito Fácil, Fácil e Claro –, esta proposta considera para cada nível aspetos linguísticos como a dimensão léxico-conceitual, a ocorrência de certos tipos de estruturas sintáticas e a coesão textual, como indutores e/ou marcadores de complexidade. Estas dimensões são detalhadas através de descritores associados aos diferentes níveis de complexidade textual estabelecidos. Paralelamente, e de modo interativo, os níveis e descritores de complexidade foram usados na compilação de um *corpus*, para a seleção de textos para cada nível a partir dos descritores propostos e da identificação dos traços linguísticos específicos de cada texto. Com o intuito de testar e validar a proposta, os níveis e descritores foram levados a avaliação e discussão por professores de EFA. Fazendo a ponte entre a investigação e a aplicação prática em sala de aula, o referencial aqui avançado contribui para a criação de ferramentas automáticas de análise de complexidade textual e para o desenvolvimento de assistentes de escrita de textos com diferentes níveis de complexidade, o que poderá ajudar a desenvolver materiais de leitura adequados a adultos de baixa literacia.

Palavras-chave: literacia, complexidade textual, descritores, referencial, adultos de baixa literacia.

1. Introdução

A literacia é essencial para os indivíduos prosperarem no mundo atual, afigurando-se, cada vez mais, como uma competência vital para o desenvolvimento pessoal e profissional. Ao mesmo tempo, estabelece uma correlação positiva com outros aspetos da vida dos indivíduos, desde melhor remuneração a melhores hábitos de saúde (OECD, 2013a). Com efeito, as questões à volta da literacia na sociedade são cada vez mais discutidas e reconhecidas como importantes para uma vida digna. Atualmente, a noção da literacia é entendida como operando num *continuum* de proficiência, e é sempre uma competência indispensável a jovens e adultos, independentemente da sua posição nesse *continuum* (OECD, 2021; Wagner, 2011). Neste contexto, surge o conceito de “aprendizagem ao longo da vida”, incorporado nos objetivos de programas de requalificação pessoal e profissional de adultos, e que tem sido posto em prática, tanto a nível internacional como nacional, com vista à melhoria das competências literácitas de adultos. No entanto, encontrar materiais de leitura adequados para este tipo de população – especialmente aqueles que são autênticos e relevantes – pode revelar-se uma tarefa difícil. Muitos programas de ensino e formação de adultos (EFA) fazem uso de materiais de leitura provenientes de programas escolares destinados a uma população infantil. O ensino de adultos através de materiais deste género revela-se, assim, uma tarefa difícil, havendo uma desadequação dos materiais de leitura ao contexto de ensino, podendo, isso, tornar-se num obstáculo à promoção do desenvolvimento de práticas de leitura.

Este problema está na origem do projeto iRead4Skills.¹ Reconhecendo a importância das competências de leitura, este projeto, iniciado em 2023, tem como meta a promoção do desenvolvimento de práticas de leitura através do desenvolvimento de um sistema que avalia a complexidade textual e sugere textos adequados a leitores com baixa literacia. O intuito é que o acesso a textos adequados proporcione uma reaproximação dos indivíduos à leitura, o que, por sua vez, estimulará o desenvolvimento das práticas e hábitos de leitura e contribuirá para o desenvolvimento pessoal e a qualificação profissional.

¹ <https://iread4skills.com/>

2. Enquadramento teórico

Nas últimas décadas, o conceito de literacia tem estado cada vez mais presente no contexto das políticas educativas europeias. A literacia é um aspeto fundamental para o acesso à informação e ao conhecimento, na possibilidade de aprender ao longo da vida, bem como para construir uma análise crítica do mundo e exercer uma cidadania ativa (Ávila, 2008). Por essa razão, os sistemas educativos, nomeadamente decisores políticos, escolas e centros de formação, têm reforçado a necessidade de promover a literacia, nas suas diversas formas (por exemplo, a literacia digital, a literacia financeira ou a literacia para a saúde), como uma competência essencial para os indivíduos na sociedade atual. Apesar disso, a relação entre o nível de escolaridade de uma população e o seu perfil de literacia não é óbvia nem direta (Rothes & Queirós, 2021; Scribner & Cole, 1981). Uma sociedade altamente escolarizada não é, necessária e obrigatoriamente, uma sociedade com mais competências literácitas. Para compreender esta competência nos indivíduos, é necessário compreender igualmente como estes usam as suas aptidões de leitura e escrita enquanto participantes ativos da sociedade.

2.1. Literacia

Quando se fala de literacia, uma abordagem descritiva das suas definições, ao invés de normativa, é recomendada (Rothes & Queirós, 2021). Isto traduz-se no entendimento da literacia não como uma competência dicotómica que se tem ou não –, mas como operando num *continuum*, que pode ser desenvolvido e variar de indivíduo para indivíduo (Hanemann, 2016; UNESCO Institute of Statistics, 2020). Assim, a literacia configura-se como um processo dinâmico e não um estado fixo (Alves, 2005), podendo assumir diferentes formas e finalidades consoante os contextos sociais e culturais em que se insere.

É igualmente importante entender a literacia como um fenómeno que ultrapassa a mera esfera individual e é, também, influenciado por fatores externos: os textos em práticas sociais, contextual e temporalmente situadas (Alves et al., 2012). Estes contextos concretos são designados por *literacy events* (Fransman, 2005). Nestes, a interação do indivíduo com uma prática social é feita através de um texto, e, de forma reflexa, a interação com o texto é feita no âmbito da prática social. Por conseguinte, os usos e práticas dos textos não apresentam sempre as mesmas características, variando conforme as esferas de vida (Dionísio et al., 2014). Desta maneira, torna-se claro que a dimensão social influencia a interação com o texto, sendo a competência em literacia de cada indivíduo dependente do contexto social no qual é posta em prática.

É, pois, na relação recíproca entre o indivíduo e a sociedade que a literacia assume um papel cimeiro. Atualmente, o impacto da literacia é visível ao nível socioeconómico e na competitividade entre países (Cunha & Dionísio, 2019). A literacia é vista como uma prática social, na qual os falantes de uma língua utilizam as suas aptidões de leitura e escrita para o seu desenvolvimento pessoal e profissional (OECD, 2021; Stromquist, 2006). Complementarmente, a Organização para a Cooperação e Desenvolvimento Económico (OECD) salienta a relação entre baixos níveis de literacia e uma população adulta desempregada (OECD, 2013a). Nesta senda, múltiplas campanhas de literacia por parte da UNESCO têm estabelecido uma ligação entre indicadores mensuráveis da literacia e outros indicadores de desenvolvimento, por exemplo, de economia ou de modernização (Wagner, 2011). Da mesma forma, uma análise baseada nos resultados do *Survey of Adult Skills-PIAAC* (OECD, 2013a), constata que, em média, indivíduos com níveis de literacia considerados altos têm salários 60% superiores a indivíduos com níveis de literacia considerados baixos (Schleicher, 2012). Comparando estas duas populações no que diz respeito à situação profissional, foi ainda demonstrado que indivíduos de baixa em literacia têm o dobro da probabilidade de se encontrarem desempregados (Schleicher, 2012). De facto, as consequências da literacia noutros aspetos da vida, à primeira vista sem relação, parecem não ficar por aqui: uma população com altos níveis de literacia goza de um maior poder de intervenção política, de uma maior participação em atividades de voluntariado, bem como de

uma saúde boa ou excelente (Schleicher, 2012). Complementarmente, baixos índices de literacia são frequentemente associados a disfuncionalidade profissional e a défices sociais.

Importa salientar que níveis de escolaridade formais não garantem, por si sós, níveis elevados de literacia. Os dados do mais recente estudo PIAAC (OCDE, 2024a) indicam que Portugal continua abaixo da média da OCDE em termos de proficiência em literacia, situando-se no penúltimo lugar entre os países participantes. Este défice abrange não apenas adultos com baixa escolaridade, mas também indivíduos com níveis de escolaridade mais avançados. Isto confirma que uma sociedade escolarizada não é automaticamente uma sociedade mais letrada (Scribner & Cole, 1981), e reforça a importância de estudar a literacia como prática situada, mais do que como mera acumulação de anos de escolaridade.

2.2. Ensino e formação de adultos

De modo a contrariar os efeitos negativos da baixa literacia nos indivíduos e na sociedade, têm sido postos em prática vários programas de incentivo à requalificação pessoal e profissional, tanto a nível internacional como nacional. Estes programas baseiam-se em valores humanísticos, democráticos e emancipatórios e têm como objetivo requalificar e integrar uma população menos letrada (Hanemann, 2016). Através do ensino formal e informal, destinados a várias faixas etárias e com uma vasta abrangência social, providenciam-se oportunidades de aprendizagem, seguindo o princípio que aqui tem vindo a ser mencionado: que a competência de literacia de cada indivíduo pode ser desenvolvida, independentemente dos fatores externos. De acordo com dados da EDUSTAT – Fundação Belmiro de Azevedo (2024), no ano letivo 2022-2023 estavam inscritos mais de 37.000 adultos em cursos de formação EFA, a grande maioria (78%) com idades superiores a 23 anos. Segundo Correia et al. (2024), este perfil caracteriza-se por ter idade entre 18 e 40 anos, residir na região de Lisboa, ser do sexo feminino, empregado, casado, com filhos e com ensino básico concluído.

O sucesso destes programas está, assim, fortemente condicionado à capacidade de responder às necessidades reais dos indivíduos, as quais passam, muitas vezes pela contextualização dos processos de ensino. (Alves et al., 2012; Cunha & Dionísio, 2019). Um dos principais obstáculos identificados pelos profissionais da área prende-se com a escassez de materiais de leitura adequados, tanto em conteúdo como em grau de complexidade. Este problema foi identificado durante o desenvolvimento do projeto iRead4Skills, através de contributos diretos da Direção-Geral da Educação - Plano Nacional de Leitura 2027 (PNL), entidade parceira do consórcio. O PNL reportou que, em muitos contextos de ensino de adultos, é prática comum a reutilização de materiais pedagógicos concebidos para crianças ou jovens, nomeadamente fichas de trabalho, textos escolares e exercícios de manuais do 1.º e 2.º ciclos. Esta reutilização, ainda que prática, levanta sérios desafios pedagógicos, uma vez que os textos utilizados nem sempre estão ajustados às vivências, interesses ou competências linguísticas dos adultos, podendo resultar em frustração ou desmotivação. Estes materiais são, assim, desajustados: ou por serem excessivamente simplistas ou infantis, ou por utilizarem uma linguagem ou estrutura desfasada da experiência de vida dos adultos. Por exemplo, os padrões de leitura (i.e., o hábito, frequência e preferências de leitura) tornam-se tipicamente mais ricos com a idade e, por isso, são mais avançados nos adultos do que nas crianças (Flor et al., 2013). Isto sugere que há uma possibilidade de “amadurecimento” a partir de certas estratégias de leitura, que são desenvolvidas de forma implícita em cada indivíduo. Assim, a consequência de materiais inadequados é dupla: perda de motivação e ausência de identificação com os textos propostos.

Esta inadequação é particularmente preocupante num contexto em que os cursos EFA acolhem públicos heterogêneos em termos de idade, contexto profissional e escolarização. Tal diversidade exige um olhar mais fino

sobre os níveis de proficiência reais em leitura, que nem sempre se correlacionam com a idade ou o grau de escolaridade formal. A identificação de um perfil mais preciso do público-alvo e das suas práticas de leitura é, portanto, essencial para a produção de materiais adequados.

2.3. Complexidade textual

Cunningham e Mesmer (2014, p. 256) definem complexidade textual como um conjunto de características inerentes e exclusivas ao texto, que podem ser analisadas e estudadas de um modo distinto das dificuldades do leitor. Por outro lado, o *Common Core State Standards/English Language Arts* (CCSS/ELA) (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010) define a noção de complexidade como algo subjetivo. Desta forma, o entendimento do que torna um texto complexo é visto como um fenómeno relativo, que varia de pessoa para pessoa (Kusters, 2003). O CCSS/ELA propõe ainda um sistema tripartido para avaliar a complexidade de um texto: a sua dimensão quantitativa, qualitativa, e a tarefa do leitor para o compreender.

Numa abordagem quantitativa, a dificuldade de um texto advém, sobretudo, da quantidade de ocorrências de algo que induz complexidade. Esta ideia é, essencialmente, executada através de fórmulas de *leiturabilidade*,² cuja aferição passa por contar ocorrências de determinadas propriedades linguísticas (e.g., palavras compostas, número de sílabas numa palavra, número de palavras numa frase). Esta contagem resulta, depois, num valor numérico associado a um nível de complexidade específico (Goldman & Lee, 2014). Por outro lado, a abordagem qualitativa tenta ir para além da quantidade de características linguísticas, que resultam em índices de leiturabilidade, fazendo com que estas estejam dependentes de uma interpretação posterior (Hiebert, 2011). Fundamentalmente, uma análise qualitativa da complexidade procura simular a inteligência humana e a forma como esta pode percecionar a complexidade. Nesta abordagem, estão incluídas as teorias de esquemas textuais,³ modelos mentais representativos do processo de leitura,⁴ assim como gramáticas de histórias.⁵ São abordagens deste género que procuram aferir a complexidade textual de uma forma mais holística e qualitativa (Harley, 2013).

Neste contexto, a proposta de referencial de complexidade apresentada neste trabalho é desenvolvida para adultos, falantes nativos de português europeu de baixa literacia, através de uma combinação de medidas de aferição quantitativa, qualitativa e subjetiva. Por meio de uma abordagem *in tandem*, direcionada para um público-alvo concreto, pretende-se colmatar as lacunas de cada abordagem específica (Goldman & Lee, 2014; Hiebert & Pearson, 2014), enquanto se emprega uma forma inovadora, potencialmente capaz de identificar manifestações de complexidade textual não analisadas até ao momento.

3. Metodologia

Nesta seção, é detalhado o processo de desenvolvimento de uma infraestrutura de complexidade textual destinada a adultos de baixa literacia, para avaliar a dificuldade de textos e sugerir leituras apropriadas para estes indivíduos. É, também, descrito o processo de elaboração do referencial com descritores de complexidade, de forma

² Tradução de *readability*.

³ O conceito de teorias de esquemas textuais baseia-se na organização (esquematização) dos elementos do texto com base no conhecimento prévio do leitor destes mesmos elementos e a forma como se podem relacionar com outros elementos (Anderson, 2000).

⁴ Os modelos mentais de leitura pretendem representar a informação transmitida por um texto à medida que esta informação é dada. O objetivo é reproduzir as imagens mentais criadas pelos leitores na sua memória de curto prazo, e como estas percecionam elementos como o tempo, espaço, causalidade, etc. (Harley, 2013).

⁵ Gramáticas de histórias identificam e analisam a estrutura de uma história, assim como os seus elementos.

a contribuir para a classificação dos níveis de complexidade textual. Para alcançar esta meta, o nosso trabalho passou por três fases distintas:

- i. Compilação e análise de referenciais com descritores e orientações relevantes para a avaliação da complexidade textual.
- ii. Definição de uma proposta de níveis de complexidade, a partir das características linguísticas de textos adequados a adultos nativos de baixa literacia.
- iii. Validação da proposta por parte de um *focus group* constituído por profissionais de EFA, com experiência em ensino de adultos de baixa literacia.

3.1. Análise de referenciais

Nesta primeira fase, foi efetuado um levantamento dos documentos orientadores pertinentes à complexidade textual. Quatro referenciais foram considerados relevantes: o *Quadro Europeu Comum de Referência para as Línguas* (QECR) (Council of Europe, 2020), o *Inquérito às Competências dos Adultos* – PIAAC (OECD, 2013b), o *Can Do Project*, da Associação de Examinadores de Línguas na Europa (ALTE, 2002) e o *Referencial de Competências-Chave de Educação e Formação de Adultos* – Nível Básico da Agência Nacional para a Qualificação e o Ensino Profissional (ANQEP) (Valente, 2021). À exceção do último, os documentos selecionados não tinham como público-alvo adultos de baixa literacia, abordando as competências comunicativas e linguísticas numa ótica mais geral ou no contexto de aprendizagem de língua segunda (L2)/língua estrangeira (LE).

Embora o público-alvo do presente estudo seja constituído por adultos com baixos níveis de literacia, falantes nativos do português europeu, a opção metodológica seguida foi a de abordar os textos a partir da noção de complexidade textual, em vez da noção de dificuldade textual. Esta escolha implica uma mudança de foco conceptual: enquanto a dificuldade textual parte da perspetiva do leitor e da sua interação com o texto, a complexidade textual é tratada como um conjunto de propriedades linguísticas e discursivas inerentes ao próprio texto, categorizáveis independentemente do leitor (Hiebert & Mesmer, 2013).

Optámos por esta abordagem com o intuito de avaliar e descrever os textos em termos das suas propriedades formais e estruturais, sem partir de pressupostos sobre a capacidade de leitura de um grupo específico de indivíduos. Só num momento posterior da investigação foi feita a correspondência entre os níveis de complexidade e a acessibilidade para o público-alvo. Neste sentido, a inclusão de referenciais concebidos para o ensino do português como língua segunda ou estrangeira (L2/LE) – nomeadamente descritores de proficiência do QECR – não teve como objetivo equiparar o nosso público-alvo a falantes não nativos. Tais referenciais foram usados como ferramentas operacionais, amplamente testadas, que permitiram descrever níveis de complexidade linguística com base em critérios já estabilizados na investigação aplicada, tais como o domínio de estruturas sintáticas, uso lexical ou coesão. Assim, a abordagem adotada é intencionalmente heurística. Ao descrever a complexidade dos textos com base em critérios independentes do/a leitor/a, é possível, numa fase posterior, observar com maior precisão quais níveis de complexidade se associam a que tipos de dificuldades no público-alvo. Esta metodologia também permite maior transferibilidade do referencial para outros contextos.

Na análise destes quatro documentos, foi dada especial atenção aos três níveis iniciais propostos em cada um. A razão para esta abordagem está relacionada, novamente, com o facto de estarmos perante referenciais que, na sua maioria, não incidem exclusivamente sobre adultos de baixa literacia. O facto de o conceito da literacia ser algo abrangente, desdobrável e passível de ser abordado de várias maneiras, faz com que seja difícil definir o que é um

indivíduo com baixa literacia. Ademais, ao contrário dos níveis de escolaridade formais, não é fácil fazer corresponder níveis de literacia pré-estabelecidos a leitores/as. Desta forma, a nossa abordagem passa por enquadrar a literacia no contexto pessoal e social (e.g., argumentar um ponto de vista, ler e compreender notícias, captar informação sobre transportes públicos, etc.) (Fransman, 2005; Street, 1993).

Assim, é preciso considerar que adultos de baixa literacia não se confinam apenas a um nível, mas que existe uma escala capaz de encapsular a heterogeneidade desta população. Com efeito, nos referenciais analisados, os primeiros níveis de proficiência em leitura alinham-se com as competências literácitas e géneros/tipos de texto associados com este tipo de população. Da mesma forma, estes níveis dão pistas em relação aos limiares do enquadramento da baixa literacia de adultos. Por exemplo, no QECR, o nível B2 – nível imediatamente acima dos que foram considerados pelo projeto –, descreve um indivíduo capaz de “ler artigos e reportagens sobre assuntos contemporâneos em relação aos quais os autores adotam determinadas atitudes ou pontos de vista particulares” (Council of Europe, 2020, p. 177). Uma proficiência de leitura deste género é aqui vista como não associada a um adulto de baixa literacia. No ponto oposto, o nível A0 corresponde ao estágio de alfabetização, no qual não existe qualquer atividade de leitura autónoma. Assim, a proficiência de leitura nestes dois extremos não é relevante para o público-alvo em questão, cingindo a nossa análise às características dos níveis intermédios – A1, A2 e B1.

3.2. Proposta de níveis de complexidade

Depois de analisados os referenciais e designados os níveis de complexidade apropriados a um contexto de baixa literacia, foram compiladas e analisadas as características associadas a cada nível. O objetivo passou por discutir descrições subjetivas e pouco precisas (e.g., *Knowledge and skill in recognizing basic vocabulary determining the meaning of sentences, and reading paragraphs of text is expected* (OECD, 2013b, p. 71)) e, sempre que possível transpô-las para – ou associá-las a – descrições que contivessem propriedades linguísticas objetivas (e.g., ausência de verbos conjugados no Modo Conjuntivo). Estas propriedades servem, assim, como descritores de cada nível.

Depois do levantamento destes descritores linguísticos (ver Anexo 1), partiu-se para uma análise da sua presença em textos autênticos de modo a:

- i. Determinar quão prototípicos de um nível seriam (i.e., se um certo descritor ocorreria frequentemente nos níveis ou seria exclusivo de um nível ou género textual)
- ii. Identificar outros descritores nos textos e especificar os já listados.

Este processo de validação e recolha de descritores em textos autênticos contribuiu também para o desenvolvimento de um *corpus* por nível de complexidade (Pintard et al., 2024). Este trabalho consistiu numa abordagem *in tandem*, na qual uma parte da equipa se encarregava de compilar textos para o *corpus* e outra os utilizava para a identificação e validação dos descritores (cf. Reis et al., 2024). Esta última tarefa pautava-se pela identificação de descritores de cada nível de complexidade num certo texto, de modo a atribuir um nível a esse texto. Uma tabela de ocorrências de descritores acompanhava cada texto e era nesta que as ocorrências de cada descritor eram anotadas. Depois, o número total de ocorrências era utilizado para decidir a que nível correspondia aquele texto ou se o mesmo seria descartado. O resultado deste trabalho foi um *corpus* de textos autênticos organizado por níveis de complexidade.

Por último, para além dos níveis de baixa literacia propostos, foi também postulado um nível de complexidade superior – o nível Mais Complexo. A consideração deste nível está ligada à necessidade de estabelecer um limiar de complexidade superior. Como referido anteriormente, existe também uma fronteira inferior natural de

complexidade, que distingue adultos de baixa literacia de adultos em processo de alfabetização. O objetivo destes limiares inferior e superior de complexidade é funcional, enquadrando a população de baixa literacia. O nível Mais Complexo, apesar de estar fora do escopo do projeto, foi usado para compilar outros textos, de modo a servir de referência para os níveis de complexidade considerados neste projeto.

O resultado desta etapa de trabalho foi uma proposta de três níveis de complexidade (juntamente de um quarto nível – Mais Complexo), assim como a sua descrição detalhada e os descritores linguísticos associados a cada um.

3.3. Validação dos níveis e descritores de complexidade por formadores de adultos

De modo a validar a proposta desenvolvida, foi pedido que pessoal formador de EFA (nomeadamente, em Centros Qualifica)⁶ fizessem uma avaliação dos descritores para cada nível. Este foi um processo iterativo cujo objetivo principal não foi o de realizar uma validação estatística ou normativa, mas sim o de recolher contributos especializados para o aperfeiçoamento linguístico, funcional e pedagógico dos descritores propostos. Esta etapa visou validar a proposta inicial através da experiência de profissionais, assegurando assim uma validação situada e contextualizada.

Este processo incluiu a apresentação da proposta e o *input* dos/as formadores/as foi recolhido em duas fases:

- i. Tarefa Preparatória – Esta tarefa, sob a forma de um questionário, consistiu em recolher as suas opiniões sobre os descritores de complexidade (cf. Monteiro et al., 2023). Neste questionário, eram apresentados, primeiramente, os níveis de complexidade e, depois, os descritores, estes últimos de modo aleatório. O que se pediu aos/às participantes foi que fizessem uma correspondência entre os descritores e os níveis de complexidade previamente apresentados. De modo a garantir a compreensão por parte de todas as pessoas envolvidas, independentemente da disciplina que lecionavam (e.g., matemática, língua portuguesa, etc.), os descritores foram apresentados sem recurso a terminologia linguística e acompanhados de exemplos ilustrativos.
- ii. Sessão de *focus group* 1 – Esta primeira reunião com as/os formadoras/es teve como objetivo discutir os resultados da Tarefa Preparatória. Nesta reunião, foram abordados os conceitos de complexidade textual, assim como os níveis e descritores de complexidade da nossa proposta, com base nos resultados da tarefa preparatória. Foram discutidos, em particular, os descritores relativamente aos quais não houve consenso por parte das/os formadoras/es.
- iii. Sessão de *focus group* 2 – A segunda reunião teve como objetivo a validação das alterações aos descritores, provenientes das conclusões da primeira reunião. Foram apresentados os descritores finais para cada nível, considerando o *feedback* dos participantes, e foi pedida uma validação final às/aos formadoras/es.

No total, 8 formadoras/es EFA participaram nas sessões de *focus group*. As sessões decorreram em ambiente colaborativo, com base em guiões estruturados.

O processo de validação e refinamento do nosso trabalho será continuado em fases posteriores do projeto, onde está prevista a aplicação de questionários e inquéritos a formadoras/es e formandas/os dos cursos EFA, com o objetivo de recolher informação empírica sobre a adequação da futura ferramenta digital, cuja base será o referencial

⁶ Os Centros Qualifica oferecem educação e formação a adultos, no âmbito do programa Qualifica. Para mais informação, consulte <https://www.qualifica.gov.pt>.

aqui exposto. Nesta fase, a validação será centrada nos utilizadores finais e permitirá validar a proposta em termos de eficácia, clareza e relevância pedagógica.

Assim, a validação da ferramenta deve ser entendida como parte de uma estratégia de desenvolvimento cooperativo e progressivo, que articula contributos especializados na fase de construção com testes de aplicabilidade no terreno, a decorrer ao longo do projeto. Esta abordagem está alinhada com metodologias de investigação aplicada, em que a validação se realiza em ciclos sucessivos de desenvolvimento, aplicação, avaliação e reformulação.

4. Apresentação e discussão dos resultados

Nesta secção, são apresentados os resultados do processo de estabelecimento dos níveis, descritores de complexidade textual para o português europeu, desenvolvidos no projeto iRead4Skills, assim como a validação dos mesmos. Primeiramente, a análise dos referenciais supramencionados conduziu à definição de três níveis de complexidade textual: **Muito Fácil**, **Fácil** e **Claro**. A estes, junta-se o nível superior Mais Complexo, que tem apenas um papel delimitativo e de referência.

Para cada nível, é apresentada uma lista de marcadores linguísticos indutores de complexidade textual. Desta forma, a nossa proposta identifica um conjunto de características textuais que contribuem para a dificuldade de compreensão de um texto, incluindo características lexicais, sintáticas e de coesão textual.

4.1. Resultado da análise dos referenciais

Os resultados apresentados nesta secção correspondem à análise realizada sobre um conjunto de referenciais existentes no domínio da leitura e da proficiência linguística. Os referenciais analisados incluíram documentos destinados ao ensino do português como língua materna, língua segunda e língua estrangeira, bem como instrumentos internacionais de avaliação de literacia em adultos. A partir desta análise, foi possível isolar descritores representativos de diferentes níveis de complexidade textual, organizando-os em quatro dimensões principais, que servem de base à proposta apresentada nas seguintes secções.

4.1.1. Níveis de complexidade

Os três níveis de complexidade propostos são: **Muito Fácil**, **Fácil** e **Claro**. Importa referir, a este propósito, que estas designações para os níveis de complexidade foram estabelecidas de forma a classificar os textos e não público-alvo. A nomenclatura escolhida resultou de um processo iterativo e colaborativo no seio do projeto, em que foram testadas diferentes alternativas, acabando por ser validada com base na sua funcionalidade pedagógica e na sua inteligibilidade prática, mesmo que não perfeitamente homogénea do ponto de vista conceptual.⁷ A estrutura progressiva está garantida não apenas pela ordem dos níveis, mas sobretudo pelos descritores linguísticos associados a cada nível, que refletem, de forma rigorosa, a crescente densidade lexical, complexidade sintática, riqueza coesiva e sofisticação discursiva dos textos.

A Tabela 1 apresenta uma descrição muito breve e acessível dos três níveis. Esta descrição, usada na tarefa de validação por formadoras/es de EFA, teve como objetivo apresentar uma descrição muito geral dos níveis de complexidade, para efeitos de contextualização da tarefa sem referência direta às propriedades dos textos. Não pretende apresentar de forma cabal as características de cada nível.

⁷ Entende-se que os conceitos de “facilidade”, “clareza” e “complexidade” podem ser vistos como distintos, operando em campos semânticos diferentes.

Tabela 1*Descrição Acessível dos Níveis de Complexidade Alvo*

Nível	Descrição
Muito Fácil	Textos que são totalmente ou quase totalmente compreendidos por todas as pessoas, incluindo pessoas com escolaridade muito baixa (i.e., que não chegaram ao 6º ano do ensino básico) e que não têm quase nenhuma prática de leitura.
Fácil	Textos que são totalmente ou quase totalmente compreendidos por pessoas com escolaridade baixa (i.e., que completaram o 6º ano do ensino básico, mas que não têm mais que o 9º ano) e que têm pouca prática de leitura.
Claro	Textos que são compreendidos à primeira leitura por pessoas que completaram o 9º ano e que têm uma prática de leitura média.

Nestas descrições, cada nível de complexidade textual é associado a um nível de escolaridade para fornecer um ponto de referência familiar e pragmático. A definição dos níveis de complexidade fundamentou-se na noção de que a complexidade dos textos pode revelar-se sob a forma de três problemas distintos: problemas superficiais, de conhecimento e de conteúdo (Beck et al., 1984). O primeiro – problemas superficiais – relaciona-se com questões estritamente linguísticas, como é o caso de estruturas sintáticas complexas ou um uso inconsistente de conjunções. O segundo – problemas de conhecimento – assenta em algo potencialmente fora da dimensão linguística, ligado à falta de conhecimento prévio sobre o tópico presente (ou envolvente) no texto. O último – problemas de conteúdo – prende-se com a forma como as ideias são apresentadas e interligadas no texto, ou seja, com a forma como o texto está organizado e é comunicado. Nesta última dimensão, ideias e conceitos pouco familiares ou que se contrariam ou que levam a conclusões contraintuitivas podem levantar problemas.

A partir desta tipologia de problemas textuais, a complexidade foi analisada sob o prisma da observação de quatro dimensões textuais distintas: a dimensão léxico-conceptual, a dimensão das estruturas sintáticas, a dimensão da coesão e, por fim, a dimensão das formas verbais. Estas quatro dimensões distinguem-se daquelas existentes em referenciais de proficiência em leitura por incidirem numa população leitora diferente das desses referenciais - tipicamente, falantes não nativos ou crianças. Por exemplo, no caso de falantes não-nativos, estes tendem a encontrar complexidade, principalmente, em questões gramaticais (Heilman et al., 2007), e crianças tendem a ter menos proficiência em leitura em comparação com adultos, devido simplesmente à sua menor experiência de leitura (Mesmer et al., 2012).

A **dimensão léxico-conceptual** é a mais relevante das quatro. Questões de compreensão lexical são as primeiras a surgir quando um/a leitor/a interage com um texto (Harley, 2013; Hu & Nation, 2000; Laufer, 2020). Com efeito, para atingir 70% da compreensão de um texto, a pessoa precisa de conhecer cerca de 95% a 98% das palavras nesse mesmo texto (Schmitt et al., 2011). Tal implica que mesmo uma pequena percentagem de vocabulário desconhecido pode comprometer seriamente a compreensão global. Adultos de baixa literacia, ainda que experientes em língua oral, não têm práticas de leitura habituais e diversas e, por conseguinte, têm uma escassez de exposição a “textos formais”, isto é, a textos nos quais são requeridas estratégias de compreensão diversas e mais avançadas. Desta forma, a **dimensão sintática** considera tanto medidas quantitativas como qualitativas da complexidade. Entre os marcadores incluem-se o **comprimento das frases**, o **número de orações subordinadas** e a **presença de construções sintáticas não canónicas**, como as orações passivas ou inversões. Embora adultos nativos com baixa

literacia possam dominar essas estruturas na oralidade, a sua presença em textos escritos pode representar obstáculos à fluidez da leitura e à construção de sentido.

No que toca à **coesão textual**, analisam-se sobretudo os mecanismos de **coesão referencial**, como as relações anafóricas, e **os marcadores temporais**, como a linearidade da narrativa textual (Fang & Pace, 2013; Glenberg et al., 1987). A ausência de pistas coesivas claras ou a sua ambiguidade dificulta a articulação entre frases e parágrafos, forçando o/a leitor/a a fazer inferências mais exigentes para manter a coerência do texto.

Por último, as formas e construções verbais, ainda que fazendo parte da dimensão lexical e morfossintática, são, aqui, consideradas independentemente. A razão para esta separação está relacionada com o facto de estas formas integrarem **informações aspetuais, temporais e modais**, manifestando, assim, complexidade textual de uma forma independente da estrutura sintática.

Como resultado, cada nível de complexidade possui uma descrição que engloba estas quatro dimensões, como se pode ver na Tabela 2. As informações detalhadas encontram-se sistematizadas sob a forma de descritores no Anexo 1.

Tabela 2

Níveis de Complexidade iRead4Skills

Nível Muito Fácil
Textos curtos e simples até cerca de 50 palavras, que se refiram ao desempenho de tarefas do quotidiano ou apresentem nova informação (ex.: textos didáticos). Conceitos simples do dia a dia, com o tópico da comunicação a ser apresentado inicialmente. Contextos básicos de comunicação: quotidianos (horários de transportes/listas/menus/informação de preço), comunicações pessoais (cartas, mensagens). Palavras simples e frequentes, comumente usadas em contextos do dia a dia (transporte, comida, família, trabalho). Uso predominante de palavras concretas que fazem referência a objetos concretos no mundo real (<i>maçã, cadeira, fogão</i>). Ausência de abreviaturas, siglas e estrangeirismos, à exceção daqueles muito frequentes (internet, hamburger). Palavras compostas comuns (<i>couve-flor, guarda-chuva</i>). Palavras com afixos frequentes (<i>infeliz, calmamente, trabalhador</i>). Frases curtas e simples, utilizando conjunções simples e na ordem direta (Sujeito-Verbo-Complemento). Estruturas de coordenação frequentes (copulativas, disjuntivas e adversativas) estão presentes (ex.: <i>O Rui <u>e/ou</u> a sua irmã foram à festa, <u>mas</u> o João não foi</i>). Estruturas de subordinação frequentes estão presentes (ex.: oração subordinada adverbial temporal – <i>Desde que voltou, o Rui não dormiu</i>). Uso predominante de frases ativas. Tempos verbais no modo Infinitivo, Indicativo e Imperativo, nomeadamente o Presente e Pretéritos Perfeito e Imperfeito. Ocorrência de construções perifrásticas comuns como <i><u>estar/começar/andar + a + infinitivo</u></i> ou <i><u>deixar/acabar + de + infinitivo</u></i>. Não figuram tempos compostos, à exceção do Futuro que é transmitido através de <i><u>ir + infinitivo</u></i>. Cadeia referencial completa sem omissão de elementos, com repetição em vez da utilização de pronomes (ex.: <i>O Rui gosta da <u>Maria</u>, porque a <u>Maria</u> é simpática</i>).

Localização temporal linear, dada por advérbios e conjunções simples (ex.: *Ontem*, *o Rui foi à praia*. *Os bombeiros chegaram depois da polícia*).

Ausência de figuras de estilo.

Nível Fácil

Textos curtos até 150 palavras, que são interessantes para o leitor **se informar/aprender algo novo ou para momentos de lazer**.

Conceitos mais específicos, mas do interesse do leitor (profissional, desporto, hobbies).

Contextos de comunicação profissionais e dos media (ex.: instrução profissional, notícias, anúncios, desporto).

Palavras e expressões frequentes que fazem referência a assuntos do dia a dia, mas também a outros âmbitos mais específicos (ex.: curso, ano escolar). O texto contém algumas palavras que referem ideias abstratas mais frequentes (*felicidade, imaginação, carinho*).

O texto utiliza abreviaturas comuns, depois de as ter explicado (ex.: *litro – l/metro – m*).

O texto apenas utiliza siglas e acrónimos conhecidos (CP, IVA, RTP, SIC, TVI).

Uso de estrangeirismos menos frequentes (*site, souvenir*).

O texto utiliza palavras compostas comuns com o seu sentido a poder ser compreendido pela combinação das duas partes (*óculos de sol, mal-educado, meia-idade, maus-tratos*).

Palavras com afixos menos frequentes (*pré-cozinhado, embalagem, finíssimo*).

Frases simples e complexas através de conjunções coordenadas e subordinadas (ex.: *Os músicos estavam tocando, mas poucos clientes ouviam. Quando pararam de tocar, apenas um senhor se manifestou*). Uso de orações relativas de sujeito, mas não de objeto (ex.: *O senhor que vive na casa ao lado ouviu o barulho*). O texto pode conter frases apresentadas numa ordem não linear (Complemento-Verbo-Sujeito) (ex.: *A escola, chegaram os autocarros da visita de estudo*).

Possível uso de frases passivas.

Uso de modos e tempos verbais variados.

Uso de tempos verbais compostos e de expressões verbais com *ter/poder/dever* (*posso/devo/tenho de gostar de bolo*). Modos Conjuntivo, Condicional e Gerúndio.

Expressões com os verbos *permanecer, querer, poder, conseguir* em tempos simples (ex.: *permaneça calmo, queria almoçar, consigo ir*).

Utilização do Futuro simples (ex.: *ele gostará de te ver*).

O texto tem alguns elementos omitidos e utiliza pronomes (ex.: *Eu gostava da Maria vs. Gostava da Maria/Mas ela não gosta dele*).

Sequência ou localizações temporais lineares, dadas através de advérbios de tempo ou conjunções, **mas também por tempos verbais** (ex.: *Ele está a caminho e a festa já começou*).

Figuras de estilo frequentes são utilizadas.

O texto faz uso da **metonímia** (ex.: *beber um copo*), assim como **metáforas familiares** (ex.: *O meu pai tem olhos de águia*).

Nível Claro

Textos com cerca de 250 palavras sobre vários tópicos, sejam estes de informação ou lazer, com os leitores a conseguirem sair fora da sua zona de conforto.

Conceitos variados que necessitam de ser interligados para a sua compreensão total (ciência, política).

Contextos comunicativos variados: lazer (literatura, reportagens); **artes** (poesia); **contextos especializados** (fóruns online, comunicação profissional formal).

Palavras e expressões menos frequentes que fazem referência a âmbitos específicos (ex.: *doença cardíaca, artéria, tratamento*). As palavras remetem para **conceitos mais abstratos** (*envelhecimento, razão, fidelidade*).

Utilização de siglas e acrónimos não conhecidos pelo leitor juntamente com o nome a que se referem (ex.: *Organização das Nações Unidas (ONU), Instituto Público (IP)*).

O texto utiliza abreviaturas comuns, sem explicação prévia (ex.: *m, km, g., l, Exmo.*).

Uso frequente de estrangeirismos comuns (ex.: *link, hobby, marketing, mail, gay*).

O texto contém palavras que são ambíguas e/ou pouco frequentes. (ex.: *interno, representar, forma, paciente*).

Palavras com sufixos e prefixos variados (ex.: *nervosismo, lavável, curiosidade, altitude, superpoder, preconceituoso*).

Frases frequentemente combinadas através de conjunções coordenativas e subordinativas e, por vezes, com conjunções menos comuns (ex.: *apesar da confusão provocada pelos animais à solta, a situação foi controlada sem os animais serem atropelados.*)

Presença frequente de ordem não linear de constituintes (sujeito-verbo-complementos) (ex.: *de ar puro, todos precisamos.*).

Utilização habitual de frases passivas (ex.: *o almoço foi pago pela empresa.*), incluindo construções impessoais comuns e passivas de -se (ex.: *O trabalho faz-se bem.*).

Utilização de todos os modos e tempos verbais compostos e simples, à exceção do Pretérito Mais-que-perfeito.

Elementos omitidos (ex.: *[] mostrámos em estudos anteriores que a gordura solidifica quando [] arrefece. vs. Nós mostrámos em estudos anteriores que a gordura solidifica quando ela arrefece*). Utilização de **pronomes e anáforas lexicais** para indicar o que é referido (ex.: *no livro, nós damos uma definição complete de uma tempestade de neve e como esta se forma. Também indicamos onde o fenómeno pode ser visto.*)

Sequências ou localizações temporais não lineares (ex.: *após analisarem as respostas dos doentes que tinham sofrido um ataque cardíaco, os médicos confirmaram que a gordura é o alimento mais perigoso.*)

Uso de **figuras de estilo** menos frequentes.

Metáforas incomuns (ex.: *A sociedade está doente; ele é um lobo em pele de cordeiro*).

4.1.2. Descritores de complexidade

A partir da sistematização das dimensões textuais identificadas e dos níveis avançados, foram elaborados descritores de complexidade. Estes descritores visam traduzir, de forma operacional, as características linguísticas, conceptuais e discursivas associadas a cada dimensão, permitindo assim avaliar e classificar textos em função do seu grau de complexidade. A proposta de complexidade considera os seguintes aspetos como indutores de complexidade textual, que incluem a sua maior ou menor frequência no texto:

- Comprimento do texto (e.g., curto, longo)
- O propósito do texto (e.g., realização de tarefas quotidianas; lazer)
- O(s) conceito(s) mencionado(s) no texto (e.g., conceitos do dia a dia; noções especializadas),

- O(s) contexto(s) comunicativo(s) em que o texto circula (e.g., contextos do dia a dia, como bilhetes de transportes públicos; contextos mais específicos, como textos de opinião política)
- O léxico empregue no texto (i.e., relacionado com os conceitos e/ou com um domínio específico)
- O tipo de frase utilizadas no texto (e.g., frases com ordem não linear SVO)
- O tipo de formas e construções verbais no texto (e.g., tempos compostos, Modo Conjuntivo)
- A coesão textual (e.g., presença de conectores textuais e organização temporal)

Através da análise dos referenciais e de textos de cada nível, estes aspetos foram desdobrados e detalhados em descritores concretos (ver Anexo 1). Isto resultou na proposta de 81 descritores, contemplando os quatro níveis de complexidade (Monteiro et al., 2023). A noção de *descritores de complexidade textual* está relacionada com as manifestações concretas de complexidade no texto. O conjunto de descritores abrange dimensões textuais e permite uma análise da sua complexidade com maior granularidade.

No caso da dimensão léxico-conceptual, foram consideradas medidas tradicionais como a densidade lexical (i.e., frequência de palavras funcionais/gramaticais por palavras lexicais) e a relação entre complexidade lexical e a complexidade das ideias expressadas no texto (e.g., palavras atípicas – pouco frequentes – contribuem para tornar o texto mais abstrato (Fang & Pace, 2013). De igual forma, a linguagem figurativa também é considerada como indutora de complexidade, uma vez que sobrecarrega o processamento cognitivo do texto (Fang & Pace, 2013; Harley, 2013). A complexidade proveniente da linguagem figurativa está também relacionada com a concretude e grau de abstração da palavra: palavras que denotam conceitos mais abstratos são, geralmente, mais difíceis de processar do que palavras que denotam entidades concretas, especialmente para leitores de baixa literacia, uma vez que as primeiras não têm um referente concreto no mundo, mas as segundas têm (James, 1975). No entanto, em alguns casos, conceitos abstratos como *amor* ou *felicidade* contrariam este entendimento, possivelmente devido à familiaridade das/os leitoras/es com esses conceitos. Além disso, a frequência de conceitos abstratos nos textos é também um fator considerado (e.g., é o fator distintivo entre os descritores 21 e 22 ou a maior ou menor transparência e frequência de sentidos figurados, descritores 76 e 77).

Em relação à dimensão das estruturas sintáticas, esta proposta tem em conta medidas de *leiturabilidade* tradicionais, como a extensão da frase, responsáveis por induzir complexidade num texto, mas também manifestações de sintaxe complexa, de avaliação mais qualitativa, como estrutura não linear e ocorrências de subordinação e coordenação frequente e combinada. Para além disto, construções particulares, como frases passivas, são também consideradas como indutoras de complexidade. As construções na passiva tendem a manifestar-se em textos de carácter mais formal (Street & Dąbrowska, 2014). Por esta razão, pessoas mais proficientes nestes textos formais, isto é, com práticas de leitura mais desenvolvidas e referentes a uma maior diversidade de textos, têm mais facilidade em lidar com frases passivas do que indivíduos de baixa literacia.⁸

A flexão verbal rica de uma língua românica como o português conduziu ao entendimento de que esta dimensão teria de ser considerada independentemente. Assim, a complexidade verbal foi analisada à luz dos traços de tempo, modo e aspeto. Por exemplo, formas verbais no Infinitivo ou Presente do Indicativo são consideradas

⁸ Para além da frequência, as razões por detrás de a passiva ser indutora de complexidade podem ser de natureza distintas: Harley (2013) afirma que, para a passiva ser processada cognitivamente, esta requer que a pessoa conceba, primeiramente, o equivalente na frase ativa, para depois converter a frase de volta para a passiva. Por outro lado, Ambridge et al. (2016) destacam as propriedades semânticas do verbo, nomeadamente os seus papéis temáticos (*θ-roles*), como facilitadores ou agravantes na compreensão da estrutura passiva, sendo verbos com os papéis temáticos de Agente-Verbo-Tema os mais fáceis em compreender em frases passivas, e verbos com os papéis temáticos envolvendo Tema-Experienciador, os mais difíceis.

como menos complexas, comparativamente com formas no modo Conjuntivo. De igual modo, formas verbais compostas menos frequentes são mais desafiantes, independentemente da frequência de uso ou densidade conceptual da unidade lexical (e.g. “Quando chegar a altura do teste, ele terá estudado bastante”). Com efeito, isto pode dever-se à frequência de uso de dada forma verbal ou da associação desta com um género textual específico (e.g., o pretérito mais-que-perfeito associado a uma linguagem literária). A juntar a isto, as formas verbais, consoante o seu tempo, modo e aspeto, podem criar espaços temporais ou mundos possíveis objetivamente mais complexos e difíceis de processar cognitivamente (Carreiras et al., 1997). Por fim, a criação de sequências temporais que aludem a mais do que uma localização temporal, como acontece várias vezes com a utilização de tempos verbais compostos, conduzem a uma dificuldade na conceptualização da situação em questão (e.g., “O João tinha comido o bolo quando o Rui entrou em casa”). Os verbos são, assim, tratados como manifestações linguísticas muito flexíveis a vários níveis e que, por esta razão, podem ser geradores de uma complexidade particular.

Nesta senda, também as perífrases verbais foram contempladas. Expressões como “*estar a + Verbo no Infinitivo*”, “*começar a + Verbo no Infinitivo*”, etc., são manifestações linguísticas de aspeto gramatical, que, embora possuam uma natureza composta, se revelam bastante acessíveis para todos os falantes. A forma como o tempo e aspeto são codificados nas expressões linguísticas, seja por flexão verbal, seja por meio de advérbios ou através de perífrase verbal, parece revelar uma hierarquia de complexidade que segue o entendimento da complexidade textual de Dahl (2007). Este considera a quantidade de informação linguística embutida numa dada manifestação linguística como reveladora da sua complexidade: quanto mais informação linguística existir (e.g., marcação de género gramatical, tempo, aspeto), maior será a complexidade do enunciado/da forma em questão. Assim, tendo em conta que é através da flexão verbal que mais informação linguística é concentrada numa única palavra, e que as perífrases verbais, por vezes, diluem esta informação linguística, podemos considerar que a flexão verbal revela mais complexidade do que certas perífrases verbais.

Por último, a dimensão de coesão textual tem em conta a coesão referencial, que se reflete no uso de anáforas, elipses, assim como no tipo de relação com o antecedente. Além disto, também são consideradas as manifestações espaço-temporais, no texto. As/Os leitoras/es constroem um modelo mental das entidades presentes num texto, bem como da informação espaço-temporal associada a estas. Estas informações condicionam o processamento do texto por parte do indivíduo: quando mais emaranhada a estruturação de informações, maior dificuldade existe para compreendê-la. Efetivamente, um aumento de localizações espaço-temporais levam a que a compreensão do texto seja afetada (Glenberg et al., 1987).

Com todos estes fatores em mente, a nossa proposta foca-se, principalmente, na aferição qualitativa da complexidade de um texto, através de descritores concretos associados a manifestações de complexidade específicas. Contudo, uma dimensão quantitativa também é tida em conta, nomeadamente no que respeita à extensão do texto.

4.2. Resultados da validação pelo *focus group*

Como mencionado anteriormente, estas dimensões da complexidade textual foram desdobradas em 81 descritores concretos, agrupados em 12 categorias distintas: tamanho do texto; propósito; contexto comunicativo; diversidade conceptual; concretude; vocabulário/conjunto de léxico; morfologia; morfossintaxe; sintaxe; coesão textual: pronomes, anáforas e elipses; coesão textual: referências temporais; estilo. Estes descritores foram decorrentes da análise dos referenciais pela equipa do projeto e, posteriormente, validados por formadoras/es de EFA.

A participação e validação de profissionais de EFA contribuiu para um melhor entendimento de uma característica específica destes descritores: a sua frequência de ocorrência. Ao contrário da suposição de que uma

frequência mais alta de ocorrências de fenómenos indutores de complexidade aumenta a complexidade num texto, as reuniões com as/os formadoras/es EFA conduziram-nos à conclusão de que isto não era universalmente verdadeiro. Algumas características foram consideradas como indutoras da mesma complexidade, independentemente da sua frequência de ocorrência num texto. Ou seja, foi considerado que o nível de complexidade associado a estes traços seria o mesmo, independentemente do número de ocorrências no texto (mais/menos frequência não equivale a mais ou menos complexidade desse traço). Consequentemente, a nossa proposta inicial foi revista, de modo a privilegiar descrições centradas em características linguísticas observáveis, desligadas da frequência de ocorrência dos fenómenos, ainda que esse aspeto seja considerado posteriormente na aplicação prática do referencial em ferramentas de avaliação de complexidade textual. Porém, também foi reconhecido que existem exceções a este entendimento, podendo manifestar-se um *continuum* de frequência os seguintes descritores:

- A frequência do grau de concretude das palavras (descritores 20, 21)
- A frequência dos estrangeirismos presentes no texto (descritores 35, 36, 37).
- A frequência de orações relativas (descritores 53, 54, 55, 56)
- A frequência da presença frases passivas (descritores 57, 58, 59)

Os descritores mencionados acima enquadram-se no campo léxico-conceitual (os dois primeiros) e sintático (os dois últimos). A continuidade dos dois primeiros descritores segue em linha com a importância do conhecimento de 95% a 98% de palavras no texto, por parte do/a leitor/a (Schmitt et al., 2011). Ou seja, quando da consideração de complexidade lexical, existe uma margem – ainda que curta – para considerar a frequência de um dado fenómeno linguístico – neste caso, certas palavras. Por sua vez, os outros dois descritores, de natureza sintática, enquadram-se numa avaliação de *leiturabilidade* (Dale & Chall, 1948; Flesch, 1948; Gunning, 1952; Kincaid et al., 1975; Lorge, 1948; McLaughlin, 1969; Spache, 1953), cuja medida de aferição é o número de ocorrências presentes no texto.

O referencial de complexidade para adultos de baixa literacia aqui proposto é integralmente apresentado no Anexo 1 e textos exemplificativos de cada nível, provenientes do *corpus* iRead4Skills,⁹ estão presentes no Anexo 2.

5. Conclusão

Este trabalho apresentou uma proposta de referencial de complexidade textual para adultos, falantes nativos de baixa literacia, desenvolvida no âmbito do projeto iRead4Skills. Considera-se que os resultados e a metodologia aqui detalhados contribuem para uma melhor compreensão do fenómeno da complexidade textual, propondo uma abordagem multifacetada que combina dimensões quantitativas e qualitativas com a variabilidade de competências literárias de um público-alvo específico. Reconhecendo o fenómeno da complexidade como algo que diverge de indivíduo para indivíduo, a nossa proposta reconhece o *continuum* de competências em literacia e as diferentes posições dos leitores, através do estabelecimento de três níveis: Muito Fácil, Fácil e Claro.

A contribuição da nossa proposta assenta no estabelecimento de características linguísticas e de coesão textual num sistema de classificação da proficiência em leitura que visa adultos falantes nativos. Para além disto, a

⁹ Os textos exemplificativos são provenientes do *corpus* iRead4Skills. Todo o *corpus* está disponível, assim como outros recursos produzidos no projeto, em <https://zenodo.org/communities/iread4skills/records?q=&l=list&p=1&s=10&sort=newest>.

validação da proposta por profissionais de EFA – bem como a integração do seu conhecimento e experiência empírica – é um fator distintivo deste novo referencial, nomeadamente no que diz respeito à sua potencial aplicação em contextos reais de sala de aula. Em termos práticos, e no contexto do projeto iRead4Skills, a nossa proposta de complexidade contribui para o desenvolvimento de ferramentas inovadoras, tais como um sistema inteligente para a avaliação da complexidade de textos e uma ferramenta de adaptação de textos para níveis de complexidade adequados a públicos específicos. Contribui, também, para auxiliar profissionais da área a melhor compreender e avaliar competências e dificuldades de leitura das/os suas/seus formandas/os. Em termos teóricos, a natureza objetiva e a caracterização linguística detalhada dos descritores propostos contribuem para o avanço de conhecimento na área da complexidade textual e da literacia, permitindo também estudos comparativos entre diferentes populações-alvo.

No decorrer da elaboração deste trabalho, foram divulgados os resultados do Inquérito às Competências dos Adultos de 2023, desenvolvido pelo PIAAC, que colocam Portugal abaixo da média da OCDE no que diz respeito aos níveis de proficiência em literacia, no penúltimo lugar entre os países participantes (OECD, 2024a). Além disso, é importante realçar que as descrições dos níveis de proficiência presentes no guia-referência do PIAAC foram ligeiramente alteradas relativamente ao documento anterior (OECD, 2013b). Estas novas descrições são mais extensas e incluem elementos descritivos mais relacionados com o conteúdo dos textos. Por exemplo, no primeiro nível de proficiência (abaixo do Nível 1), há agora referências explícitas à consistência do uso de palavras e à forma como esta consistência contribui para o significado global da frase. Ao nível da frase, a adjacência dos conceitos e a forma como se relacionam entre si são também consideradas (OECD, 2024b). Estas revisões, que enfatizam a coerência interna e a conectividade semântica dos textos – em contraste com as anteriores, que se centravam mais nas tarefas que se esperava que os indivíduos realizassem em cada nível e na forma como a informação era apresentada no texto como um todo –, alinham-se com a abordagem e as conclusões da nossa proposta. No entanto, os níveis e descritores propostos neste trabalho são significativamente mais pormenorizados, oferecendo uma caracterização mais rica e mais fina da complexidade do texto. Isto assegura a nossa proposta como uma ferramenta valiosa e altamente relevante para responder às necessidades específicas de adultos com baixa literacia.

Em suma, o referencial de complexidade textual iRead4Skills representa um contributo relevante para o estudo sobre os desafios de leitura em adultos de baixa literacia, estabelecendo uma ponte entre a investigação teórica e a sua aplicação. O melhor entendimento dos fenómenos da complexidade textual e o seu condicionamento em adultos de baixa literacia proveniente deste trabalho espera-se benéfico para profissionais da educação e investigação, bem como em lugares de decisão política, que ambicionam promover o desenvolvimento das competências de literacia e potenciar as competências de leitura das pessoas, com retorno para a sociedade.

Agradecimentos

Este trabalho foi financiado pela Comissão Europeia (projeto iRead4Skills, número 1010094837, HORIZONCL2- 2022-TRANSFORMATIONS-01-07). Os pontos de vista e opiniões expressos neste trabalho são da exclusiva responsabilidade dos/as autores/as e não refletem necessariamente os da União Europeia ou da Agência de Execução Europeia da Investigação. Nem a União Europeia nem a autoridade que concedeu o financiamento podem ser responsabilizadas pelos mesmos. Parte deste trabalho é financiado por fundos nacionais através da FCT – Fundação para a Ciência e Tecnologia, I.P., no âmbito do projeto UID/03213 – Centro de Linguística da Universidade NOVA de Lisboa (CLUNL).

Referências

- ALTE (2002). *The ALTE Can Do Project (1992–2002)*. https://www.alte.org/can_do/alte_cando.pdf.
- Alves, J. (2005). Literacia e modernização económica. In H. Moura (Org.), *Diálogos com a literacia* (pp. 15–22). Lisboa Editora.
- Alves, A. C. & Arqueiro, A. S. & Dionísio, M. (2012). Vidas letradas de adultos em processo de Reconhecimento, Validação e Certificação de Competências. In I. Pereira, I. Dias, H. Pinto, H. Menino & R. Cadima (Orgs.), *I Conferência Internacional — investigação, práticas e contextos em Educação* (pp. 123–129). Escola Superior de Educação e Ciências Sociais – Instituto Politécnico de Leiria.
- Ambridge, B., Bidgood, A., Pine, J. M., Rowland, C. F., & Freudenthal, D. (2016). Is passive syntax semantically constrained? Evidence from adult grammaticality judgment and comprehension studies. *Cognitive Science*, 40(6), 1435–1459. <https://doi.org/10.1111/cogs.12277>
- Anderson, J.R. (2000). *Cognitive psychology and its implications*. W.H. Freeman.
- Beck, I. L., McKeown, M. G., Omanson R., & Pople, M. (1984). Improving the comprehensibility of stories: The effects of revisions that improve coherence. *Reading Research Quarterly*, 19, 263–277.
- Carreiras, M., Carriedo, N., Alonso, M., & Fernández, A. (1997). The role of verb tense and verb aspect in the foregrounding of information during reading. *Memory & Cognition*, 25(4), 438–446. <https://doi.org/10.3758/bf03201120>
- Correia, S., Amaro, R., & Monteiro, R. (2024). *iRead4skills – reading skills survey* [Relatório]. Zenodo. <https://doi.org/10.5281/ZENODO.10179536>
- Council of Europe. (2020). *Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume*. Council of Europe Publishing. <https://www.coe.int/lang-cefr>.
- Cunha, J. & Dionísio, M. (2019). O lugar da literacia nas ofertas de educação e formação de jovens em Portugal. In T. Dantas, M. Dionísio & M. Loffin (Orgs.), *Educação de jovens e adultos: Políticas, direitos, formação e emancipação social* (pp. 195–212). Editora da Universidade Federal da Bahia (EDUFBA).
- Cunningham, J. W., & Mesmer, H. A. (2014). Quantitative measurement of text difficulty: What’s the use? *Elementary School Journal*, 115, 255–269.
- Dahl, Ö. (2007). Definitions of complexity. *Proceedings of the Colloquium on complexity, accuracy and fluency in second language use, learning & teaching*, 41–46.
- Dale, E. & Chall, J. (1948). A formula for predicting readability. *Educational Research Bulletin*, 27(1). 11–20, 37–54.
- EDUSTAT – Fundação Belmiro de Azevedo. (2024). *Distribuição de adultos matriculados nos diferentes tipos de curso orientados para formação de adultos*. <https://www.edustat.pt/indicador?id=1>
- Fang, Z., & Pace, B. G. (2013). Teaching with challenging texts in the disciplines. *Journal of Adolescent & Adult Literacy*, 57(2). 104–108.
- Flor, M., Klebanov, B., & Sheehan, K. (2013). Lexical tightness and text complexity. In *Proceedings of the Workshop on Natural Language Processing for Improving Textual Accessibility* (pp. 29–38). Association for Computational Linguistics. <https://aclanthology.org/W13-1504/>
- Fransman, J. (2005). *Understanding literacy: A concept paper. Background paper for EFA global monitoring report*. UNESCO.
- Glenberg, A.M., Meyer, M., & Lindem, K. (1987). Mental models contribute to foregrounding during text comprehension. *Journal of Memory and Language*, 26.
- Goldman, S.R., & Lee, C.D. (2014). Text complexity. State of the art and the conundrums it raises. *Elementary School Journal*, 115, 290–300.

- Gunning, R. (1952). *The technique of clear writing*. McGraw-Hill.
- Hanemann, U. (2016). Shaping the literacy agenda from a lifelong learning perspective. *Adult Education and Development*, 83, 14–17.
- Harley, T. A. (2013). *Psychology of language: From data to theory*. Taylor and Francis.
- Heilman, M., Collins-Thompson, K., Callan, J., & Eskenazi, M. (2007). Combining lexical and grammatical features to improve readability measures for first and second language texts. *Proceedings of the NAACL Human Language Technology Conference*, 460–467.
- Hiebert, E.H. (2011). *Using multiple sources of information in establishing text complexity* [Reading Research Report #11.03]. TextProject, Inc.
- Hiebert, E. H., & Mesmer, H. A. E. (2013). Upping the ante of text complexity in the Common Core State Standards: Examining its potential impact on young readers. *Educational Researcher*, 42(1), 44–51. <https://doi.org/10.3102/0013189x12459802>
- Hu, M., & Nation, I. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language*, 13(1), 403–430.
- Kincaid, J. P., Fishburne, R. P. J., Rogers, R. L., & Chissom, B. S. (1975). *Derivation of new readability formulas Automated Readability Index, Fog Count and Flesch Reading Ease Formula for Navy enlisted personnel* [Research Branch Report 8-75]. Naval Technical Training.
- Lauffer, B. (2020). Lexical coverages, inferencing unknown words and reading comprehension: How are they related? *TESOL Quarterly*, 54(4), 1076–1085. <https://doi.org/10.1002/tesq.3004>
- Lorge, I. (1948). The Lorge and Flesch readability formulae: A correction. *School and Society*, 67, 141–142.
- McLaughlin, G. H. (1969). SMOG grading — A new readability formula. *Journal of Reading*, 128, 639–646.
- Mesmer, H. A., Cunningham, J. W., & Hiebert, E. H. (2012). Toward a theoretical model of text complexity for the early grades: Learning from the past, anticipating the future. *Reading Research Quarterly*, 47, 235–258. <https://doi.org/10.1002/rrq.019>.
- Monteiro, R., Amaro, R., Correia, S., Pintard, A., Gauchola, R., Moutinho, M., & Blanco Escoda, X. (2023). *iRead4Skills - Complexity Levels* (V1.0). Zenodo. <https://doi.org/10.5281/zenodo.10459090>
- OECD (2013a). *Technical report of the survey of adult skills (PIAAC)*. OECD Publishing.
- OECD. (2013b). *The survey of adult skills reader's companion*. OECD Publishing.
- OECD (2021). *The assessment frameworks for cycle 2 of the programme for the International Assessment of Adult Competencies*. OECD Publishing. <https://doi.org/10.1787/4bc2342d-en>
- OECD (2024a). *Do adults have the skills they need to thrive in a changing world survey of adult skills 2023*. OECD Publishing. <https://doi.org/10.1787/b263dc5d-en>
- OECD (2024b). *Survey of adult skills – Reader's companion: 2023*. OECD Publishing. <https://doi.org/10.1787/3639d1e2-en>
- Pintard, A., François, T., Nagant de Deuxchaisnes, J., Barbosa, S., Reis, M. L., Moutinho, M., Monteiro, R., Amaro, R., Correia, S., Rodríguez Rey, S., García González, M., Mu, K., & Blanco Escoda, X. (2024). *iRead4Skills Dataset 1: corpora by complexity level for FR, PT and SP* (2.1.) [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.13768477>
- Reis, M., Barbosa, S., Moutinho, M., Monteiro, R., Correia, S., & Amaro, R. (2024). Intelligent support for low literacy adults: The European Portuguese iRead4Skills corpus. *International Journal of Emerging Technologies in Learning (iJET)*, 19(08), 61–81. <https://doi.org/10.3991/ijet.v19i08.52023>
- Roths, L & Queirós, J. (2021). Práticas de leitura e competências de literacia: Resultados e ilações do PIAAC. *Entreler*, 1, 32–42.

- Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal*, 95(1), 26–43. <https://doi.org/10.1111/j.1540-4781.2011.01146>.
- Scribner, S., & Cole, M. (1981). *The psychology of literacy*. Harvard University Press.
- Spache, G. (1953). A new readability formula for primary-grade reading materials. *The Elementary School Journal*, 537, 410–413.
- Street, B. (1993). Introduction: The new literacy studies. In B. Street (Ed.), *Cross-cultural approaches to literacy* (pp. 1–21). Cambridge University Press.
- Street, J. A., & Dąbrowska, E. (2014). Lexically specific knowledge and individual differences in adult native speakers' processing of the English passive. *Applied Psycholinguistics*, 35(1), 97–118. <https://doi.org/10.1017/S0142716412000367>
- Stromquist, N. P. (2006). Women's rights to adult education as a means to citizenship. *International Journal of Educational Development*, 26(2), 140–152. <https://doi.org/10.1016/j.ijedudev.2005.07.02>
- UNESCO Institute of Statistics. (2020, 22 junho). *Literacy*. <https://uis.unesco.org/node/3079547>
- Valente, A. (Coord.) (2021). *Referencial de Competências-Chave de Educação e Formação de Adultos – Nível Básico*. ANQEP.
- Wagner, D. (2011). What happened to literacy? Historical and conceptual perspectives on literacy in UNESCO. *International Journal of Educational Development*, 42(3), 319–323.

Anexo 1 – Descritores por nível de complexidade

	Descritores	Muito fácil	Fácil	Claro	+ complexo
tamanho	1 O texto tem até 50 palavras.				
	2 O texto tem até 150 palavras.				
	3 O texto tem até 250 palavras.				
	4 O texto tem até 500 palavras.				

	Descritores	Muito fácil	Fácil	Claro	+ complexo
propósito	5 O texto serve como apoio para desempenhar tarefas familiares/quotidianas (e.g., comprar bilhetes de autocarro, identificar tipo e preço de produtos).				
	6 O texto apresenta informação nova (e.g., uma introdução sobre a importância da reciclagem como <i>A reciclagem é quando transformamos o lixo em coisas novas e úteis.</i>).				
	7 O texto tem o propósito de entreter o leitor em momentos de lazer (e.g., desporto, crónica, notícias).				
	8 O texto serve para aprender/realizar tarefas novas (e.g., instruções sobre como utilizar o e-mail da empresa).				
	9 O propósito do texto pode ser sobre vários tópicos de interesse do leitor, quer seja para o informar ou para lazer (e.g., artigo de revista sobre natureza).				

	Descritores	Muito fácil	Fácil	Claro	+ complexo
contexto comunicativo	10 O texto circula em comunicações do dia a dia (e.g., horários de transportes, ementas, instruções, preços de produtos, comunicações pessoais, etc.).				
	11 O texto é utilizado em contextos profissionais ou de comunicação dos <i>media</i> (e.g., comunicação ou instrução profissional, notícias, anúncios, desporto).				
	12 O texto é utilizado em contextos de lazer (e.g., literatura e ficção, reportagens, artigos de opinião).				
	13 O texto circula em contextos artísticos (e.g., literatura, poesia).				

	14	O texto é empregue em contextos especializados (e.g., comunicação profissional formal, fóruns <i>online</i>).				
	15	O texto é utilizado em contextos científicos e académicos (e.g., ensaios, relatórios, artigos científicos).				

Descritores		Muito fácil	Fácil	Claro	+ complexo
<i>diversidade conceptual</i>	16	O texto menciona conceitos relacionados com atividades do dia a dia (e.g., comida, lida da casa, roupa).			
	17	O texto menciona conceitos presentes em âmbitos específicos, mas que são de interesse do leitor (e.g., profissional, desporto, hobbies).			
	18	O texto refere conceitos que os leitores precisam de interligar para compreender totalmente (e.g., estrelas, galáxias, vulcões).			
	19	O texto faz referência a conceitos novos e multifacetados (e.g., matéria negra, desenvolvimento do cérebro, estatísticas).			

Descritores		Muito fácil	Fácil	Claro	+ complexo
<i>concretude</i>	20	O texto apresenta mais palavras que referem objetos concretos (e.g., <i>cadeira, maçã, martelo, computador</i>).			
	21	O texto contém poucas palavras que referem ideias abstratas (e.g., <i>felicidade, imaginação, amor</i>).			
	22	O texto emprega palavras que remetem para conceitos e ideias abstratas (e.g., <i>envelhecimento, razão, fidelidade</i>).			
	23	O texto usa palavras que referem conceitos e ideias abstratas em contextos multifacetados e/ou âmbitos especializados (e.g., <i>poesia, democracia, fusão nuclear</i>).			

Descritores		Muito fácil	Fácil	Claro	+ complexo
<i>vocabulário/ conjunto de léxico</i>	24	O texto não emprega abreviações.			
	25	O texto utiliza abreviações comuns após ter explicado o que significam (e.g., <i>l – litro, m – metro</i>).			
	26	O texto utiliza abreviações comuns (e.g., <i>m, km, g., l, Exmo.</i>).			

27	O texto não utiliza siglas.		
28	O texto apenas utiliza siglas e acrónimos conhecidos (e.g., <i>CP, IVA, RTP, SIC, TVI</i>).		
29	O texto utiliza siglas e acrónimos não conhecidos juntamente com o nome a que se referem (e.g., <i>Organização das Nações Unidas (ONU), Instituto Público (IP)</i>).		
30	O texto emprega todos os tipos de siglas (e.g., <i>ADN, UNESCO, EU, UNICEF</i>).		
31	O texto utiliza palavras e expressões simples relacionadas com assuntos do quotidiano, fazendo uso de substantivos frequentes (e.g., <i>batatas, azeite, legumes</i>).		
32	O texto tem palavras e expressões frequentes que fazem referência a assuntos do dia a dia, mas também a outros âmbitos específicos ou profissionais (e.g., <i>segunda-feira, curso, ano escolar</i>).		
33	O texto tem palavras e expressões um pouco menos conhecidas e frequentes que fazem referência a âmbitos específicos (e.g., <i>doença cardíaca, artéria, tratamento</i>).		
34	O texto utiliza palavras e expressões de âmbitos específicos, jargão ou termos arcaicos (e.g., <i>enfarte do miocárdio, artéria coronária, dissecação espontânea, genoma</i>).		
35	O texto não faz uso de estrangeirismos ou faz uso de estrangeirismos muito frequentes e conhecidos (e.g., <i>internet, donut</i>).		
36	O texto tem poucos estrangeirismos comuns (e.g., <i>site, stress</i>).		
37	O texto faz uso de estrangeirismos comuns frequentemente (e.g., <i>link, hobby, marketing, mail, gay</i>).		
38	O texto contém palavras polissémicas (palavras que possuem mais do que um sentido) que são ambíguas e/ou pouco frequentes. (e.g., <i>interno, representar, forma, paciente</i>).		
39	O texto contém palavras polissémicas (palavras que possuem mais do que um significado) ambíguas e/ou bastante frequentes (e.g., <i>base, experiência, leve, seca, aplicar, banco</i>).		

	Descritores	Muito fácil	Fácil	Claro	+ complexo
morfologia	40 O texto utiliza palavras compostas comuns (e.g., <i>feijão verde</i> , <i>couve-flor</i>) ou palavras compostas transparentes, ou seja, palavras ou expressões que podem ser facilmente compreendidas através da combinação das duas partes (e.g., <i>tira-caricas</i>).				
	41 O texto utiliza palavras compostas comuns e o seu sentido pode ser percebido pela combinação das duas partes (e.g., <i>óculos de sol</i> , <i>mal-educado</i> , <i>meia-idade</i> , <i>maus-tratos</i>).				
	42 O texto tem palavras compostas cujo sentido não é facilmente compreendido através da combinação das duas partes (e.g., <i>meio-termo</i> , <i>porta-voz</i> , <i>obra-prima</i>).				
	43 O texto possui várias palavras compostas com sentidos que não são facilmente compreendidos pela combinação das duas partes (e.g., <i>pós-graduação</i> , <i>mais-valia</i> , <i>pé-de-meia</i> , <i>cavalo de batalha</i>).				
	44 O texto tem palavras derivadas com prefixos e sufixos comuns e familiares em grande número (e.g., <i>infeliz</i> , <i>calmamente</i> , <i>organizar</i> , <i>trabalhador</i> , <i>casinha</i>).				
	45 O texto tem algumas palavras derivadas mais comuns, com prefixos e sufixos frequentes (e.g. <i>pré-cozinhado</i> , <i>desorganizar</i> , <i>embalagem</i> , <i>organização</i> , <i>jardineiro</i> , <i>finíssimo</i>).				
	46 O texto utiliza palavras derivadas com sufixos e prefixos variados (e.g., <i>nervosismo</i> , <i>lavável</i> , <i>curiosidade</i> , <i>altitude</i> , <i>superpoder</i> , <i>preconceituoso</i>).				
	47 O texto utiliza palavras derivadas com sufixos e prefixos menos comuns (e.g., <i>megatonelada</i> , <i>tendinite</i> , <i>geopolítico</i> , <i>neoliberal</i>).				

	Descritores	Muito fácil	Fácil	Claro	+ complexo
morfossintaxe	48 <u>Modos e Tempos verbais</u>				
	Verbos no infinitivo (e.g., <i>gostar</i> , <i>comer</i> , <i>fugir</i>).				
	Verbos no Presente do Indicativo (e.g., <i>gosto</i> , <i>comem</i> , <i>foge</i>).				
	Verbos no Pretérito Imperfeito do Indicativo (e.g., <i>gostavam</i> , <i>comia</i> , <i>fugias</i>).				
	Verbos no Pretérito Perfeito do Indicativo (e.g., <i>gostou</i> , <i>comi</i> , <i>fugiram</i>).				

	Verbos no Pretérito Mais-que-perfeito do Indicativo (e.g., <i>gostara, comeras, fugira</i>)		
	Verbos no Futuro Simples do Indicativo (e.g., <i>esquecerá, gostará</i>).		
	Verbos no Gerúndio (e.g., <i>gostando, comendo, fugindo</i>).		
	Verbos no Imperativo (e.g., <i>come!, falai!; não comas!; não falem!</i>).		
	Verbos no Condicional (e.g., <i>gostaria, comerias, fugiriam</i>).		
	Verbos no Presente do Conjuntivo (e.g., <i>gosta, coma, fujam</i>)		
	Verbos no Pretérito Imperfeito do Conjuntivo (e.g., <i>gostasses, comessem, fugisse</i>).		
	Verbos no Futuro do Conjuntivo (e.g., <i>se eu gostar, se tu comeres, se eles fugirem</i>).		
	Verbos no Pretérito Perfeito Composto do Indicativo (e.g., <i>tenho gostado, tem comido, têm fugido</i>).		
	Verbos no Pretérito Mais-que-perfeito Composto do Indicativo (e.g., <i>tinhas gostado, tinham comido, tinha fugido</i>).		
	Verbos no Futuro Composto do Indicativo (e.g., <i>terei gostado, terás comido, terão fugido</i>).		
	Verbos no Pretérito Perfeito Composto do Conjuntivo (e.g., <i>tenha gostado, tenha comido, tenham fugido</i>).		
	Verbos no Pretérito Mais-que-perfeito Composto do Conjuntivo (e.g., <i>tivesses gostado, tivessem comido, tivesse fugido</i>).		
	Verbos no Futuro Composto do Conjuntivo (e.g., <i>tiverem gostado, tiver comido, tiveres fugido</i>).		
	Verbos auxiliares <i>poder, ir, estar</i> no Indicativo (e.g., <i>posso gostar, vai comer, estão a fugir</i>).		
	Expressões com os verbos <i>permanecer, querer, poder, conseguir</i> em tempos simples (e.g., <i>permaneça calmo, queria almoçar, consigo ir</i>).		
	Expressões não muito extensas com os verbos <i>estar, ser, permanecer, querer, poder e dever</i> , em tempos compostos (e.g., <i>tenham estado ligados; devíamos ter acabado; temos de permanecer calmos; devíamos ir</i>).		

	Descritores	Muito fácil	Fácil	Claro	+ complexo
sintaxe	49 No texto, há mais frases curtas e diretas ou combinadas através de conjunções familiares (e.g., <i>A criança fugiu ontem. O pai trabalha na fábrica, mas a mãe fica em casa.</i>).				
	50 O texto combina frases através de conjunções coordenadas e subordinadas (e.g., <i>Os músicos estavam tocando, <u>mas</u> poucos clientes ouviam. <u>Quando</u> pararam de tocar, apenas um senhor se manifestou.</i>).				
	51 As frases no texto são frequentemente combinadas através de conjunções coordenadas e subordinadas e, por vezes, com conjunções menos comuns (e.g., <i>Apesar da confusão provocada pelos animais à solta, a situação foi controlada <u>sem</u> os animais serem atropelados.</i>).				
	52 As frases no texto são longas e compostas por diversas outras frases combinadas através de conjunções coordenadas e subordinadas (e.g., <i>Dado <u>que</u> grande parte deste livro fala na gestão das partes difíceis do trabalho a solo e analisa modos de as corrigir ou de as transcender, primeiro quero deixar bem claro que trabalhar sozinho é fantástico: é <u>por isso que</u> tantas pessoas o consideram uma opção.</i>).				
	53 O texto tem poucas orações relativas referentes ao sujeito da frase (e.g., <i>O senhor <u>que</u> vive na casa ao lado ouviu o barulho.</i>).				
	54 O texto tem orações relativas que referem o sujeito da frase (e.g., <i>O senhor <u>que</u> vive na casa ao lado ouviu o barulho.</i>).				
	55 As frases no texto incluem frequentemente orações relativas correspondentes ao sujeito da frase (e.g., <i>Francisca, <u>a quem os amigos tratam por «Kika», começou a surfar muito pequena.</u></i>).				
	56 As frases no texto incluem frequentemente orações relativas correspondentes ao sujeito ou ao objeto da frase (e.g., <i>O rapaz <u>que abraçou a mãe</u> dirigiu-se à entrada; O rapaz <u>que a mãe abraçou</u> dirigiu-se à entrada.</i>).				
	57 No texto, há mais frases ativas (e.g., <i>O cão comeu o bolo</i> (ativa) vs. <i>O bolo foi comido pelo cão.</i> (passiva)).				
	58 O texto tem poucas frases passivas (e.g., <i>A visita foi paga pelo patrão.</i>).				
	59 O texto apresenta frases passivas (e.g., <i>O almoço foi pago pela empresa</i>), incluindo construções impessoais comuns e passivas de -se (e.g., <i>O trabalho faz-se bem.</i>).				
	60 O texto tem frases simples apresentadas numa ordem básica e linear (sujeito-verbo-complementos) (e.g., <i>Eu ajudei os meus filhos a comprar casa.</i>).				

61	O texto contém poucas frases simples apresentadas numa ordem não linear (sujeito-verbo-complementos) (e.g., <i>À escola, chegaram os autocarros da visita de estudo.</i>).		
62	As frases do texto não estão necessariamente numa ordem básica e linear (sujeito-verbo-complementos) (e.g., <i>De ar puro, todos precisamos.</i>).		
63	As frases no texto são apresentadas em diversas ordens não lineares (sujeito-verbo-complementos; verbo-complementos-sujeito, etc.) (e.g., <i>Até final de março, porque antes seria impossível encontrar todos os ingredientes, haverá na ementa vários pratos de caça.</i>).		

Descritores		Muito fácil	Fácil	Claro	+ complexo
coesão textual: pronomes, anáforas e elipses	64 O texto tem nenhum ou poucos elementos omitidos (e.g., <i>O acidente causou dois feridos leves. Os bombeiros levaram os feridos para o hospital.</i>).				
	65 O texto tem elementos repetidos em vez de pronomes (e.g., <i>O João não gosta da Maria porque a Maria não é simpática.</i>).				
	66 O texto tem alguns elementos omitidos (e.g., <i>Gostava de fazer artesanato.</i> vs. <i>Eu gostava de fazer artesanato.</i>).				
	67 O texto usa pronomes (e.g., <i>O João não gosta da Maria porque ela não é simpática.</i>).				
	68 O texto apresenta elementos omitidos (e.g., <i>[] Mostrámos em estudos anteriores que a gordura solidifica quando [] arrefece.</i> vs. <i>Nós mostrámos em estudos anteriores que a gordura solidifica quando ela arrefece.</i>).				
	69 O texto apresenta pronomes e palavras relacionadas para indicar o que é referido (e.g., <i>No livro, nós damos uma definição completa de uma tempestade de neve e como esta se forma. Também indicamos onde o fenómeno pode ser visto.</i>).				

	Descritores	Muito fácil	Fácil	Claro	+ complexo
estilo	74 O texto não faz uso de sentidos figurativos (e.g., <i>Este verão os preços <u>subiram</u> vs. Este verão os preços <u>dispararam</u>.</i>)				
	75 O texto apresenta sentidos figurativos comuns (e.g., <i>Este verão os preços <u>dispararam</u>.</i>)				
	76 O texto apresenta sentidos figurativos menos comuns (e.g., <i>O Barcelona <u>varreu</u> o Real Madrid; O Estoril <u>esmagou</u> o Santa Clara.</i>)				
	77 O texto apresenta sentidos figurativos incomuns (e.g., <i>Os primeiros raios de sol <u>aconchegaram-me</u> as faces.</i>)				
	78 O texto utiliza apenas comparações, personificações e hipérboles mais comuns (e.g., <i>Ela é <u>linda como</u> uma flor; As minhas costas <u>estão a dar cabo de mim</u>; A tua mala <u>pesa uma tonelada</u>.</i>)				
	79 O texto faz uso da metonímia (e.g., <i>beber um copo</i>), assim como metáforas familiares (e.g., <i>Meu pai tem olhos de águia</i>).				
	80 O texto faz uso de metáforas incomuns (e.g., <i>A sociedade está doente; ele é um lobo em pele de cordeiro</i>).				
	81 O texto apresenta diversas figuras de estilo (e.g., <i>O cão da fábrica de curtumes acendia latidos fosforescentes nas noites de julho, quando o pólen da acácia me cobria as pálpebras, eu, morto de amores pela mulher de Sandokan, descobria-me unicórnio trancado na retrete da escola</i>).				

Anexo 2 – Textos exemplificativos por nível de complexidade

Exemplo de texto do nível Muito Fácil

Irlanda, 16 de novembro de 2014

Cara Ana,

Muito obrigado pelas fotos de Inglaterra que me mandaste! São muito bonitas.

Na Irlanda é tudo muito bonito. As pessoas cá são simpáticas, e já sei palavras e expressões novas que desconhecia. Este ano passamos bons momentos juntas. Agora seguimos caminhos diferentes. Vou mandar-te umas lembranças de cá como uns trevos e uns duendes (em peluche). Tens de vir ver a minha nova casa. Nas férias passa por cá e ficas aqui uma se maninha ou duas.

Dá notícias.

Um beijinho,

Bruna

PS. - Um dia destes vou aí visitar-te. Manda beijinhos ao Harry.

In

MAGMA studio. (2022). <https://magmastudio.pt/>

Exemplo de texto do nível Fácil

Pelas Florestas de Portugal

O verão está a chegar e o calor já começa a apertar. Felizmente, vivemos em Portugal, um país repleto de paisagens que concilia água e floresta e ainda nos oferece umas boas idas à praia. Estás pronto para entrar nesta aventura e conhecer melhor este nosso retângulo?

Vira a página!

In

Pelas florestas de Portugal. (2021). *Dá a Mão à Floresta*, 42, 3.

Exemplo de texto do nível Claro

Sentei-me no muro a comer uma bola de Berlim, uns instantes antes de caírem umas gotas grossas de chuva. As pessoas na praia começaram a arrumar as coisas apressadamente. Fechei os olhos por um instante, uma ou duas gotas caíram-me sobre a testa. O cheiro da areia molhada é diferente do petricor, aquele nome que dão ao cheiro da terra molhada. É um cheiro quente, metálico, de maresia misturada com pó. A propósito de nada lembrei-me daquela telenovela dos anos oitenta, baseada no romance do Sttau Monteiro, a "Chuva na Areia". Estendi a mão para que me caísse um pingo ou dois na palma, as pessoas corriam na areia, naquele movimento um bocado ridículo, em câmara lenta e de pernas presas na areia solta, alguns com um saco ou toalha na cabeça, a dizerem para os miúdos correrem. Continuei sentada no muro, limpei o açúcar do canto da boca, com o dedo indicador, que me esqueci de levar lenços e enquanto estendia de novo a palma da mão, muitos dos carros arrancaram, ficaram poucos mais as carrinhas dos surfistas, a praia esvaziou-se, o silêncio alargou o horizonte sobre o mar, e eu, sentada no muro, depois de uma semana cansativa, sem me chamar Guida nem agarrar mais uma vez o Verão, finalmente respirei.

In

Soares, C. N. (2023, 25 agosto). Agarra o Verão, Cristina. *Já disse que não quero nada*. <https://naoqueronada.blogs.sapo.pt/agarra-o-verao-cristina-2323>