

Framework of Textual Complexity for Low-Literacy Adults: Levels and Descriptors within the iRead4skill Project

Raquel Amaro

<https://orcid.org/0000-0002-4923-7186>; raquelamaro@fcsh.unl.pt
NOVA University (Portugal)

Susana Correia

<https://orcid.org/0000-0001-7340-7184>; susanacorreia@fcsh.unl.pt
NOVA University (Portugal)

Ricardo Monteiro

<https://orcid.org/0009-0001-8991-9130>;
ricardo.santos.monteiro@campus.fcsh.unl.pt
NOVA University (Portugal)

Alice Pintard

<https://orcid.org/0009-0003-3207-8849>; alice.pintard@gmail.com
UC Louvain (Belgium)

Michell Moutinho

<https://orcid.org/0000-0002-7665-2983>; michellmoutinho@fcsh.unl.pt
NOVA University (Portugal)

Sílvia Barbosa

<https://orcid.org/0000-0002-5617-278X> ; silviabarbosa@fcsh.unl.pt
NOVA University (Portugal)



© authors'

Suggested citation: AMARO, R., CORREIA, S., MONTEIRO, R., PINTARD, A., BARBOSA, S. (2025), Framework of textual complexity for low-literacy adults: levels and descriptors within the iRead4skills Project, *Langue(s) & Parole*, **10**, 57-120, <https://doi.org/10.5565/rev/languesparole.154>

Abstract

This paper presents the iRead4Skills framework of textual complexity, a novel approach to text difficulty tailored to low-literacy adult native speakers. Recognizing the pivotal role of literacy in personal and professional development, this framework addresses the challenges of identifying and adapting texts for this demographic. Integrating lexical, syntactic, and cohesion features, the framework delineates three levels of text complexity – Very Easy, Easy, and Plain – roughly aligned with the CEFR levels A1, A2, and B1. It bridges 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, language-specific adaptations for French, Portuguese, and Spanish, and its validation by experts in Adult Learning and Vocational Education and Training. By framing complexity as a continuum, it accommodates the heterogeneity of low-literacy adults while providing a scalable tool for educators and researchers. Practical applications of the 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: Textual complexity, complexity features, literacy, reading skills

Résumé

Cet article présente le cadre de complexité textuelle iRead4Skills, une approche novatrice de la difficulté textuelle adaptée aux adultes locuteurs natifs peu alphabétisés. Reconnaisant le rôle essentiel de la littératie dans le développement personnel et professionnel, ce cadre aborde les défis liés à l'identification et à l'adaptation des textes pour ce public. Intégrant des caractéristiques lexicales, syntaxiques et de cohésion, le cadre définit trois niveaux de complexité textuelle : Très facile, Facile et Accessible, globalement alignés sur les niveaux A1, A2 et B1 du CECR. Il relie la recherche théorique à la pratique en combinant des mesures quantitatives et qualitatives, garantissant une analyse robuste et contextuellement pertinente de la complexité textuelle. Parmi les principales contributions, citons l'alignement du cadre avec des textes authentiques, des adaptations spécifiques au français, au portugais et à l'espagnol, et sa validation par des experts en apprentissage des adultes et en formation professionnelle. En considérant la complexité comme un continuum, il s'adapte à l'hétérogénéité des adultes peu alphabétisés tout en offrant un outil évolutif aux enseignants et aux chercheurs. Les applications pratiques du cadre englobent des outils innovants tels qu'un système d'évaluation de la complexité textuelle et un assistant d'écriture pour l'adaptation de textes, comblant ainsi les lacunes en matière de supports de lecture adaptés aux adultes peu alphabétisés. Cette recherche offre des perspectives précieuses pour les études interlinguistiques sur la complexité textuelle, ouvrant la voie à de futures explorations des cadres de littératie et de complexité textuelle.

Mots clés : Complexité textuelle, caractéristiques de complexité, littératie, compétences en lecture

Resumen

Este artículo presenta el marco de iRead4Skills para la complejidad textual, un enfoque novedoso para la dificultad textual, adaptado a hablantes nativos adultos

con bajo nivel de alfabetización. Al reconocer el papel fundamental de la alfabetización en el desarrollo personal y profesional, este marco aborda los desafíos de identificar y adaptar textos para este grupo demográfico. Integrando características léxicas, sintácticas y de cohesión, el marco define tres niveles de complejidad textual: Muy Fácil, Fácil y Claro, aproximadamente alineados con los niveles A1, A2 y B1 del MCER. Conecta la investigación teórica con la aplicación práctica mediante la combinación de medidas cuantitativas y cualitativas, lo que garantiza un análisis sólido y contextualmente relevante de la complejidad textual. Entre sus contribuciones clave se incluyen la alineación del marco con textos auténticos, las adaptaciones lingüísticas específicas para francés, portugués y español, y su validación por expertos en aprendizaje de adultos y formación profesional. Al enmarcar la complejidad como un continuo, se adapta a la heterogeneidad de los adultos con bajo nivel de alfabetización, a la vez que proporciona una herramienta escalar para educadores e investigadores. Las aplicaciones prácticas del marco abarcan herramientas innovadoras, como un sistema de evaluación de la complejidad textual y un asistente de escritura para la adaptación de textos, que abordan las deficiencias en materiales de lectura adecuados para adultos con bajo nivel de alfabetización. Esta investigación ofrece información valiosa para los estudios interlingüísticos sobre la complejidad textual, allanando el camino para futuras exploraciones en marcos de alfabetización y complejidad textual.

Palabras clave: Complejidad textual, características de la complejidad, alfabetización, habilidades lectoras

Resum

Aquest article presenta el marc iRead4Skills de complexitat textual, un nou enfocament de la dificultat dels textos adaptat a parlants nadius adults amb baixa alfabetització. En reconèixer el paper fonamental de l'alfabetització en el desenvolupament personal i professional, aquest marc aborda els reptes d'identificar i adaptar textos per a aquest grup demogràfic. Integrant característiques lèxiques, sintàctiques i de cohesió, el marc delinea tres nivells de complexitat textual (Molt fàcil, Fàcil i Clar), alineats aproximadament amb els nivells A1, A2 i B1 del MEER. Uneix la recerca teòrica i l'aplicació pràctica combinant mesures quantitatives i qualitatives, garantint una anàlisi robusta i contextualment rellevant de la complexitat textual. Les contribucions clau inclouen l'alineació del marc amb textos autèntics, adaptacions específiques de cada llengua per al francès, el portuguès i el castellà, i la seva validació per part d'experts en aprenentatge d'adults i educació i formació professional. En emmarcar la complexitat com un continu, s'adapta a l'heterogeneïtat dels adults amb baixa alfabetització alhora que proporciona una eina escalar per a educadors i investigadors. Les aplicacions pràctiques del marc de treball inclouen eines innovadores com ara un sistema d'avaluació de la complexitat textual i un assistent d'escriptura per a l'adaptació de text, que cobreixen les mancances en materials de lectura adequats per a adults amb baix nivell d'alfabetització. Aquesta

investigació ofereix informació valuosa per a estudis interlingüístics sobre la complexitat textual, preparant el camí per a futures exploracions en marcs d'alfabetització i complexitat textual.

Paraules clau: Complexitat textual, característiques de complexitat, alfabetització, habilitats lectores

Introduction

Proficiency in reading is essential for individuals to thrive in today's world, reading being an essential component of literacy. High levels of literacy, encompassing oral reading fluency, reading comprehension, spelling, and writing, are strongly associated with numerous life benefits, including better employment opportunities and improved health habits (Street, 2005; OECD, 2013a). The significance of literacy and its role in society have been increasingly acknowledged as fundamental for achieving a decent quality of life. Literacy is regarded as a social practice, wherein individuals use their reading and writing skills in contexts that are relevant to their personal and professional development (Stromquist, 2006; OECD, 2021). The Organization for Economic Cooperation and Development (OECD) further underscores the connection between low literacy levels, unemployment among adults and social exclusion (OECD, 2013a).

The concept of literacy has evolved considerably over time. Initially limited to the notion of alphabetization, it expanded to encompass diverse perspectives based on specific societal contexts (Street, 2005). For an individual, achieving literacy is a complex issue, better described as a process than a state, i.e., an ongoing effort that needs to be trained and developed throughout life rather than a permanent or stable condition.

This understanding of literacy as a continuum of proficiency, varying from individual to individual, has driven discussions about what can enhance literacy in the adult population. Consequently, numerous adult learning (AL) programs with a strong emphasis on literacy as a basic skill have been developed. These programs aim to improve the qualifications of individuals with little or no formal education, enabling them to adapt to the demands of modern society and align with the principles of lifelong learning. However, identifying suitable reading materials for this demographic – particularly authentic and appropriate texts with low complexity – remains a significant challenge.

This was the challenge that motivated the iRead4Skills project. Recognizing the importance of reading practices and skills, the project seeks to promote the development of reading skills through the development of an intelligent system that evaluates text complexity and suggests appropriate readings for adults with low literacy levels, as supported by influential studies (Doronilla, 1996; Street, 2005). Moreover, the iRead4Skills project addresses a concern still frequently raised by professionals in AL centers: the difficulty of finding suitable materials for adult learners with low literacy and the impact of this gap in the motivation and success of this target population. Additionally, the project aims to develop a writing assistant to help create or adapt texts to appropriate levels of complexity and establish an open-access database serving as a catalogue of reading materials (books, magazines, newspaper articles, etc.) classified by complexity level. Ultimately, the goal is to facilitate access to suitable texts, thereby encouraging reading practices among low-literacy adults, fostering their personal and professional development. The successful accomplishment of the project requires several essential steps, ranging from the identification of the target audience and their needs to developing the intelligent system via Natural Language Processing and machine learning techniques.

The project's central aim is, thus, the definition of a linguistic framework for textual complexity tailored to low-literacy adults. This paper presents this framework, detailing the motivations behind its elaboration and the methodology employed.

To guide this effort, the study is driven by the following research questions: i) what are the most relevant linguistic features in gauging textual complexity in texts, for this target population? ii) how to design a framework of textual complexity specifically for low-literacy adult native speakers of Portuguese, French, and Spanish?

These questions lay the foundation for the development of the iRead4Skills framework.

Conceptual overview

The foundational concept of lifelong learning is that literacy operates on a *continuum*. Individuals occupy different points along this *continuum*, with variations between them (Goodman, 1996; Hanemann, 2012;

OECD, 2021; UNESCO, 2020). This notion also assumes that an individual's position on the continuum can shift, either upward or downward, implying an inherent potential for progress. This perspective not only supports the idea that someone with low literacy today can achieve higher literacy in the future but also reframes the understanding of what it means to be literate. In fact, determining whether the process of advancing along this continuum ever truly ends is challenging, as the “finishing line” continually evolves in response to societal and personal changes.

In this context, literacy is conceptualized as an information-processing skill encompassing oral reading fluency, reading comprehension, spelling, and writing. It is not merely a binary attribute that individuals either possess or lack; instead, it represents a spectrum of proficiency. However, the correlation between high levels of schooling and high literacy within a population remains elusive (Scribner, Cole, 1981; Rothes, Queirós, 2021). While a society with high levels of formal education may not necessarily exhibit higher literacy levels, schooling and specialized training address the core competences that literacy entails.

To define literacy comprehensively, it must be considered within its societal context (Street, 2005). This approach to literacy assumes that social practices are mediated by texts (and vice versa) within a specific temporal and spatial context (Barton *et al.*, 2000). In this view, literacy is perceived as it is practiced in everyday situations (Fransman, 2005). Consequently, understanding what it means to be literate is essential for addressing the challenge of defining text complexity and its manifestations in texts.

Complexity, defined as the inherent quality that hinders the process of retrieving meaning, can also be conceptualized as a *continuum* (Cunningham, Mesmer, 2014). From this perspective, a linguistic feature can occupy various points along a complexity scale. For instance, linguistic features¹ and values such as tense and mood (Gillie, 1957; Carreiras *et al.*, 1997) or syntactic structures with several dependencies (MacGinitie, Tretiak, 1971; Frazier, 1985) are likely to be related to complexity. This aligns with Dahl's (2007) interpretation of linguistic complexity, which

¹ *Feature* is here understood as a characteristic that is relevant for determining or inferring patterns or classes.

measures complexity based on the amount of information embedded in specific linguistic productions.

However, understanding how complexity manifests in texts, and measuring it objectively are two distinct challenges. Some approaches adopt a quantitative analysis of complexity, for instance relying on readability formulas (Goldman, Lee, 2014; Hiebert, 2011). Readability refers to the properties of a text that influence a reader's ability to identify, understand, interpret, and engage with printed or written materials, which depend on both intrinsic text features and reader-related factors (Bailin, Grafstein, 2016). Other approaches take a qualitative perspective, addressing the limitations of quantitative methods by simulating how a reader perceives and comprehends a text, for instance, through text schema (Anderson, 1995), mental reading models, story grammars, etc. (Harley, 2013).

Another crucial consideration is the potential to apply these approaches as a one-size-fits-all'. Frequently, they overlook the unique characteristics of specific population groups, or individuals, which can be significant. For instance, in reading, non-native speakers of a language often face greater grammar-based challenges compared to native speakers (Heilman *et al.*, 2007). Similarly, children typically exhibit lower reading proficiency than adults, primarily due to their limited reading experience (Mesmer *et al.*, 2012). In fact, reading habits tend to become richer and more advanced with age (Flor *et al.*, 2013).

Additionally, linguistic competence and language also differ significantly. These differences are evident between native and non-native speakers, as well as between adult and child native speakers. Consequently, assessing the reading proficiency of an adult native speaker using the same metrics applied to second-language learners or children can yield misleading results.

Understanding textual complexity as a subjective phenomenon requires acknowledging the idiosyncrasies of readers. This means recognizing that readers possess varying levels of literacy, which directly influence their interactions with texts.

The complexity framework we propose addresses text complexity in relation to low-literacy adult native speakers and examines how this

complexity manifests or can be detected in texts. We posit that the most effective approach combines both qualitative, and quantitative measures, as these methods are complementary. Quantitative measures lack “real meaning” until they are interpreted within a qualitative context (Hiebert, Pearson, 2014). Conversely, qualitative approaches are inherently holistic and somewhat impressionistic, as they do not explicitly identify the specific complexity features being considered (Pearson, Hiebert, 2014). By integrating these measures, we do not only enhance the validity of results but also provide an innovative method for identifying and analyzing text complexity features that remain unexplored (Hiebert, Pearson, 2014; Goldman, Lee, 2014).

Methodology

This section outlines the methodology employed to develop a framework for textual complexity tailored to low-literacy adults. In alignment with the overarching goals of the iRead4Skills project, the methodology comprised three key steps:

- i. Compiling and evaluating existing guidelines and frameworks relevant to text readability and complexity assessment and, from there, identifying the pertinent descriptors related to the assessment of reading skills and text complexity.
- ii. Proposing an initial framework for three levels of text complexity suitable for adult native speakers with low literacy skills. This framework, aligned with the project’s objectives, included the collection of specific linguistic features to provide a detailed and objective characterization of the targeted text complexity levels.
- iii. Implementing a focus group with professionals with expertise in adult education to validate the proposed framework. Their feedback was used to refine and further adjust the framework to ensure its relevance and applicability.

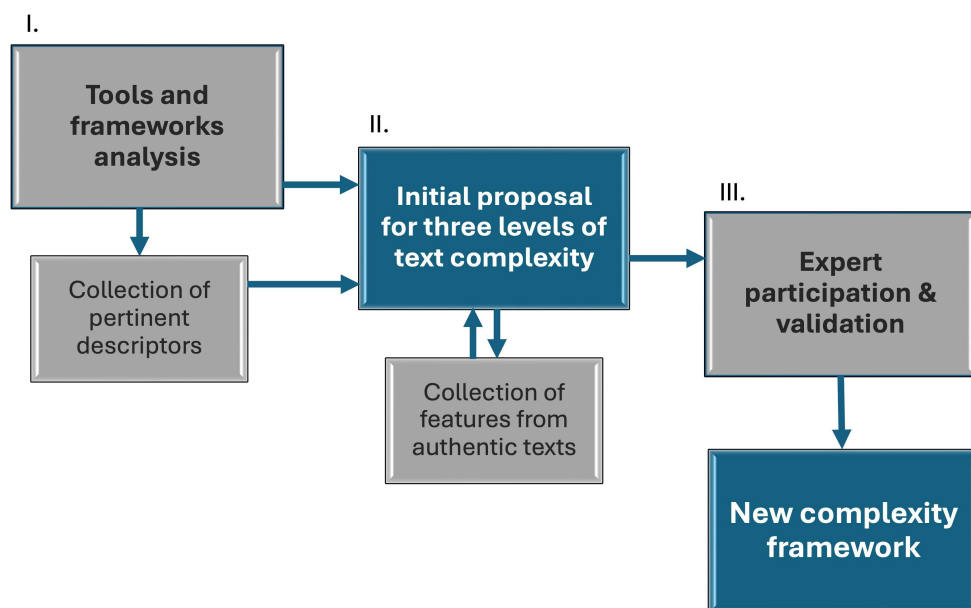


Figure 1. General process

Analysis of guidelines and frameworks

To define the different textual complexity levels relevant for low-literacy adults, we conducted a review of existing frameworks and benchmarks related to adults' key competences such as language skills and linguistic competence, which included descriptors for literacy and reading skills. Four key documents were analyzed to achieve a cross-linguistic understanding of reading proficiency: the Common European Framework of Reference for Languages (CEFR) (CEFR, 2020); the Programme for the International Assessment of Adult Competencies (PIAAC) (OECD, 2013b); the Can Do Project – Association of Language Tests in Europe (ALTE) (ALTE, 2002); and the Benchmark of Key Competences in Adult Education and Training (Valente, 2021), from the Portuguese National Agency for Qualification and Vocational Education and Training (ANQEP), which assesses language-related skills in adults attending AL Centers.

These four literacy frameworks were chosen based on their prominence in language assessment, relevance to adult literacy, and alignment with the project's goals. The decision was informed by a systematic review of the existing literature, focusing on frameworks widely recognized for addressing the complexities of adult literacy across disciplines. Additionally, given that the project is conducted within the

contexts of Portuguese, French, and Spanish languages, this linguistic and cultural dimension played a significant role in framework selection to ensure applicability and relevance across these language communities. While other frameworks exist, we prioritized those offering the most comprehensive and adaptable approaches suitable for our particular multilingual context.

In analyzing these referentials, special attention was paid to the initial levels described, for two primary reasons: Firstly, the broad and intricate concept of literacy and the varying proficiency levels outlined in these documents make it difficult to determine which frameworks are appropriate for individuals with low literacy. Unlike formal schooling levels, there is no universally established formula or index to assign a literacy level to individuals.

Consequently, our approach frames low literacy within the context of societal life (Street, 1993; Fransman, 2005). From this perspective, literacy is assessed in terms of its pragmatic goals (e.g., argumentation, conveying of news, etc.), and the contexts in which it is practiced (e.g., professional, familial, formal, etc.). These factors both shape and are critical to evaluating an individual's literacy level.

Unlike political and institutional actors, who often define literacy in quantifiable and instrumental terms – as a skill measurable across populations and correlated with employment, economic growth, and productivity –, we acknowledge these associations as reminders of the real consequences of low literacy, but do not build our work on them. A sociocultural perspective, as well as language-specific one, is here considered, with the aim of seeing our object of study through a critical lens that values culturally-situated literacy practices, particularly among marginalized adult learners.

Secondly, considering literacy as a flexible and scalable concept led us to recognize that adults with low literacy are not confined to a single proficiency level. Instead, variability and heterogeneity within this population must be acknowledged, as low literacy encompasses a range of incremental levels. The initial levels outlined in the analyzed benchmarks align with the literacy skills and text genres/types expected for adults with low literacy. For instance, in the CEFR framework, the B2 level – the level

immediately above those considered in this project (A1-B1) – describes an individual capable of “*reading articles and reports concerned with contemporary problems in which the authors adopt particular attitudes or viewpoints*” (CEFR, 2020, 117), which requires advanced reading skills. Conversely, Pre-A1 corresponds to an early stage of alphabetization and very basic reading skills, prior to autonomous reading activities. Thus, reading proficiency at these extremes is not relevant to the target population of low-literacy adults. From the analysis of the existing benchmarks and frameworks, the CEFR levels A1, A2 and B1 were identified as relevant for our framework. The focus, therefore, was on understanding the characteristics of each level and determining what adaptations were necessary to make them suitable for low-literacy adults.

Complexity levels initial proposal

The textual complexity characteristics associated with each level were compiled from the four documents analyzed and adapted to low-literacy contexts. To minimize reliance on subjective descriptions, the properties described in the frameworks (e.g., “*simple texts used in everyday contexts to perform familiar tasks*”) were mapped, whenever possible, to objective linguistic properties (e.g., “*verbs in simple indicative tenses*”). These features thus serve as specific descriptors for each level.

Given that our target audience consists of native speakers, and that most existing frameworks are designed for second- and/or foreign-language learners, the compiled descriptors were tested against authentic texts to: i) assess prototypicality (i.e., whether a given feature occurs frequently or rarely within and across levels); ii) determine whether the presence of a given feature contributes to complexity or does not significantly affect it; and iii) identify additional features not described in existing frameworks.

This process also contributed to the development of the project’s corpus. The goal was to provide an initial validation of the proposed levels and descriptors through the analysis of texts and the identification of the level-specific descriptors within them. We followed an *in tandem* approach in which one team selected and initially classified the texts, while another team validated the classifications. This collaborative workflow allowed for constant communication and iterative refinement of both tasks. As a result, the teams developed a clearer understanding of which texts

corresponded to the proposed levels, leading to more efficient corpus compilation and validation. Additionally, involving multiple contributors in each team reduced the impact of individual bias.

The classification and validation process involved the following steps:

- i. Preliminary classification: A text was initially assigned to one of the three levels, based on a “first glance” assessment of its complexity.
- ii. Validation process: The validation team reviewed the text by checking the presence of descriptors corresponding to the initially assigned level, as well as descriptors from other levels, if present. A table was created to code the occurrence of the descriptors across the three levels (0=not observed; 1=observed). The total number of observations was used to determine whether the text was appropriately classified, required reclassification to a different level, or should be discarded.

As part of this process, a fourth complexity level, termed “More Complex” was introduced. This level established an upper threshold for complexity in relation to the levels targeted by the project.

This procedure resulted in an initial proposal for describing the three relevant levels for the project, including a detailed set of associated features for each level. Starting with the proposal developed for Portuguese, the project team – comprising experts in linguistics, complexity assessment, and low-literacy contexts – extended the framework to the project’s other target languages, French and Spanish.

Expert validation

To further validate the devised complexity framework, we collected input from teachers and trainers from AL centers in Portugal, France, Belgium, and Spain. Given the expertise of these professionals and their experience teaching and addressing the challenges faced by low-literacy adults, we consider this validation process essential for the success of project.

The selection of these validation teams was deliberate to include educators from natural sciences and mathematics because adult literacy in

these disciplines often presents unique challenges related to content-specific language and literacy demands. These trainers, although not so versed in language-teaching, continue to bring classroom experience and knowledge of disciplinary discourse and instructional practices, which complements our expertise. Their involvement ensures that the framework is not only theoretically sound but also applicable within adult literacy settings across diverse content areas.

The validation process involved gathering insights on text complexity and its descriptors through focus groups of experts for each language. These experts were recruited from AL and VET centers in Belgium, France, Portugal and Spain, where they worked as teachers in various fields, including languages (French, Portuguese or Spanish), natural sciences, mathematics, ICT, among others. The input of these professionals was obtained through discussions following a standardized protocol across the three languages. The process consisted in three main steps:

1. Preparatory task (approximately 1 hour): prior to the first focus group meeting, participants were assigned a task (administered as an online survey, cf. Monteiro *et al.*, 2023) designed to validate and gather their opinions on the descriptors. The task included i) a brief presentation of the complexity levels targeted by the project, without disclosing the associated descriptors, and ii) a random presentation of the descriptors, with the participants tasked to match each descriptor to one (and only one) of the presented complexity levels. To ensure accessibility for participants without a background in linguistics or grammar (e.g., mathematics teachers), the descriptors were paraphrased using minimal specialized terminology and accompanied by one or more illustrative examples. Participants were asked to complete the survey within 10 days of receiving it.
2. Focus group session 1 (approximately 1.5 hours): the first focus group session aimed to discuss the descriptors and the results of the preparatory task. Before the session, the preparatory task results were analyzed to create a table categorizing descriptors. The session included i) covering the

concepts of complexity, text complexity, and readability, as well as the framework's complexity levels; ii) short exercises using texts that did not perfectly align with the descriptors to stimulate discussion; iii) a detailed discussion of the preparatory task results and the complexity level descriptions, focusing on descriptors that showed low agreement among participants during the preparatory task.

3. Follow-up session & focus group session 2 (approximately 0.5 hour): Following the analysis of the discussions in the first focus group sessions, participants were consulted to confirm the finalized version of the complexity level descriptions. The final descriptors for each level, incorporating feedback from the focus group discussions, were sent to participants along with example texts for each level. Participants were invited to validate the finalized descriptors or suggest minor adjustments.

Lastly, it is important to note that, in addition to the validation stages involving AL trainers, AL students also participated in later phases of the validation process. Given the complexity of the earlier stages, which were focused on the conceptual development of the framework, the input of AL trainers was prioritized due to their pedagogical expertise and familiarity with instructional design. However, in subsequent stages, which centered on assessing the framework's practical applicability, feedback was gathered from AL students, who represent the project's primary audience. These later validation phases, while essential to the overall project, fall outside the scope of this paper.

Results and discussion

In this section, we present the results obtained following the previously presented methodology. The analysis conducted through the integration of existing guidelines and frameworks and textual features culminated in the definition of three distinct complexity levels: Very Easy, Easy, and Plain. These levels aim to categorize texts based on their accessibility to readers with varying degrees of literacy.

A comprehensive description of each level is provided. The descriptions serve not only as a classification framework but also as a foundation for understanding the factors that contribute to text complexity.

These levels were developed with a focus on textual elements that impact comprehension, including lexical, syntactic, and cohesion features.

Through this approach, the iRead4Skills framework offers a nuanced perspective on text complexity, bridging theoretical insights with practical relevance, and aligning with language-specific needs for French, Portuguese, and Spanish. This section provides an in-depth discussion of these findings, their implications, and their applications, supported by detailed descriptors and expert validation.

iRead4Skills text complexity levels

The analysis of existing guidelines and frameworks, together with the collection of features from authentic texts, resulted in three different levels of text complexity: *Very Easy*, *Easy* and *Plain*. These can be briefly described as presented in Table 1.

<i>Level name</i>	<i>Brief description</i>
Very Easy	Texts that are fully or almost fully understood by everyone, including people with very low schooling (i.e., that did not finish primary school (ca. 6th year)) and almost no reading experience.
Easy	Texts that are fully or almost fully understood by people with low schooling (i.e., that completed primary school but do not have more than the 9th year) and with poor reading experience.
Plain	Texts that are understood the first time they are read by people who completed the 9th year and have a functional-to-average reading experience.

Table 1: Complexity levels informal description

It is important to note that, in these informal descriptions, each level refers to a specific educational stage, providing a more objective and familiar reference framework for validation purposes. However, the aim is not to characterize the readers themselves, but rather classify the texts, suggesting the education level most appropriate for texts of a given complexity level. As further discussed below, the complexity levels are defined based on textual complexity.

The elaboration of the descriptions was also grounded in the notion that complexity in texts can pose three types of comprehension challenges for readers: superficial, knowledge and content problems (Beck *et al.*,

1984). Superficial problems pertain to strictly linguistic issues, such as complex syntactic structures or inconsistent use of conjunctions within the text. Knowledge problems are linked to an extralinguistic dimension, arising from the reader's lack of prior knowledge about the topic(s) addressed in the text. Lastly, content problems stem from the way ideas are structured, interwoven, and presented within the text, that is, how the text is organized and conveyed. Contradictory concepts, counterintuitive conclusions and ideas that are difficult to infer can create significant comprehension challenges for readers.

Building upon these textual problems, our proposal analyzes textual complexity across four dimensions: lexical-conceptual aspects, syntactic structures, verb forms and cohesion. These dimensions differ from those outlined in existing benchmarks for reading proficiency, which target a different audience. For instance, non-native speakers tend to experience complexity related primarily to grammatical issues (Heilman *et al.*, 2007).

The lexical-conceptual dimension is the most relevant dimension in our proposal. Lexical comprehension difficulties are the first to arise when a reader engages with a text (Harley, 2013). As highlighted in the literature from its early stages (e.g., Lively, Pressey, 1923, to Pireli *et al.*, 2023), the ability to understand a text is severely impaired if the reader does not know the meaning of the words it contains. Several studies demonstrate a direct correlation between lexical coverage and reading comprehension (Schmitt *et al.*, 2011; Hu, Nation, 2000; Laufer, 2020), showing that achieving a 70% level of comprehension requires knowing between 95% to 98% of the words in the text (Schmitt *et al.*, 2011). This means that even a very small proportion of unknown words in a text can significantly hinder comprehension.

Low-literacy adult native speakers, while proficient and exposed to a variety of language usages and contexts, have access to a broader range of constructions than children or second-language learners. However, readers with higher proficiency in reading “formal texts”, that is, those with more developed and diverse reading practices, encounter fewer issues with specific syntactic structures compared to low-literacy readers².

² Besides reader's frequency of contact with the passive voice, the reasons behind the complexity-inducing nature of this feature can vary: Harley (2013) states that, in order to process the passive voice, the reader must firstly process the equivalent in the active voice and then convert it back to

Consequently, the framework we propose considers syntactic readability measures, such as sentence length, as contributors to text complexity, alongside qualitative manifestations of complex syntax.

Regarding cohesion, our proposal examines referential cohesion through anaphoric relations and the temporal markers within the text (Fang, Pace, 2013; Glenberg *et al.*, 1987).

Finally, verb forms and verb constructions are a distinct dimension in our framework. Although these could be viewed as part of the lexical and syntactic domains, we argue that verb forms integrate aspectual, temporal and modal information that pertain to verbal morphology alone, thus contributing to complexity in a unique way and deserving independent consideration in this work³.

Following this approach, and as outlined below, each level encompasses detailed descriptions that also consider language-specific characteristics. This implies that some features defining each level may vary between the languages targeted by the project (French, Portuguese, and Spanish).

Very Easy level

Short (approximately 50 words) **and simple texts** for the purpose of performing familiar tasks OR short and simple texts introducing new information (e.g., didactic texts).

Typically, simple and everyday concepts. It is assumed that the speaker has limited access to communication domains. Ideally, the topic is presented earlier.

Basic communication contexts: basic day-to-day (transport schedules/lists, menus/general instructions, price information); family/personal communications; simple/basic information.

Absence of figures of speech.

Basic lexicon (active lexicon): known words and simple expressions memorized and used in everyday matters (e.g., transport, food, family, work). Frequent and concrete main and copulative verbs and frequent and concrete nouns, that is, concepts/ideas with a higher level of concreteness than abstraction. Rare affixation (except frequent affixes such as FR *dé-*; *-age*; PT *-mente*; SP *-mente*, *-ción*, *re-*).

the passive voice. On the other hand, Ambridge *et al.* (2016) highlights the semantic properties of the verb, namely its thematic roles (θ -roles), as facilitating or aggravating factors in the processing of the passive voice, with verbs with thematic roles of Agent-Verb-Theme as easier to process than verbs with roles Theme-Experiencer.

³ This understanding is even more pertinent when considering languages with high verb inflectional paradigms.

Short periods with simple conjunctions and in direct word order (Subject-Verb-Object).

Rare auxiliary verbs (except for copulative verbs) and few anaphoric references: the referential chain is complete and does not occur in an elliptical way (e.g., ‘Tina doesn’t like her because she’s not nice’ vs. ‘Tina doesn’t like her for [-] not being nice’.).

Coordination structures (Noun Phrase & Noun Phrase, Adjective & Adjective...): copulative, disjunctive and adversative conjunctions are admitted. Some frequently used subordination structures (subordinate temporal adverbials, for example), except for those less frequent (reduced adverbials of infinitive), are admitted.

Occurrence of some **periphrastic constructions**, more usual and therefore more decipherable (e.g., FR: *aller à, penser à, venir de, jouer de, parler de*; PT: *estar a, começar a, andar a, deixar de, acabar de, ir*; SP: *estar a, empezar a, ir a, dejar de, acabar de*). No compound tenses.

Indicative verb tenses. Personal Infinitive and Gerund are admitted.

Simple temporal location. Temporal cohesion is given by means of temporal adverbs or connectors (today, tomorrow, before, after ...) and not by verb tenses (e.g., ‘She left before John arrived.’ vs. ‘John arrived when she had already left’).

Easy Level

Short texts (approximately 150 words) that are interesting for readers to inform themselves, or in moments of leisure, or with the purpose of carrying out tasks.

Some presence of abstract concepts (such as feelings, states of mind, religiosity, qualities, and defects, etc.). Concepts linked to the world personally and professionally familiar to the reader.

Usual communication contexts: work context (specific instructions); media (news of interest, e.g., sports); commercial communication (ads).

Commonly used figures of speech (e.g., ‘the prices skyrocketed’).

Basic lexicon (active lexicon) and expanded passive lexicon with frequent words. Main and copulative verbs and frequent nouns in different domains where the reader routinely interacts or is interested in. Some affixation, frequent and productive prefixes, and suffixes (e.g., FR: *ex: re-, dé-, in-, micro-, -age, -tion, -ment, -ité*; PT: *-agem, -ção, -eiro, -mente, -ade, -íssimo, -inha, -ice*; SP: *-aje, -ción, -ero, -mente, -ísimo, -izar, -sub, super-, in-*).

Short periods, with coordinated conjunctions and most of the subordinate conjunctions, both in direct (Subject-Verb-Object) and indirect word order. Subject relative subordinate clauses are admitted, but not object clauses (e.g., ‘The boy who hugged his mother.’ vs. ‘The boy who his mother hugged.’). Subordinates with Indicative, Subjunctive, and Infinitive are admitted.

Verbs in simple tenses, including simple future (but not frequently). Some **periphrastic constructions** such as the passive voice (especially in the Indicative) are admitted. Compound tenses are present (e.g., FR: Plus-que-

parfait (<i>avait mangé</i>); PT: Pretérito Mais que Perfeito Composto (<i>tinha comido</i>); SP: Pretérito Pluscuamperfecto (<i>había comido</i>)). More complex temporal references in linear sequence are present. Temporal cohesion can be given via verb tenses. The reader can link different parts of the text and make a global sense of them.
Plain Level
Texts of different sizes (approximately 250 words) and on varied topics of interest to the reader for information or leisure. Varied concepts, not necessarily familiar to the reader. More contact with the online world. The reader can infer at a more complex level (e.g., infer opinions from opinion texts), including at a multimodal level, with texts and information in less common formats (e.g., infographics). Various communication contexts: leisure (stories; travel diaries, fiction); professional (reports, instructions); media (reportage, opinion articles); online (forums) communication contexts. Varied lexicon to express subjects in any of the communication domains. More active lexicon, due to diverse contact and familiarity. Presence of polysemic words. Occurrence of frequent foreign words and borrowings (e.g., <i>timing, hobby, show</i>). Nouns that express both concrete and abstract concepts to describe situations, reactions, emotions, thoughts, etc. are admitted. Some frequent domain-specific verbs and nouns, describing trendy situations, or situations known to the reader, occur. Some specific main and auxiliary verbs are admitted. The occurrence of most affixations (e.g., FR: <i>a-, auto-, co-, ex-, inter-, sur-, -able, -esse, -eur, -iste</i>; PT: <i>-ite, -itude, -ume, -vel, -oso, -ismo</i>) is admitted, except for erudite and less frequent affixes. Presence of frequent compound words which can have their meanings retrieved from each component. The reader can infer the meaning of derived words and some more frequent and non-idiomatic compound nouns. Longer periods, with simple and compound sentences and a greater variety of conjunctions and syntactic order. Some high-frequency irregular verbs are admitted. Presence of modal verbs, with uses and meanings in common expressions and in unusual contexts. Indicative, Subjunctive and Imperative and Conditional moods, both in the active and in the passive voice, are admitted. FR: Passive structures with <i>se faire</i>; PT and SP: Passives with <i>-se</i>. Complex temporal reference in non-linear sequence. The reader can infer information (albeit basic) that is not explicit in the text.

Table 2: *iRead4Skills Complexity Levels General Description*

Descriptors and features

As reflected in the previous section, our proposal considers the following aspects in the analysis of text complexity:

- length of the text (e.g., short; long),
- purpose of the text (e.g., performing familiar tasks; provide information on leisure),
- domains and topics addressed by the text (e.g., everyday concepts; technical information),
- communication contexts in which the text is typically used (e.g., daily activities; media such as opinion articles; literature),
- conceptual range covered by the texts (e.g., simple everyday concepts; familiar abstract concepts; domain-specific concepts),
- lexicon used (i.e., the actual set of words occurring in the text),
- type of sentences used (e.g., simple, direct; long, coordinated; long, subordinated),
- type of verb inflection and constructions used (e.g., verb periphrases, compound tenses, modal verbs),
- structure of the text (e.g., temporal sequences, argumentation).

The lexical-conceptual aspect considered the traditional measures of lexical rarity (word frequency and the presence of simple/complex words), as well as lexical density and the relationship between lexicon complexity and complexity of the ideas expressed in a text. It is understood that atypical wordings contribute to rendering a text more abstract (Fang, Pace, 2013). Figurative language is also accounted for and measured as a factor inducing complexity, as it imposes an additional cognitive processing load (Fang, Pace, 2013; Harley, 2013). The complexity of figurative language at lexical level can be assessed based on the concreteness of the words used. Less abstract concepts are generally harder to process cognitively than concrete ones, especially for low-literacy readers, as concrete concepts have a material referent in the real world, whereas abstract concepts do not (James, 1975). However, familiar abstract concepts (e.g., *love*, *happiness*), are not typically difficult to understand.

The syntactic structures aspect of our framework considers traditional readability measures – such as sentence length – as contributors to complexity, as well as qualitative manifestations of complex syntax, such as non-linear structures and frequent or intricate subordination and coordination. Moreover, specific features such as the passive voice are considered complexity-inducing elements. Sentences constructed using the passive voice are often associated to more formal texts (Street,

Dąbrowska, 2014). Structures such as verb periphrases and modality constructions are similarly analyzed for their potential to increase text complexity.

The rich verb inflection systems of the target Romance languages in this project necessitated an independent measure. Verbs complexity was analyzed in relation to tense and mood. For instance, common and simple verb forms, such as the infinitive or the present indicative, are considered less complex, while less common verb moods, such as the subjunctive, or compound verb forms, are more challenging for readers. These forms can express temporal shifts and hypothetical constructs that impose additional cognitive processing demands. Verbs are, therefore, regarded as highly flexible linguistic manifestations capable of generating complexity in particular ways. For instance, verb forms inflected in the present tense, or the infinitive mood are seen as simple when compared to compound forms or the subjunctive mood, regardless of the verb's frequency or conceptual weight. This is primarily because many complex verb forms are relatively rare or associated with specific genres (e.g., the pluperfect tense is linked to literary discourse). Additionally, these forms often create temporal spaces that involve intricate temporal relations (e.g., “John had eaten the cake when Mark opened the door”), making it more challenging to conceptualize the situation described.

In line with this, verbal periphrases, commonly used in romance languages such as Portuguese and Spanish, were also included in our framework. These expressions are typically composed of an auxiliary verb, a preposition, and a main verb in the infinitive (e.g., “to be about to/to start to” + Verb Infinitive). Despite their compound nature, several verbal periphrases were considered generally easy for most readers (e.g. *passé composé* in French or verbal periphrases expressing the future - *aller/ir/* + infinitive, cf. Appendixes II, III and IV). The way tense and aspect are encoded – whether through verb inflection, adverbs or verbal periphrases – suggests a hierarchy of complexity that aligns with Dahl's (2007) view of textual complexity. According to Dahl (2007), the complexity of a linguistic unit is determined by the amount of information it encodes: the more information it contains (e.g., grammatical gender marking, tense, aspect), the more complex the unit becomes. In this regard, verb inflection often condenses more linguistic information into a single unit, whereas

verbal periphrases “distribute” this information across multiple units, making inflected forms more complex than periphrases.

Regarding cohesion, its absence in a text can lead to significant reading difficulties. A lack of cohesion often arises when the connections between sentences within a paragraph – or between paragraphs – are unclear. These links are frequently conveyed not through explicit conjunctions, but implicitly through nouns, verbs, adverbs, and other linguistic features. Another key aspect of cohesion is the use of referential devices, such as pronouns, which depend on accurate identification of their antecedents. This can be challenging when multiple potential referents exist or when the referent is distant in the text. Our proposal focuses on referential cohesion, specifically examining anaphoric relations and the use of temporal markers within texts (Fang, Pace, 2013; Glenberg *et al.*, 1987).

Considering all these factors, our framework primarily evaluates complexity qualitatively, using descriptors that concretely refer to specific manifestations of complexity. However, a quantitative dimension is also considered, particularly in relation to text length.

As outlined in the methodology section, the analyzed aspects of complexity were further detailed into 81 concrete descriptors or features, grouped into 12 categories (see Appendix I): text size; purpose; communication context; conceptual diversity; concreteness; vocabulary/lexicon subset; morphology; morphosyntax; syntax; textual cohesion: pronouns, anaphora and ellipsis; textual cohesion: temporal references; and style. These descriptors were derived through text analysis conducted by the project team and subsequently validated by experts in AL and VET, to ensure their appropriateness for the target audience.

In addition to informing the final set of descriptors and validating the detailed description of each level, expert participation allowed us to better understand the contribution of feature frequency to text complexity. Contrary to the assumption that higher frequencies of the complexity-inducing phenomenon would consistently increase text complexity, discussion with the experts revealed that this was not universally true. Certain features were found to contribute to complexity regardless of their frequency of occurrence in a text.

However, there are exceptions perceived as operating on a *continuum* of frequency. These are:

- degree of concreteness of words (descriptors 20, 21)
- foreign words (descriptors 35, 36, 37)
- relative clauses (descriptors 53, 54, 55, 56)
- passive voice constructions (descriptors 57, 58, 59)

These descriptors can be classified within the categories of lexical-conceptual features (the first two) and syntactic structures (the latter two). The results for the first two descriptors are consistent with the notion that very small thresholds of unfamiliar words in a text can hinder comprehension (Hu, Nation, 2000; Schmitt *et al.*, 2011). In other words, when considering lexical complexity, there exists a limited margin for understanding abstract and unfamiliar words within a frequency continuum. Conversely, the latter two descriptors, of a syntactic nature, align with traditional readability frameworks (Dale, Chall, 1948; Flesch, 1948; Gunning, 1952; Kincaid *et al.*, 1975; Lorge, 1948; McLaughlin, 1969; Spache, 1953), where complexity is gauged by the frequency of their occurrence within a text.

The complete descriptions of the three complexity levels for the three languages, along with the full set of descriptors associated with each level, are provided in Appendix II (French), Appendix III (Portuguese) and Appendix IV (Spanish).

Conclusion

This paper presents a comprehensive framework for analyzing text complexity tailored to low-literacy adult native speakers, developed within the scope of the iRead4Skills project. The findings and methodology detailed herein contribute significantly to advancing the understanding of textual complexity, offering a multifaceted approach that integrates both quantitative and qualitative dimensions to address the challenges faced by this specific population. The framework's focus on real-world applications, including the creation of tools for assessing and adapting text complexity, underscores its relevance in promoting literacy and reading practices among low-literacy adults.

This framework's main contribution lies in the integration of linguistic, syntactic, and cohesion features into a detailed classification

system covering three complexity levels: Very Easy, Easy, and Plain. By incorporating insights from existing frameworks, linguistic theory, authentic text analysis, and experts in AL and VET, the proposed framework addresses gaps in current research, particularly concerning the needs of adult learners with limited formal education. Unlike many traditional readability metrics, this framework accounts for the nuanced challenges experienced by low-literacy adults, such as lexical-conceptual issues, syntactic structures, and cohesive markers. These dimensions align with the real-world contexts in which literacy is practiced, making the framework not only theoretically robust but also practically applicable.

Another key contribution is its description of language-specific characteristics of French, Portuguese, and Spanish. The research presented here highlights the variability in complexity features and their manifestations, offering valuable insights for cross-linguistic studies of text complexity. The validation process, involving teachers and trainers, highlighted the importance of aligning theoretical descriptors with practical relevance, fostering a framework that is both academically rigorous and user-friendly for practitioners.

The framework proposed underscores the importance of viewing text complexity as a continuum, reflecting the diverse literacy needs and capacities of low-literacy adults. By framing complexity as a scalable concept, the research acknowledges the heterogeneity within this population, providing a flexible tool for educators and researchers alike.

Moreover, the creation of leveled texts with carefully controlled complexity aims to address a key challenge identified by the project: the lack of suitable reading materials for adults with low literacy. At the same time, these texts are designed to engage the target population effectively. This approach aligns with the concept of a reader's optimal difficulty level – an idea related to Vygotsky's Zone of Proximal Development – which suggests that there is a specific level of text difficulty that keeps readers both interested and appropriately challenged. While this concept is often applied to second language learners, research demonstrates its effectiveness in promoting learning without causing frustration (Chiang, 2015).

In practical terms, the framework contributes to the development of innovative tools, including an intelligent system for evaluating text

complexity and a writing assistant for adapting texts to appropriate levels of difficulty. These tools hold the potential to transform literacy education for adults, addressing a critical gap in the availability of suitable reading materials. By fostering greater engagement with texts, the framework and its associated tools aim to enhance reading practices, supporting the personal and professional development of low-literacy adults.

This research also opens new avenues for further exploration. Future studies could expand the framework to additional languages, investigate its applicability to other target populations, or explore its integration with emerging. Additionally, longitudinal studies tracking the impact of the framework and its tools on literacy outcomes could provide valuable evidence of their effectiveness.

Furthermore, since the elaboration of this work, the results from PIAAC's 2023 Survey of Adults Skills were released. According to this latest survey, Portugal, France, and Spain all ranked below the OECD average in adult literacy proficiency. Portugal, in particular, occupied the penultimate position among all participating countries, and all three countries experienced a decline in adult literacy levels over the last decade (OECD, 2024a). These findings underscore the continued relevance and urgency of the present research in addressing low literacy among adult populations in these countries.

Additionally, it is important to highlight that the descriptions of proficiency levels in PIAAC's framework, which served as one of the foundations for the development of our complexity levels and descriptors, were slightly revised (OECD, 2013b). The new descriptions are longer and include more content-specific elements. For example, at the first proficiency level (below Level 1), there are now explicit references to the consistency of word usage and how this consistency contributes to the overall meaningfulness of the sentence. At the sentence level, the adjacency of concepts and the way they relate to one another are also considered (OECD, 2024b). These revisions, which emphasize the internal coherence and semantic connectivity of texts – in contrast to the previous ones, which focused more on the tasks readers were expected to perform at each level and how information was presented in the text as a whole –, align with the approach and findings presented in this paper. However, the levels and descriptors proposed in this work remain

significantly more detailed, offering a richer and more fine-grained characterization of text complexity. This ensures that our framework continues to provide a valuable and highly relevant tool for addressing the specific needs of low-literacy adult readers.

In conclusion, the iRead4Skills framework represents a relevant advancement in the study of text complexity, bridging the gap between theoretical research and practical application. By addressing the unique challenges faced by low-literacy adult native speakers, this framework contributes to a deeper understanding of literacy as a continuum and provides a valuable resource for educators, researchers, and policymakers striving to enhance literacy outcomes worldwide.

Acknowledgements

We would like to thank Roser Gauchola for her valuable contribution to the development of complexity descriptors and features for Spanish within the iRead4Skills project.

Part of the research presented here is supported by the Portuguese national funding through the FCT – Portuguese Foundation for Science and Technology, I.P. as part of the project UID/03213/2025 – Linguistics Research Centre of NOVA University Lisbon (CLUNL) (<https://doi.org/10.54499/UID/03213/2025>).

References

- ALTE, The ALTE CAN DO PROJECT (1992–2002), 2002. Consulted on January 6th 2023. Available at: www.alte.org/can_do/alte_cando.pdf.
- AMBRIDGE, B., BIDGOOD, A., PINE, J. M., ROWLAND, C. F., FREUDENTHAL, D., Is Passive Syntax Semantically Constrained? Evidence From Adult Grammaticality Judgment and Comprehension Studies, *COGNITIVE SCIENCE*, 2016, **40(6)**, 1435-1459. <https://doi.org/10.1111/cogs.12277>
- ANDERSON, J.R., *Cognitive psychology and its implications*, New York, W.H. Freeman, 1995.
- ANQEP, News, 2023. Consulted on January 6th 2023. <https://www.anqep.gov.pt/np4/home>
- BAILIN, A., GRAFSTEIN, A., *Readability: Text and Context*, London, Palgrave Macmillan, 2016. <https://doi.org/10.1057/9781137388773>
- BARTON, D., HAMILTON, M., IVANIC, R., *Situated Literacies: Reading and Writing in Context*, Abingdon, Routledge, 2000. <https://doi.org/10.4324/9780203984963>
- BECK, I.L., MCKEOWN, M.G., OMANSON R., POPLE, M., Improving the comprehensibility of stories: The effects of revisions that improve coherence, *READING RESEARCH QUARTERLY*, 1984, **19**, 263-277, <https://doi.org/10.2307/747821>

- CARREIRAS, M., CARRIEDO, N., ALONSO, M., FERNÁNDEZ, A., The role of verb tense and verb aspect in the foregrounding of information during reading, *MEMORY & COGNITION*, 1997, **25**(4), 438-446. <https://doi.org/10.3758/bf03201120>
- CHIANG, M. H., Effects of varying text difficulty levels on second language (L2) reading attitudes and reading comprehension, *JOURNAL OF RESEARCH IN READING*, 2015, **39**(4), 448-468. <https://doi.org/10.1111/1467-9817.12049>
- COUNCIL OF EUROPE, *Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume*, Strasbourg, Council of Europe Publishing, 2020. Available at: www.coe.int/lang-cefr.
- CUNNINGHAM, J. W., MESMER, H. A., Quantitative measurement of text difficulty: What's the use? *ELEMENTARY SCHOOL JOURNAL*, 2014, **115**, 255-269. <https://doi.org/10.1086/678292>
- DAHL, Ö., Definitions of complexity, *Proceedings of the Colloquium on complexity, accuracy and fluency in second language use, learning & teaching*, 2007, 41-46.
- DALE, E., CHALL, J., A Formula for Predicting Readability, *EDUCATIONAL RESEARCH BULLETIN*, 1948, **27**, **28**, 11-20, 37-54.
- DORONILLA, M., *Landscapes of Literacy: an Ethnographic Study of Functional Literacy in Marginal Philippine Communities*, Hamburg, UIE, 1996.
- FANG, Z., PACE, B. G., Teaching with challenging texts in the disciplines, *JOURNAL OF ADOLESCENT & ADULT LITERACY*, 2013, **57**(2), 104-108. <https://doi.org/10.1002/JAAL.229>
- FLOR, M., KLEBANOV, B., SHEEHAN, K., Lexical tightness and text complexity, *Proceedings of the Workshop on Natural Language Processing for Improving Textual Accessibility*, Atlanta, GA, 2013, 29-38.
- FRANSMAN, J., *Understanding Literacy: A Concept Paper*, Background paper for EFA global monitoring report, UNESCO, 2005.
- FRAZIER, L., Syntactic complexity, in DOWTY, D. R., KARTTUNEN, L., ZWICKY, A. M. (eds.), *Natural Language Parsing: Psychological, Computational, and Theoretical Perspectives*, Cambridge University Press, 1985, 129-189. <https://doi.org/10.1017/CBO9780511597855.005>
- GILLIE, P. J., A simplified formula for measuring abstraction in writing, *JOURNAL OF APPLIED PSYCHOLOGY*, 1957, **41**(4), 214-217. <https://doi.org/10.1037/h0045036>
- GLENBERG, A.M., MEYER, M., LINDEM, K., Mental models contribute to foregrounding during text comprehension, *JOURNAL OF MEMORY AND LANGUAGE*, 1987, **26**. [https://doi.org/10.1016/0749-596X\(87\)90063-5](https://doi.org/10.1016/0749-596X(87)90063-5)
- GOODMAN, K., *On Reading*, University of Michigan, Heinemann, 1996.
- GOLDMAN, S., LEE, C., Text complexity. State of the art and the conundrums it raises, *THE ELEMENTARY SCHOOL JOURNAL*, 2014, **115**, 290-300. <https://doi.org/10.1086/678298>

- GUNNING, R., *The Technique of Clear Writing*, New York, McGraw-Hill, 1952.
- HANEMANN, U., Shaping the literacy agenda from a lifelong learning perspective, *Better Skills, Better Jobs, Better Lives*, OECD, 2012. <https://doi.org/10.1787/9789264177338-en>
- HARLEY, T. A., *Psychology of Language: From Data to Theory*, London, Taylor and Francis, 2013. <https://doi.org/10.4324/9781315859019>
- HEILMAN, M., COLLINS-THOMPSON, K., CALLAN, J., ESKENAZI, M., Combining lexical and grammatical features to improve readability measures for first and second language texts, *Proceedings of the NAACL Human Language Technology Conference*, 2007, 460-467.
- HIEBERT, E.H., Using multiple sources of information in establishing text complexity, *Reading Research Report*. 11, 2011. <https://doi.org/10.1177/002205741119100206>
- HU, M., NATION, I., Vocabulary density and reading comprehension, *READING IN A FOREIGN LANGUAGE*, 2000, **13(1)**, 403-430. <https://doi.org/10.64152/10125/66973>
- KINCAID, J. P., FISHBURNE, R. P. J., ROGERS, R. L., CHISSOM, B. S., *Derivation of new readability formulas Automated Readability Index, Fog Count and Flesch Reading Ease Formula for Navy enlisted personnel*, University of Central Florida, Naval Technical Training, U.S., 1975. <https://doi.org/10.21236/ADA006655>
- LAUFER, B., Lexical coverages, inferencing unknown words and reading comprehension: How are they related? *TESOL QUARTERLY*, 2020, **54(4)**, 1076-1085. <https://doi.org/10.1002/tesq.3004>
- LIVELY B., PRESSEY, S., A method for measuring the vocabulary burden of textbooks, *EDUCATIONAL ADMINISTRATION AND SUPERVISION*, 1923, **9(7)**, 389-398.
- LORGE, I., The Lorge and Flesch Readability Formulae: A Correction, *SCHOOL AND SOCIETY*, 1948, **67**, 141-142.
- MACGINTIE, W., TRETIAK, R., Sentence depth measures as predictors of reading difficulty, *READING RESEARCH QUARTERLY*, 1971, **6 (3)**, 364-377. <https://doi.org/10.2307/747125>
- MCLAUGHLIN, G. H., SMOG Grading - a New Readability Formula, *JOURNAL OF READING*, 1969, **128**, 639-646.
- MESMER, H. A., CUNNINGHAM, J. W., HIEBERT, E. H., Toward a theoretical model of text complexity for the early grades: learning from the past, anticipating the future, *READING RESEARCH QUARTERLY*, 2012, **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., iRead4Skills - Complexity Levels (V1.0). Zenodo, 2023. <https://doi.org/10.5281/zenodo.10459090>
- OECD, Technical report of the survey of adult skills (PIAAC), OECD Publishing, 2013a.

- OECD., The Survey of Adult Skills Reader's Companion. OECD Publishing, 2013b.
- OECD, The Assessment Frameworks for Cycle 2 of the Programme for the International Assessment of Adult Competencies, OECD Publishing, 2021. <https://doi.org/10.1787/4bc2342d-en>.
- ROTHES, L., QUEIRÓS, J., Práticas de leitura e competências de literacia: resultados e ilações do PIAAC. Plano Nacional de Leitura, *ENTRELER*, 2027, **1**, 32-42.
- SCHMITT, N., JIANG, X., GRABE, W., The Percentage of Words Known in a Text and Reading Comprehension, *THE MODERN LANGUAGE JOURNAL*, 2011, **95(1)**, 26-43. <https://doi.org/10.1111/j.1540-4781.2011.01146>. <https://doi.org/10.1111/j.1540-4781.2011.01146>
- SCRIBNER, S., COLE, M., *The Psychology of Literacy*, Harvard, Harvard University Press, 1981. <https://doi.org/10.4159/harvard.9780674433014>
- SPACHE, G., A new readability formula for primary-grade reading materials, *THE ELEMENTARY SCHOOL JOURNAL*, 1953, **537**, 410-413. <https://doi.org/10.1086/458513>
- STREET, B., Introduction: the new literacy studies, in STREET, B. (org.), *Cross-cultural Approaches to Literacy*, Cambridge University Press, 1993.
- STREET, B., Understanding and defining literacy. Background Paper for the Education for All Global Monitoring Report 2006: Literacy for Life, 2005.
- STREET, J., DĄBROWSKA, E., Lexically specific knowledge and individual differences in adult native speakers' processing of the English passive, *APPLIED PSYCHOLINGUISTICS*, 2014, **35(1)**, 97-118. <https://doi.org/10.1017/S0142716412000367>
- STROMQUIST, N. P., Women's rights to adult education as a means to citizenship, *INTERNATIONAL JOURNAL OF EDUCATIONAL DEVELOPMENT*, 2006, **26(2)**, 140-152. <https://doi.org/10.1016/j.ijedudev.2005.07.021>
- UNESCO, Literacy. [Uis.unesco.org](https://uis.unesco.org). Consulted on 31/10/2023, 2020. Available at: <https://uis.unesco.org/node/3079547>
- VALENTE, A. (coord.) Referencial de Competências-chave de Educação e Formação de Adultos – Nível Básico. ANQEP, 2021.

Appendix I

<i>Category</i>	<i>Descriptor (complexity level)</i>
<i>text size</i>	1. The text has up to 50 words. (Very Easy) 2. The text has up to 150 words. (Easy) 3. The text has up to 250 words. (Plain) 4. The text has up to 500 words. (+Complex)
<i>purpose</i>	5. The text serves the purpose of performing common/familiar tasks (e.g., instruction to buy bus tickets, usage instructions on product labels) (Very Easy) 6. The text serves the purpose of presenting a new topic (e.g., initiatory text on the importance of recycling). (Very Easy) 7. The text is interesting for the reader to inform him/herself in moments of leisure (e.g., sports chronicle, daily news). (Easy) 8. The text to learn/perform new tasks (e.g., instructions on how to use the company email). (Easy) 9. The text can be about varied topics of interest to the reader for information or leisure. (e.g., nature magazine article). (Plain)
<i>communication context</i>	10. The text is used in everyday communications (e.g., transport schedules, lists, menus, general instructions, item price information, personal communication, etc.). (Very Easy) 11. The text is used in work contexts or media contexts (e.g., work communication or instructions, news, ads, sports). (Easy) 12. The text is used in leisure contexts (e.g., literature and fiction, reportage, opinion articles). (Plain) 13. The text is used artistic contexts (e.g., literature, poetry). (Plain) 14. The text is used in specialized contexts (e.g., formal professional communication, online forums). (Plain) 15. The text is used in academic or scientific contexts (e.g., essays, reports, academic papers) (+Complex)

<i>conceptual diversity</i>	16. The text refers to concepts of everyday activities. (e.g., food, housekeeping, clothes). (Very Easy) 17. The text refers to concepts related to domain-specific uses (e.g., work, sports, hobbies). (Easy) 18. The text refers to concepts that require readers to puzzle together pieces of information. (e.g., stars, volcanos). (Plain) 19. The text refers to multifaceted or new concepts. (e.g., black matter, brain development, statistics). (+Complex)
<i>concreteness</i>	20. The text is mostly composed of words that refer to concrete objects (e.g., chair, apple, hammer, computer). (Very Easy) 21. The text has a minor percentage of words that refer to common abstract ideas (e.g., happiness, imagination, love). (Easy) 22. The text uses words that refer to abstract concepts and ideas (e.g., aging, reasoning, loyalty). (Plain) 23. The text uses words that refer to multifaceted and/or domain-specific abstract concepts and ideas (e.g., poetry, democracy, nuclear fusion). (+Complex)
<i>vocabulary/ lexicon subset</i>	24. The text does not use abbreviations, except the frequent forms of treatment. (Very Easy) 25. The text uses common abbreviations after explanation in the text. (Easy) 26. The text uses common abbreviations. (Plain) 27. The text does not use acronyms or initialisms. (Very Easy) 28. The text uses acronyms and initialisms that are commonly known, i.e., that can be understood in isolation. (Easy) 29. The text uses acronyms and initialisms after describing them. (Plain) 30. The text uses all types of acronyms (+Complex) 31. The text has words and simple expressions of everyday matters, using frequent and commonly known nouns. (Very Easy) 32. The text has frequent words and expressions that can refer to everyday matters but also to other specific domains or to work/professional matters. (Easy) 33. The text has words and expressions less known that refer to specific domains. (Plain) 34. The text has words and expressions that refer to specific domains, specialized jargon or archaic terms. (+Complex) 35. The text has no foreign words, except those very common and widely known. (Very Easy) 36. The text has common foreign words, but these are not the majority in the text. (Easy) 37. The text uses common foreign words frequently. (Plain) 38. The text has polysemic words (i.e., words that can have more than one

	meaning and can be ambiguous), but these are not the majority in the text. (Plain) 39. The text has highly polysemic words (i.e., words that can have more than one meaning and can be ambiguous). (+Complex)
<i>morphology</i>	40. The text has common compound words or transparent compound words (words or expressions whose meaning can be easily understood by combining the meaning of the parts). (Very Easy) 41. The text has common compound words whose meaning can be more or less understood by combining the meaning of the parts. (Easy) 42. The text has compound words whose meaning cannot easily be understood by combining the meaning of the parts). (Plain) 43. The text has diverse compound words whose meaning cannot easily be understood by combining the meaning of the parts. (+Complex) 44. The text has derived words, with common and familiar prefixes and suffixes, but these are not the majority in the text. (Very Easy) 45. The text has derived words with frequent prefixes and suffixes, but these are not the majority in the text. (Easy) 46. The text has derived words, with varied suffixes and prefixes. (Plain) 47. The text has derived words with less common suffixes and prefixes. (+Complex)
<i>morphosyntax</i>	48. Verb Moods and Tenses (language-dependent: cf. Appendixes II, III and IV): verbs in the Infinitive (Very Easy) verbs in the Indicative Present (Very Easy) verbs in the Indicative Past Imperfect (Very Easy) verbs in the Indicative Past Perfect (Very Easy) verbs in the Indicative Pluperfect (+Complex) verbs in the Indicative Simple Future (Easy) verbs in the Gerund (Easy) verbs in the Imperative (Very Easy) verbs in the Conditional (Easy) verbs in the Subjunctive Present (Plain) verbs in the Subjunctive Past Imperfect (Easy) verbs in the Subjunctive Future (Plain) verbs in the Compound Past Perfect Indicative (Plain) verbs in the Indicative Compound Pluperfect (Plain) verbs in the Indicative Compound Future (Plain) verbs in the Subjunctive Compound Past Perfect (Plain) verbs in the Subjunctive Compound Plusperfect (Plain) verbs in the Subjunctive Compound Future (Plain) common auxiliary verbs in the Indicative tenses (Easy) expressions with common modal and aspectual verbs in simple tenses (Easy) expressions with the auxiliary verbs in compound tenses (+Complex)

syntax	49. Most of the sentences in the text are short and direct or combined through frequent and familiar connectors. (Very Easy) 50. The text has sentences that are combined through coordination and subordination connectors. (Easy) 51. The sentences in the text are frequently combined through coordination and subordination connectors, and sometimes with less common connectors. (Plain) 52. The sentences in the text are long and composed of several other sentences combined through coordination and subordination connectors. (+Complex) 53. The text has relative sentences related to the subject of the sentence, but these are not the majority of the text. (Very Easy) 54. The text has relative sentences related to the subject of the sentence. (Easy) 55. The sentences in the text frequently include relative sentences related to the subject of the sentence. (Plain) 56. The sentences in the text frequently include relative sentences related to the subject or the object of the sentence. (+Complex) 57. The majority of the sentences in the text are in the active voice. (Very Easy) 58. The text has sentences in the passive voice. (Easy) 59. The text has sentences in the passive voice, including common impersonal and specific passive constructions. (Plain) 60. The text has simple sentences and in the basic linear order (subject-verb-complements). (Very Easy) 61. The text has simple sentences not in the basic linear order (subject-verb-complements), but these are not the majority of the text. (Easy) 62. The sentences of the text are not necessarily in the basic linear order (subject-verb-complements). (Plain) 63. The sentences of the text have diverse linear orders (subject-verb-complements; verb-complements-subject, etc.). (+Complex)
textual cohesion: Pronouns, anaphora	64. The text has few or none omitted elements. (Very Easy) 65. There are repeated elements instead of pronouns. (Very Easy) 66. The text has some omitted elements. (Easy) 67. The text uses pronouns. (Easy) 68. The text has omitted elements. (Plain) 69. The text uses pronouns and related words to indicate what is being referred. (Plain)
textual cohesion: temporal references	70. The text has simple and linear temporal sequence or location, expressed by common temporal adverbs or connectors. (Very Easy) 71. The text has a linear temporal sequence or location, expressed by temporal adverbs or connectors and by verb tenses. (Easy) 72. The text can have a non-linear temporal sequence or location. (Plain) 73. The text has diverse and non-linear temporal sequences or locations. (+Complex)

<i>style</i>	74. The text does not use figurative senses. (Very Easy) 75. The text uses common figurative senses. (Easy) 76. The text uses less common figurative senses. (Plain) 77. The text uses uncommon figurative senses. (+Complex) 78. The text uses only comparison, personification, and hyperbole. (Very Easy) 79. The text uses metonymy, familiar metaphors. (Easy) 80. The text uses non common metaphors. (Plain) 81. The text uses diverse figures of speech. (+Complex)
--------------	--

Appendix II - iRead4Skills Complexity Levels for French

1) Description des niveaux

NIVEAU 1 – TRÈS FACILE

Textes courts et simples jusqu'à environ 50 mots, visant à accomplir des tâches habituelles (ex : horaires de transport, étiquettes de produits, communications personnelles).

Concepts du quotidien avec de nombreuses références à des **objets concrets** (ex : *chaise, pomme, ordinateur*).

Mots simples et fréquents, avec une absence de mots étrangers, d'abréviations et d'acronymes à l'exception des formes très fréquentes (ex : *Internet, Mme*).

Présence d'expressions et mots composés fréquents (ex : *grand-mère, pomme de terre*) mais affixation rare, à l'exception des affixes très fréquents (*re-, -ment, -tion*).

Phrases courtes et simples, parfois composées à l'aide de conjonctions fréquentes et familières (ex : *et, mais*).

Les phrases et propositions suivent **l'ordre de base sujet-verbe-complément**, même si un complément de temps ou de lieu peut parfois se trouver avant le sujet (ex : *Hier, j'ai pris le train*).

Textes principalement écrits au présent de l'indicatif et au passé composé, avec utilisation des périphrases verbales pour décrire le futur proche (*aller* + infinitif), le passé récent (*venir de* + infinitif) et le présent progressif (*être en train de* + infinitif). Présence de l'impératif et de l'infinitif présent. Le gérondif est également admis. Pas ou très peu de phrases à la voix passive.

Chaîne référentielle complète et produite au moyen d'éléments répétés ou d'anaphores pronominales, mais uniquement avec des pronoms personnels sujet (ex : *Je comprends bien Maria parce qu'elle parle fort*).

Pas d'ellipse (ex : *Jean préfère la voile et sa sœur préfère le kayak*. vs *Jean préfère la voile ; sa sœur, le kayak*).

Cadre spatio-temporel simple et linéaire. La cohérence temporelle est donnée au moyen d'adverbes ou de connecteurs temporels courants (ex : *aujourd'hui, demain, avant*) et non par les temps verbaux (ex : *Elle est partie avant l'arrivée de Jean*. vs *Elle était partie et Jean est arrivé*).

Absence de figures de style.

NIVEAU 2 – FACILE

Textes courts d'environ 150 mots, permettant au lecteur d'accomplir des tâches habituelles, de **s'informer ou se distraire pendant son temps libre** (ex : presse quotidienne, information de proximité, courtes fictions).

Concepts du quotidien ou liés à un domaine spécifique proche du lecteur (sport, travail, loisirs) et présence d'idées abstraites courantes (ex : sentiments, états d'esprit, qualités et défauts).

Mots et expressions fréquents qui peuvent se référer au contexte professionnel (ex : *réunion, lundi, sortie*) avec quelques abréviations et mots étrangers courants (ex : *km, rdv, email*).

Présence d'acronymes et de sigles passés dans le langage courant (*SNCF, EDF, CDD, PDG*).

Utilisation d'expressions et mots composés dont le sens peut être compris en combinant le sens des parties (ex : *sac à dos, non-fumeur*). Présence de mots dérivés contenant des affixes courants (ex : *dé-, in-, -age, -ité, -eur*).

Phrases simples et composées, pouvant s'éloigner de l'ordre de base sujet-verbe-complément avec la présence de tous types de compléments circonstanciels.

Présence de propositions subordonnées relatives, temporelles et causatives introduites par des connecteurs fréquents (ex : *qui, que, quand, comme, parce que, si*).

Utilisation de la plupart des conjonctions de coordination (*et, mais + donc, car, où*).

Variété de temps verbaux et de modes : imparfait, futur simple, passé composé et plus-que-parfait de l'indicatif, présent du subjonctif, conditionnel de politesse (ex : *je voudrais, vous pourriez*), infinitif passé, gérondif. Le passé simple est admis lorsque conjugué à la 3ème personne de singulier. Phrases à la voix passive et structures passives avec *se faire* à l'indicatif.

Chaîne référentielle suivie au moyen d'**anaphores pronominales** (tous types de pronoms). Très peu d'ellipses.

Cadre spatio-temporel plus complexe mais toujours en séquences linéaires. La cohérence temporelle peut être assurée par les temps verbaux (ex : *Piero était parti quand la police est arrivée*).

Utilisation de mots au sens figuré si le sens est courant (ex : *la nuit tombe en hiver*) et de **quelques figures de style** : comparaisons, métonymies et métaphores courantes (ex : *grand comme une montagne, prendre un verre, suivre un chemin différent*).

NIVEAU 3 – ACCESSIBLE

Textes de tailles variées, jusqu'à environ 250 mots, visant à présenter **un nouveau sujet, une nouvelle tâche** ou à s'informer sur des sujets variés dans **les contextes professionnels, médiatiques, artistiques et de loisirs** (ex : lecture plaisir).

Présence de **concepts complexes traités de manière introductive et d'idées abstraites** (ex : *vieillesse, loyauté*).

Mots et expressions liés à des domaines spécifiques (travail, média, arts, loisirs) mais non spécialisés ou académiques.

Utilisation d'abréviations, sigles et acronymes après explication dans le texte (ex : *Parti Socialiste (PS)*).

Nombreux mots dérivés avec une **grande variété d'affixes**, sauf quelques affixes très peu courants (ex : *méga-, géo-, -(i)fier*).

Expressions et mots composés dont le sens ne peut pas être compris en combinant le sens des parties, mais qui restent concrets ou courants (ex : *troisième âge, rez-de-chaussée* vs *ça et là, année-lumière, à tout bout de champ*).

Phrases simples et composées, suivant une grande **variété d'ordres syntaxiques**. Présence de propositions subordonnées relatives, temporelles et causales introduites par

des **connecteurs variés** (ex : *dont, avant que, après que, à cause de, alors*). Utilisation de toutes les conjonctions de coordination.

Présence de presque tous les temps et modes, à l'exception des passé antérieur, subjonctif passé, subjonctif imparfait et impératif composé. Phrases à la voix passive et structures passives avec *se faire*.

Ellipses fréquentes et utilisation d'**anaphores pronominales et nominales** (ex : *Jean est boulanger. Il a repris l'échoppe de son père. Cet enfant du pays est la fierté du village.*).

Cadre spatio-temporel complexe et séquences temporelles non linéaires, le nombre de temps différents par phrase reste néanmoins limité (ex : *Les médecins ont analysé les réponses de patients ayant subi une crise cardiaque. Ils ont ensuite confirmé que la graisse est l'aliment le plus dangereux.* vs *Après avoir analysé les réponses de patients ayant subi une crise cardiaque, les médecins ont confirmé que la graisse est l'aliment le plus dangereux.*).

Utilisation de mots au sens figuré même si le sens est peu courant (ex : *un manteau de neige*) et présence de **nombreuses figures de styles** : comparaisons, métonymies, métaphores, personnifications (ex : *Les maladies se cachent partout*) et hyperboles (ex : *Cette valise pèse une tonne.*).

2) Tableau de niveaux

	Descripteurs	Très Facile	Facile	Accessible	complexe+
Taille	1 Le texte peut contenir jusqu'à 50 mots.				
	2 Le texte peut contenir jusqu'à 150 mots.				
	3 Le texte peut contenir jusqu'à 250 mots.				
	4 Le texte peut contenir jusqu'à 500 mots.				

	Descripteurs	Très Facile	Facile	Accessible	complexe+
But	5 Le texte est pour accomplir des tâches habituelles/quotidiennes (e.g. cuisine, horaires de bus, infos pratiques).				
	6 Le texte est pour accomplir des tâches quotidiennes (e.g. notices d'utilisation, tutoriels).				
	7 Le texte est pour s'informer pendant son temps libres (e.g. presse quotidienne, information de proximité, météo locale).				
	8 Le texte est pour présenter un nouveau sujet ou apprendre des nouvelles tâches.				
	9 Le texte est pour s'informer sur des sujets variés ou des lectures de plaisir.				

	Descripteurs		Très Facile	Facile	Accessible	complexe+
Domaine/ contexte	10	Communications quotidiennes (e.g. horaires de transport, menus).				
	11	Contextes de loisirs (fiction, reportages).				
	12	Contextes professionnels, médiatiques et artistique.				
	13	Contextes spécialisés et académiques.				

	Descripteurs		Très Facile	Facile	Accessible	complexe+
Concepts	14	Du quotidien (e.g. nourriture, ménage, vêtements).				
	15	Concepts liés à un domaine spécifique (e.g. travail, sports, loisirs).				
	16	Concepts qui exigent de rassembler différents éléments d'information (e.g. concepts complexes, texte introductif (volcans, étoiles)).				
	17	Concepts à multiples facettes ou nouveaux (e.g. la matière noire, les statistiques).				

	Descripteurs		Très Facile	Facile	Accessible	complexe+
Concrétude	18	Le texte a principalement des objets concrets (e.g. <i>pomme</i> , <i>ordinateur</i>).				
	19	Le texte a quelques idées abstraites courants (e.g. <i>bonheur</i> , <i>amour</i> , <i>imagination</i>).				
	20	Le texte contient idées et concepts abstraits (e.g. <i>vieillesse</i> , <i>loyauté</i>).				
	21	Le texte est pour présenter un nouveau sujet ou apprendre de nouvelles tâches.				
	22	Le texte a textes et concepts abstraites à multiples facettes e/ou spécifiques à un domaine (e.g. <i>poésie</i> , <i>démocratie</i> , <i>fusion nucléaire</i>).				

	Descripteurs		Très Facile	Facile	Accessible	complexe+
Lexique	23	Mots et expressions de la vie quotidienne, simples et fréquents.				
	24	Mots et expressions fréquents qui se réfèrent au contexte professionnel (<i>lundi</i> , <i>réunion</i> , <i>sortie</i>).				

	25	Mots et expressions qui se réfèrent à des domaines spécifiques.				
	26	Jargon spécialisé et termes archaïques (<i>génomé, moult</i>).				
	27	Pas de mots étrangers, sauf mots largement connus (<i>internet</i>).				
	28	Peu de mots étrangers, et ces mots sont courants (<i>email</i>).				
	29	Nombreux mots étrangers (<i>marketing, feedback, lowcost</i>).				
	30	Pas d'abréviations, sauf formes très fréquentes (<i>Mme., prof.</i>).				
	31	Abréviations courantes (<i>km., g., rdv., svp.</i>).				
	32	Abréviations après explications dans le texte (<i>millilitre (mL.)</i>).				
	33	Pas d'acronymes.				
	34	Acronymes et sigles passés dans le langage courant (<i>SNCF, EDF, CDD, PDG</i>).				
	35	Acronymes et sigles après explication (<i>Parti socialiste (PS)</i>).				
	36	Tous types de acronymes et sigles (<i>ADN, UNESCO, UNICEF</i>).				

		Descripteurs	Très Facile	Facile	Accessible	complexe+
Morphologie nominale	37	Pas ou peu de mots dérivés et affixes très fréquents (<i>re-, -ment, -tion</i>).				
	38	Présence de mots dérivés et affixes courants (<i>dé-, in-, -age, -ité, -eur</i>).				
	39	Affixes variés (<i>auto-, anti-, co-, ex-, inter-, mini-, pré-, sou(s)-, sur-, -able, -aire, -esse, -ien, -iste, rie</i>).				
	40	Affixes peu courants (<i>méga-, géo-, -(i)fier</i>).				
	41	Expressions et mots composés fréquents (<i>grand-mère, pomme de terre</i>).				
	42	Expressions et mots composés dont le sens peut être compris en combinant le sens des parties (<i>sac à dos, non-fumeur</i>).				
	43	Expressions et mots composés dont le sens ne peut pas être facilement compris en combinant le sens des parties (<i>troisième âge, rez-de-chaussée</i>).				
	44	Expressions peu fréquentes (<i>ça et là, année lumière, à tout bout de champ</i>).				

		Descripteurs	Très Facile	Facile	Accessible	complexe+
Ordre syntaxique	45	Phrases simples, dans l'ordre de base sujet-objet-complément.				
	46	Quelques phrases simples ne suivant pas l'ordre base (i.e. avec compléments de temps et de lieu).				
	47	Phrases simples e composées et variété d'ordre syntaxique.				

		Descripteurs	Très Facile	Facile	Accessible	complexe+
Phrases/ Propositions	68	Phrases/Propositions courtes et directes ou combinées à l'aide de connecteurs fréquents et familiers (<i>et, mais</i>).				
	69	Peu de propositions subordonnées relatives, temporelles et causales. Connecteurs fréquents (<i>qui, que, quand, comme, parce que</i>).				
	70	Prépositions combinées à l'aide de connecteurs de coordination et subordination (<i>donc, si</i>). Connecteurs variés (<i>dont, avant que, après que, à cause de, alors</i>).				
	71	Prépositions combinées à l'aide de connecteurs parfois peu fréquents (e.g. <i>malgré</i>).				
	72	Majorité de phrases longues et complexes + présence de connecteurs peu fréquents (e.g. <i>étant donné que</i>).				

		Descripteurs	Très Facile	Facile	Accessible	complexe+
Voix passive	73	Pas ou très peu de phrases à la voix passive.				
	74	Phrases à la voix passive et structures passives avec ' <i>se faire</i> '.				
	75	Phrases à la voix passive à l'indicatif.				

<i>Cohésion textuelle</i>	Descripteurs		Très Facile	Facile	Accessible	complexe+
	76	Présences d'éléments répétés à place des pronoms.				
	77	Très peu de ellipses (<i>Marc aime la voile, sa sœur le kayak.</i>).				
	78	Anaphores pronominales.				
	79	Ellipses fréquentes.				
	80	Anaphores pronominales et nominales (<i>Jean est boulanger. Il a repris l'échoppe de son père. Cet enfant du pays est la fierté du village</i>)				

<i>Cohésion spatio-temporelle</i>	Descripteurs		Très Facile	Facile	Accessible	complexe+
	81	Cadre spatio-temporel linéaire avec adverbess et connecteurs courants (<i>aujourd'hui, demain, avant</i>).				
	82	Séquences spatio-temporelles non-linéaires avec adverbess, connecteurs et temps verbaux (<i>Piero était parti quand la Police est arrivée</i>).				
	83	Séquences spatio-temporelles non-linéaires avec divers temps verbaux utilisés.				

<i>Style</i>	Descripteurs		Très Facile	Facile	Accessible	complexe+
	84	Pas de mots au sens figuré.				
	85	Utilisation de mots courants au sens figuré (<i>la nuit tombe en hiver</i>).				
	86	Utilisation de mots peu courants au sens figuré (<i>un manteau de neige</i>). Personnifications, hyperboles et métaphores.				
	87	Tout type de figures de styles + métaphores non courantes.				

Appendix III - iRead4Skills Complexity Levels for Portuguese

1) Descrição dos níveis

NÍVEL 1 – MUITO FÁCIL

Textos curtos e simples até cerca de 50 palavras, com o objetivo de desempenhar tarefas do quotidiano ou de apresentar 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 de um produto), 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 (*maçã, cadeira, fogão*).

Ausência de abreviaturas, siglas e estrangeirismos, à exceção daqueles muito frequentes (*internet, hamburger*).

Palavras compostas comuns (*couve-flor, guarda-chuva*).

Palavras com afixos frequentes (*infeliz, calmamente, trabalhador*).

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.: *O Rui e/ou a sua irmã foram à festa, mas o João não foi*).

Estrutura de subordinação frequentes estão presentes (ex.: oração subordinada adverbial temporal – *Desde que voltou, O Rui não dormiu*).

Uso predominante da voz ativa.

Textos utilizam tempos verbais no modo Infinitivo, Indicativo e Imperativo, nomeadamente o Presente, e Pretéritos Perfeito e Imperfeito. Ocorrência de construções perifrásticas comuns como *estar/começar/andar + a + infinitivo* ou *deixar/acabar + de + infinitivo*.

Não figuram tempos compostos, à exceção do Futuro que é transmitido através de *ir + infinitivo*.

Cadeia referencial completa sem omissão de elementos, com repetição em vez da utilização de pronomes (ex.: *O Rui gosta da Maria, porque a Maria é simpática*).

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 2 – 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 abreviações 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 (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.: *À escola, chegaram os autocarros da visita de estudo*). Possível uso da voz passiva.

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 3 – 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 abreviações comuns, sem explicação prévia (ex.: *m, km, g, l, Exmo.*).

Uso de estrangeirismos comuns frequentemente (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 da voz passiva (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 completa 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*).

2) Quadro dos níveis

	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+
<i>Propósito</i>	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
<i>Contexto comunicativo</i>	10	O texto circula em comunicações do dia a dia (e.g., horários de transportes, menus, 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 algumas 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>Exmo., m, km, g., l</i>).				
	27	O texto não utiliza siglas.				
	28	O texto apenas utiliza siglas e acrónimos conhecidos (e.g., <i>CP, IV/A, RTP, SIC, TVT</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>miocárdio, artéria coronária, dissecação espontânea, enfarte do 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, bobby, 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, 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 seu sentido pode ser percebido pela combinação das duas partes (e.g., <i>óculos de sol, mal-educado, meia-idade, 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, porta-voz, 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, mais-valia, pé-de-meia, cavalo de batalha</i>).				
	44	O texto tem palavras derivadas com prefixos e sufixos comuns e familiares em grande número (e.g., <i>infeliz, calmamente, organizar, trabalhador, casinha</i>).				
	45	O texto tem palavras algumas derivadas mais comuns, com prefixos e sufixos frequentes (e.g. <i>desorganizar, embalagem, organização, jardineiro, pré-cozinhado, finíssimo</i>).				
	46	O texto utiliza palavras derivadas com sufixos e prefixos variados (e.g., <i>nervosismo, lavável, curiosidade, altitude, superpoder, preconceituoso</i>).				
	47	O texto utiliza palavras derivadas com sufixos e prefixos menos comuns (e.g., <i>megatonelada, tendinite, geopolítico, neoliberal</i>).				

		Descritores	Muito fácil	Fácil	Claro	+ complexo
Morfossintaxe	48	<u>Modos e Tempos verbais</u>				
		Verbos no infinitivo (e.g., <i>gostar, comer, fugir</i>).				
		Verbos no Presente do Indicativo (e.g., <i>gosto, comem, foge</i>).				
		Verbos no Pretérito Imperfeito do Indicativo (e.g., <i>gostavam, comia, fugias</i>).				
		Verbos no Pretérito Perfeito do Indicativo (e.g., <i>gostou, comi, fugiram</i>).				
		Verbos no Pretérito Mais-que-perfeito do Indicativo (e.g., <i>gostara, comeras, fugira</i>)				
		Verbos no Future 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!, fale!, 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, fuja</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>ontem. O pai trabalha na fábrica, mas a mãe fica em A criança fugiu casa</i>).				
	50	O texto combina frases através de conjunções coordenadas e subordinadas (e.g., <i>poucos clientes ouviam os músicos tocando. Quando 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 sem 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 que 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 por isso que tantas pessoas o consideram uma opção.</i>).				
	53	O texto tem poucas orações relativas referentes ao sujeito da frase (e.g., <i>O senhor que 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 que 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, a quem os amigos tratam por «Kika», começou a surfar muito pequena.</i>).				
	56	As frases no texto incluem frequentemente orações relativas correspondentes ao sujeito ou ao objeto da frase (e.g., <i>O rapaz que abraçou a mãe dirigiu-se à entrada; O rapaz que a mãe abraçou dirigiu-se à entrada.</i>).				
	57	No texto, há mais frases na voz ativa (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 na voz passiva (e.g., <i>A visita foi paga pelo patrão.</i>).				
	59	O texto apresenta frases na voz passiva (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 não numa ordem 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 elípses	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 <u>Maria</u> porque a <u>Maria</u> não é simpática.</i>).				
	66	O texto tem alguns elementos omitidos (e.g., <i>Gostava de fazer artesanato. vs. Eu gostava de fazer artesanato.</i>).				
	67	O texto usa pronomes (e.g., <i>O João não gosta da <u>Maria</u> porque <u>ela</u> não é simpática.</i>).				
	68	O texto apresenta elementos omitidos (e.g., <i>□ Mostrámos em estudos anteriores que a gordura solidifica quando □ arrefece. vs. Nós mostrámos em estudos anteriores que a gordura solidifica quando <u>ela</u> 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 complete de uma tempestade de neve e como <u>esta</u> se forma. Também indicamos onde o fenómeno pode ser visto.</i>).				

		Descritores	Muito fácil	Fácil	Claro	+ complexo
Estilo	70	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>).				
	71	O texto apresenta sentidos figurativos comuns (e.g., <i>Este verão os preços <u>dispararam</u>.</i>).				
	72	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>).				
	73	O texto apresenta sentidos figurativos mais incomuns (e.g., <i>Os primeiros raios de sol <u>aconchegaram-me</u> as faces.</i>).				
	74	O texto utiliza apenas comparações, personificações e hipérboles mais comuns (e.g., <i>As minhas costas <u>estão a dar cabo de mim</u>; Ela é linda como uma flor; A tua mala <u>pesa uma tonelada</u>.</i>).				
	75	O texto faz uso da metonímia (e.g., <i>beber um copo</i>), assim como metáforas familiares (e.g., <i>O meu pai tem olhos de águia.</i>).				
	76	O texto faz uso de metáforas incomuns (e.g., <i>A sociedade está doente; ele é um lobo em pele de cordeiro.</i>).				

	77	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>).				
--	----	--	--	--	--	--

Appendix IV - iRead4Skills Complexity Levels for Spanish

1) Descripción de los niveles

NIVEL 1 – MUY FÁCIL

Textos breves y sencillos, de hasta 50 palabras, con el objetivo de realizar tareas cotidianas.

Conceptos simples de la vida cotidiana, presentando en primer lugar el tema de la comunicación.

Contextos comunicativos básicos: cotidianos (horarios de transporte/ listas/ menús/ información sobre precios de productos)

Palabras sencillas y frecuentes, de uso común en contextos cotidianos (transporte, comida, familia, trabajo). Uso predominante de **palabras concretas que se refieren a objetos concretos** del mundo real (*manzana, silla, cocina*). **Ausencia de abreviaturas, acrónimos y extranjerismos**, salvo algunos muy comunes (*internet, hamburguesa*).

Palabras compuestas comunes (ej. *baloncesto, bienestar*).

Palabras con afijos frecuentes (*infeliz, tranquilamente, trabajador*).

Frases cortas y sencillas, utilizando **conjunciones simples y en orden directo** (Sujeto-Verbo-Complemento). Presencia de estructuras de **coordinación frecuentes** (copulativa, disyuntiva y adversativa) (por ejemplo, *los músicos tocaban, pero pocos clientes escuchaban*). Estructuras de **subordinación frecuentes** (por ejemplo, oraciones subordinadas adverbiales temporales - *Desde que volvió, Paco no ha dormido*). Uso predominante de la **voz activa**.

En los textos se utilizan los **tiempos verbales en infinitivo, indicativo**, sobre todo el presente. Aparición de construcciones con verbos modales.

Cadena referencial completa sin omitir elementos.

Localización temporal lineal, dada por adverbios simples y conjunciones (ejemplo: *Paco fue ayer a la playa. Los bomberos llegaron después que la policía*).

Ausencia de figuras retóricas.

NIVEL 2 – FÁCIL

Textos breves de hasta 250 palabras, interesantes para que el lector descubra/aprenda algo nuevo o para su tiempo de ocio.

Conceptos más específicos, pero de interés para el lector (profesional, deporte, aficiones).

Contextos profesionales y de comunicación mediática (por ejemplo, instrucción profesional, noticias, anuncios, deporte).

Palabras y expresiones frecuentes que se refieren a asuntos cotidianos, pero también a otros ámbitos más específicos (por ejemplo, *curso, año escolar*). El texto contiene algunas palabras que se refieren a **ideas abstractas** más frecuentes (*felicidad, imaginación, afecto*).

En el texto se utilizan **abreviaturas comunes después de explicarlas** (por ejemplo, *litro - l / metro - m*). Uso de **extranjerismos comunes**, pero no en la mayoría (*estrés, rock*).

En el texto se utilizan **palabras compuestas comunes cuyo significado puede entenderse combinando las dos partes** (*gafas de sol, grosero, mediana edad, maltrato*). Palabras con afijos frecuentes (*jardinero, organización, finísimo*).

Oraciones simples y complejas utilizando conjunciones coordinantes y subordinantes (por ejemplo: *los músicos estaban tocando, pero pocos clientes escuchaban. Cuando dejaron de tocar, sólo habló un señor*). El texto contiene **oraciones de relativo que se refieren al sujeto de la frase** (por ejemplo, *El señor que vive al lado oyó el ruido*). El texto puede contener frases presentadas en un orden no lineal (Complemento-Verbo-Sujeto) (por ejemplo: *Los autobuses de la excursión escolar llegaron al colegio*).

Posible uso de la **voz pasiva**.

Uso de **diferentes tiempos verbales y modos**.

Modos **Imperativo y Gerundio**.

En el texto **se omiten algunos elementos**.

Secuencias o localizaciones temporales lineales, dadas a través de **adverbios de tiempo o conjunciones**, pero también mediante tiempos verbales (p. ej. *él está de camino y la fiesta ya ha empezado*).

Se utilizan **figuras retóricas**.

En el texto se recurre a la **metonimia** (por ejemplo, *El palacio no quiso negociar la rendición*), así como a metáforas familiares (por ejemplo, *mi padre tiene ojos de águila*).

NIVEL 3 – CLARO

Textos hasta 500 palabras, que enseñan a llevar a cabo tareas nuevas para el lector.

El texto se refiere a **conceptos que requieren juntar distintas informaciones**. (por ej., *estrellas, volcanes*) y a **temas variados** de interés para el lector para su información o su ocio.

El texto remite a **conceptos o ideas abstractos**. Se usan palabras polisémicas y abreviaciones, siglas y acrónimos corrientes, así como palabras extranjeras corrientes.

En el texto hay alguna palabra compuesta con significado no decodificable analíticamente o palabras derivadas con morfemas (prefijos y sufijos) corrientes.

El texto presenta una **estructura mayoritariamente paratáctica**; no se sigue siempre el orden básico de palabras. Las oraciones del texto con frecuencia se articulan con **conectores de coordinación y de subordinación** y en ocasiones con conectores menos comunes (por ej., *La casa es bonita, pero resulta cara de acuerdo con lo que dicta el mercado inmobiliario*), así como incluyen oraciones de relativo que complementan al sujeto o al objeto de la oración (por ej., *El chico que me abrazó se dirigió a la entrada. El chico al que su madre abrazó se dirigió a la entrada*).

Los verbos están conjugados en subjuntivo (presente, pretérito perfecto o pluscuamperfecto), en futuro compuesto y lo condicional.

El texto presenta un uso frecuente de la **pronominalización** y elementos omitidos.

El texto presenta distintas **secuencias temporales no lineales** (por ej., *María y Pablo fueron muy felices mientras estuvieron casados; de hecho, nunca habían sido tan felices como desde que se casaron, pero la felicidad duró hasta que ella se enamoró de un compañero de trabajo al que había conocido dos años antes*).

En el texto hay sentidos figurados no muy comunes (por ej., *El Barça barrió al Madrid*).

1) Cuadro de los niveles

		Descriptores	Muy fácil	Fácil	Claro	Más complejo
<i>Tamaño</i>	1	El texto puede tener hasta 50 palabras.				
	2	El texto puede tener hasta 250 palabras.				
	3	El texto puede tener hasta 500 palabras.				
	4	El texto contiene más de 500 palabras.				

		Descriptores	Muy fácil	Fácil	Claro	Más complejo
<i>Diversidad conceptual</i>	5	El texto menciona conceptos relacionados con actividades cotidianas (por ejemplo, <i>comida, tareas domésticas, ropa</i>).				
	6	El texto hace referencia a algunas ideas abstractas.				
	7	El texto se refiere a conceptos que requieren juntar distintas informaciones. (por ej., <i>estrellas, volcanes</i>).				
	8	El texto se refiere a conceptos nuevos o con múltiples facetas. (por ej., <i>materia oscura, desarrollo cerebral, estadística</i>).				

		Descriptores	Muy fácil	Fácil	Claro	Más complejo
<i>Finalidad</i>	9	El texto sirve para llevar a cabo tareas corrientes/domésticas (por ej., instrucciones para comprar billetes de autobús, instrucciones de uso en la etiqueta de un producto).				
	10	El texto se utiliza en el ámbito laboral o en los medios de comunicación (por ej., comunicación o instrucciones en el trabajo, noticias, anuncios, deportes).				
	11	El texto enseña a llevar a cabo tareas nuevas (por ej., instrucciones sobre cómo usar el correo electrónico de la empresa).				
	12	El texto se utiliza en contextos de ocio (por ej., literatura y ficción, reportaje, artículos de opinión).				

		Descriptores	Muy fácil	Fácil	Claro	Más complejo
<i>Contexto comunicativo</i>	13	El texto se utiliza en comunicaciones corrientes (por ej., horarios de transportes, listas, menús, instrucciones generales, información sobre el precio de un producto, comunicación personal, etc.).				

	14	El texto sirve para llevar a cabo tareas corrientes/domésticas (por ej., instrucciones para comprar billetes de autobús, instrucciones de uso en la etiqueta de un producto).				
	15	El texto puede abordar temas variados de interés para el lector para su información o su ocio (por ej., artículo de revista sobre la naturaleza).				
	16	El texto se utiliza en contextos especializados (por ej., comunicación profesional formal, foros en Internet).				

		Descriptores	Muy fácil	Fácil	Claro	Más complejo
Concreción	17	El texto se compone sobre todo de palabras que se refieren a conceptos concretos (por ej., <i>silla, manzana, martillo, ordenador</i>).				
	18	El texto contiene algunas palabras que hacen referencia a ideas abstractas (por ejemplo, <i>felicidad, imaginación, amor</i>).				
	19	El texto utiliza palabras que se refieren a conceptos e ideas abstractas (por ejemplo, <i>envejecimiento, razón, lealtad</i>).				
	20	El texto se refiere a conceptos relacionados con usos de ámbitos específicos (por ej., <i>trabajo, deportes, aficiones</i>).				
	21	El texto utiliza palabras que se refieren a conceptos e ideas abstractas en contextos polifacéticos y/o áreas especializadas (por ejemplo, <i>poesía, democracia, fusión nuclear</i>).				

		Descriptores	Muy fácil	Fácil	Claro	Más complejo
Léxico	22	El texto tiene palabras y expresiones simples relativas a cuestiones cotidianas y se usan nombres frecuentes y comúnmente conocidos (por ej., <i>patatas, aceite de oliva, verduras</i>).				
	23	En el texto no se utilizan abreviaciones (por ej., <i>gr., km., kg.</i>), salvo las formas frecuentes de tratamiento (por ej., <i>Sr., Dra.</i>).				
	24	En el texto no se utilizan ni acrónimos ni siglas.				
	25	En el texto se utilizan pocos sentidos figurados, acrónimos, siglas y abreviaciones.				
	26	En el texto se utilizan abreviaciones corrientes tras explicarlas (por ej., <i>l. – litro, m. – metro</i>).				

	27	En el texto se utilizan acrónimos y siglas que son comúnmente conocidos, esto es, se pueden entender de manera aislada (por ej., <i>RENFE, IVA, PSOE, PP</i> (<i>compañía ferroviaria, impuestos, partidos políticos...</i>)).				
	28	En el texto se utilizan abreviaciones corrientes (por ej., <i>m., km., gr., l., Excmo., ...</i>).				
	29	En el texto se utilizan acrónimos y siglas después de describirlos (por ej., <i>Organización de las Naciones Unidas (ONU), instituto de enseñanza secundaria (IES)</i>).				
	30	En el texto se utiliza todo tipo de siglas (por ejemplo, <i>ADN, UNESCO, UE, UNICEF</i>).				
	31	El texto tiene palabras y expresiones menos conocidas que se refieren a ámbitos específicos (por ej., <i>enfermedad cardíaca, arteria, tratamiento</i>).				
	32	En el texto no se utiliza ninguna palabra extranjera o se utilizan palabras extranjeras muy comunes y conocidas (por ejemplo, <i>internet, donut</i>).				
	33	El texto tiene palabras extranjeras comunes (por ej., <i>estrés, rock</i>), pero no son la mayoría en el texto.				
	34	En el texto se utilizan palabras extranjeras comunes con frecuencia (por ej., <i>link, hobby, marketing, mail, gay</i>).				

		Descriptor	Muy fácil	Fácil	Claro	Más complejo
Morfología nominal	35	En el texto se utilizan palabras compuestas comunes o palabras compuestas transparentes, es decir, palabras o expresiones que pueden entenderse fácilmente combinando las dos partes.				
	36	El texto tiene palabras compuestas corrientes (por ej., <i>judías verdes</i>) o palabras compuestas transparentes (palabras o expresiones cuyo significado se puede entender fácilmente combinando el significado de sus partes (por ej., <i>abrelatas, sacacorchos</i>).				
	37	El texto tiene palabras derivadas con prefijos y sufijos corrientes y frecuentes (por ej., <i>infeliz, tranquilamente, organizar, trabajador, casita</i>), pero no son mayoritarias en el texto.				
	38	El texto tiene palabras derivadas con prefijos y sufijos que son frecuentes (por ej., <i>precocinado, desorganizar, embalaje, organización, jardinero, finísimo</i>), pero no son mayoritarias en el texto.				

	39	El texto tiene alguna palabra compuesta cuyo significado no se puede fácilmente entender a partir de la combinación del significado de sus partes (por ej., <i>portaaviones</i> , <i>portavoz</i> , <i>ópera prima</i>).				
	40	El texto tiene palabras derivadas con sufijos y prefijos muy variados (por ejemplo, <i>nerviosismo</i> , <i>lavable</i> , <i>curiosidad</i> , <i>altitud</i> , <i>superpoder</i> , <i>preconcebido</i>).				
	41	El texto tiene distintas palabras compuestas cuyo significado no se puede entender fácilmente a partir de la combinación del significado de sus partes (por ej., <i>plusvalía</i> , <i>caballo de batalla</i> , <i>en pie de guerra</i>).				
	42	El texto tiene palabras derivadas con sufijos y prefijos muy poco comunes (por ej., <i>megafiesta</i> , <i>tendinitis</i> , <i>geopolítico</i>).				

		Descriptores	Muy fácil	Fácil	Claro	Más complejo
Morfología verbal	43	Modos e tiempos verbales				
		Verbos en infinitivo (por ej., <i>gustar</i> , <i>comer</i> , <i>huir</i>).				
		Verbos en presente de indicativo (por ej., <i>gusta</i> , <i>come</i> , <i>huye</i>).				
		Verbos auxiliares <i>poder</i> , <i>ir</i> , <i>estar</i> en los tiempos del indicativo (por ej., <i>puede gustar</i> , <i>voy a comer</i> , <i>está huyendo</i>).				
		Verbos en imperativo (por ej., <i>mira</i> , <i>oíd</i> , <i>estudiad</i>).				
		Verbos en gerundio (por ej., <i>gustando</i> , <i>comiendo</i> , <i>huyendo</i>).				
		Verbos en imperfecto de indicativo (por ej., <i>gustaba</i> , <i>comía</i> , <i>huías</i>).				
		Verbos en pretérito indefinido (por ej., <i>gustó</i> , <i>comí</i> , <i>huyeron</i>).				
		Verbos en pretérito perfecto (por ej., <i>ha gustado</i> , <i>has comido</i> , <i>han huido</i>).				
		Verbos en futuro simple (por ej., <i>comerán</i> , <i>gustará</i>).				
		Verbos en presente del subjuntivo (por ej., <i>guste</i> , <i>coma</i> , <i>huyan</i>).				
		Verbos en imperfecto de subjuntivo (por ej., <i>gustara</i> , <i>comiesen</i> , <i>huyera</i>).				
		Verbos en pretérito perfecto del subjuntivo (por ej., <i>haya ido</i> , <i>hayas oído</i> , <i>hayan estudiado</i>).				
		Verbos en futuro compuesto del indicativo (por ej., <i>habrá gustado</i> , <i>habrás comido</i> , <i>habré huido</i>).				
		Verbos en condicional simple (por ej., <i>gustaría</i> , <i>comerías</i> , <i>huirían</i>).				

	Expresiones con verbos como <i>permanecer, querer, poder, conseguir</i> en tiempos simples (por ej., <i>permanezca callado, quería almorzar, consigo ganar</i>).				
	Verbos en pluscuamperfecto de subjuntivo (por ej., <i>hubiera gustado, hubiese comido, hubiera huido</i>).				
	Verbos en condicional compuesto (por ej., <i>habría gustado, habrías comido, habrían huido</i>).				
	Expresiones con verbos como <i>estar, ser, permanecer, querer, poder, deber</i> en tiempos compuestos (por ej., <i>habían querido estudiar; han estado buyendo, habíamos debido huir</i>).				

		Descriptores	Muy fácil	Fácil	Claro	Más complejo
Sintaxis	44	La mayoría de las oraciones del texto son cortas y directas o se articulan por medio de conectores frecuentes y comunes (por ej., <i>La casa es grande, bonita y además es barata</i>).				
	45	El texto combina oraciones utilizando conjunciones coordinantes y subordinantes (por ejemplo: <i>los músicos tocaban, pero pocos clientes escuchaban. Cuando dejaron de tocar, sólo un caballero habló</i>).				
	46	Las oraciones del texto con frecuencia se articulan con conectores de coordinación y de subordinación y en ocasiones con conectores menos comunes (por ej., <i>La casa es bonita, pero resulta cara de acuerdo con lo que dicta el mercado inmobiliario</i>).				
	47	Las oraciones del texto son largas y se componen de varias oraciones que se articulan por medio de conectores de coordinación y de subordinación (por ej., <i>Dado que gran parte de este libro habla de la gestión de las partes difíciles del trabajo en solitario y analiza los modos de corregir o de ir más allá, primero quiero dejar bien claro que trabajar en solitario es fantástico, por lo que muchas personas lo consideran una opción</i>).				
	48	La mayoría de las oraciones del texto están en voz activa (por ej., <i>El gato ha cazado el ratón</i> (activa) vs. <i>El ratón ha sido cazado por el gato</i> (pasiva)).				
	49	El texto presenta frases en voz pasiva.				
	50	El texto contiene construcciones impersonales comunes y pasivas -se (p. ej., <i>El trabajo se hace bien</i>) (por ej., <i>Se detuvo a los ladrones en dos horas</i>).				
	51	El texto contiene oraciones de relativo que se refieren al sujeto de la frase (por ejemplo, <i>el señor que vive al lado oyó el ruido</i>).				

52	Las oraciones del texto con frecuencia incluyen oraciones de relativo que complementan al sujeto o al objeto de la oración (por ej., <i>El chico que me abrazó se dirigió a la entrada. El chico al que su madre abrazó se dirigió a la entrada.</i>).				
53	El texto contiene oraciones simples y en el orden básico de palabras (sujeto-verbo-complementos) (por ej., <i>Pablo ayudó a sus hijos a comprar una casa.</i>).				
54	El texto contiene oraciones simples que no siguen el orden básico de palabras (sujeto-verbo-complementos) (por ej., <i>Al niño le han regalado una bicicleta sus padres</i>), pero no son mayoritarias en el texto.				
55	Las oraciones del texto siguen distintos órdenes de palabras (sujeto-verbo-complementos; verbo-complementos-sujeto, etc.) (por ej., <i>Compraron los niños azúcar para hacer la tarta; Perdieron la guerra los que la habían iniciado.</i>).				
56	Las frases del texto se presentan en diferentes órdenes no lineales (sujeto-verbo-complemento; verbo-complemento-sujeto, etc.).				

		Descriptores	Muy fácil	Fácil	Claro	Más complejo
Cohesión textual: pronombres, anáforas y elipsis	57	El texto tiene pocos o ningún elemento omitido (por ej., <i>Comienza una semana que te traerá abundantes alegrías y realizaciones tanto en los asuntos laborales, materiales y sociales como también en tu vida íntima, todo ello gracias a un tránsito muy favorable del Sol. Viajes felices y afortunados, tanto si son de trabajo como por motivos de ocio.</i>).				
	58	Hay elementos repetidos en lugar de pronombres (por ej., <i>A Juan no le gusta Maria porque María no es amable.</i>).				
	59	El texto tiene algunos elementos omitidos.				
	60	El texto tiene elementos omitidos (por ej., <i>La diabetes es un grave problema de salud pública en los países del primer mundo. Además, es una enfermedad silenciosa vs. La diabetes es un grave problema de salud pública en los países del tercer mundo. Además, esta dolencia es una enfermedad silenciosa.</i>).				
	61	El texto presenta una secuenciación temporal lineal simple, expresada por medio de adverbios temporales o conectores como <i>desde que</i> , <i>en cuanto</i> , <i>después de que</i> y por medio de tiempos verbales (por ej., <i>Cuando llegó María, la fiesta ya había comenzado.</i>).				

	62	El texto presenta distintas secuencias temporales no lineales (por ej., <i>María y Pablo fueron muy felices mientras estuvieron casados; de hecho, nunca habían sido tan felices como desde que se casaron, pero la felicidad duró hasta que ella se enamoró de un compañero de trabajo al que había conocido dos años antes</i>).				
	63	El texto utiliza pronombres y palabras relacionadas para indicar a qué se está haciendo referencia (por ejemplo: <i>En el libro, damos una definición completa de una tormenta de nieve y de cómo se forma. También indicamos dónde se puede ver el fenómeno</i>).				
	64	En el texto se usan pronombres y distintos tipos de palabras para indicar a qué se refiere (por ej., <i>Tal y como especifica la plataforma especializada en proporcionar información sobre el tiempo, las temperaturas en España “van a ir subiendo” día a día, y no solo las máximas, sino también las mínimas. Un hecho que provocará que sea más difícil que éstas disminuyan durante la noche, propiciando la aparición de las conocidas como “noches tropicales”, que, por el momento, ya se han dado en diversos puntos de nuestro país</i>).				
	65	El texto presenta secuenciaciones cronológicas no lineales.				

		Descriptores	Muy fácil	Fácil	Claro	Más complejo
Estilística	66	En el texto no hay sentidos figurados (por ej., <i>Este verano los precios subirán vs. Este verano se dispararán</i> (fig.)).				
	67	En el texto hay sentidos figurados (por ej., <i>Este verano los precios se dispararán</i>).				
	68	El texto tiene algunos casos de metáfora y metonimia.				
	69	En el texto hay sentidos figurados no muy comunes (por ej., <i>El Barça barrió al Madrid</i>).				
	70	El texto sólo utiliza las comparaciones, personificaciones e hipérboles más comunes (por ejemplo: <i>Es tan bella como una flor; Tu maleta pesa una tonelada</i>).				
	71	En el texto se utilizan metáforas poco comunes (por ej., <i>Se destejen los días. Las noches se consumen antes de darnos cuenta</i>).				

Raquel AMARO is an Assistant Professor with tenure at NOVA University of Lisbon (NOVA FCSH) and Researcher at the Linguistics Research Center of NOVA University of Lisbon (CLUNL), since 2015. She served on the Board of Directors of CLUNL from 2017 to 2025, during which time the research unit progress to Excellent, and she became the coordinator of the Lexicology, Lexicography and Terminology group at CLUNL in 2025. She is currently the coordinator of the Horizon project iRead4Skills (2023-2026) and a researcher at the Horizon project MultiPod (2024-2027). She participated in more than 30 projects in the fields of Corpus Linguistics, Computational Lexical Semantics, Lexicography and Language Technologies.

Susana CORREIA completed her PhD in Linguistics in 2010 at the University of Lisbon, School of Arts and Humanities, where she later worked as a postdoctoral researcher. She is currently an Assistant Professor at the Faculty of Social Sciences and Humanities, NOVA University Lisbon (NOVA FCSH), where she teaches undergraduate and graduate courses in Linguistics, including Psycholinguistics, Phonetics and Phonology, Language Acquisition and Disorders, Methods in Language Teaching, and Experimental Methods in Language Acquisition Research.

She supervises several master's and doctoral students, two of whom hold grants from the Portuguese Foundation for Science and Technology (FCT). For six years, she directed the Portuguese as a Foreign Language program at NOVA FCSH, coordinating courses offered both at the university and through partnerships with national and international institutions such as the Camões Institute, Emory University, and the Luso-American Development Foundation.

Her research focuses on how children and adults acquire and use the sound structure of first and second languages, as well as on assessing linguistic competence to promote better language learning and communication outcomes. She has published in journals including *Language and Speech*, *First Language*, *Frontiers in Psychology*, *Written Language & Literacy*, and *Studies in Second Language Acquisition*.

Susana collaborates with international research teams across Europe and North America and has led the FCT-funded project ProPerL2 on L2 speech learning. She is also co-Principal Investigator of iRead4Skills, a Horizon Europe project uniting five universities and three non-academic partners.

Ricardo MONTEIRO is a PhD student in the branch of Lexicology, Lexicography and Terminology. His MA's dissertation was developed in the context of the topic of Textual Complexity through the creation of a framework that aimed at gauging complexity in texts for low-literacy adults. Beyond this topic of interest, his PhD project revolves around Lexical Semantics and Computational Linguistics, namely the modelling of Metaphor.

Alice PINTARD holds a Master's degree in Language Didactics (2019) from Le Mans Université, where she specialized in research and higher education, focusing on text readability and cultural mediation, and a Master's in French as a Foreign Language (2014). She worked as a teaching assistant at the Université catholique de Louvain, where she joined the iRead4Skills project as a researcher. Her research interests lie at the intersection of language technology and linguistics, with a particular focus on machine learning and automatic readability assessment, lexical complexity prediction, and the development of

corpora and language resources to support research in language learning and accessibility. She is currently an aeronautical technical writer at Daher.

Michell GADELHA MOUTINHO holds a PhD and a Master's degree in Language/Culture Learning and Teaching from the Graduate Program in Letters at the Federal University of Pará (UFPA), Brazil. He has been teaching German as a foreign language to children and adolescents at Escola de Aplicação da UFPA since 2011 and also has taught English and German in other educational contexts. Since 2012, he has served as an assessor for German language examinations administered by the Goethe-Institut. His research interests include: reading fluency, readability, curriculum-based measurement, and language learning and teaching.

Sílvia BARBOSA holds a PhD in Lexicology, Lexicography and Terminology from the School of Social Sciences and Humanities, NOVA University Lisbon, supervised by Professors Raquel Amaro, William Martinez, and Teresa Lino. Her doctoral research focuses on the proposal for the development of a specialized dictionary of descriptors used in wine tasting notes. She also holds a Master's degree in Phonetics and Phonology and a Bachelor's degree in Linguistics from the Faculty of Letters, University of Lisbon. She has contributed as a researcher to several linguistic projects, including the Corpus of Spontaneous Speech in European Portuguese (2006), the Portuguese Language Portal (2006–2010), the Vocabulary of the Portuguese Language (2009–2010), the Portuguese Unisyn Lexicon Project (2010–2013), the European Portuguese–Standard Arabic Dictionary Project (2017–2020), the Humanities Going Digital Project (2021–2023), and the iRead4Skills Project (2023–2024). She is currently an Invited Assistant Professor at the School of Social Sciences and Humanities, NOVA University Lisbon, where she teaches Lexicology and Lexicography.