

Horizon Europe



D3.1 iRead4Skills Complexity Levels

Document Control

Information:

Settings	Value
Deliverable No.	D3.1
Document Title:	D3.1 iRead4Skills Complexity Levels
Author(s):	Ricardo Monteiro, Raquel Amaro, Susana Correia, Alice Pintard, Roser Gauchola, Michell Moutinho
Reviewer(s):	Xavier Blanco Escoda
Sensitivity:	Public
Date:	20/12/2023

Document Location: The latest version of this controlled document is stored in OneDrive-
 fcsh.unl.pt/iRead4Skills/Project/WorkPackages/WP3/D3.1_iRead4SkillsComplexityLevels

TABLE OF CONTENTS

- 0. GLOSSARY3
- 1. INTRODUCTION3
- 2. TEXT COMPLEXITY4
 - 2.1. Reader-task dimension4
 - 2.2. Quantitative dimension4
 - 2.3. Qualitative dimension5
 - 2.4. In tandem approaches6
- 3. TOOLS FOR LANGUAGE PROFICIENCY ASSESSMENT7
 - 3.1. Common European Framework of Reference - CEFR7
 - 3.2. Survey of Adult Skills – Program for the International Assessment of Adult Competencies (PIAAC)7
 - 3.3. Association of Language Testers in Europe (ALTE)8
- 4. METHODOLOGY8
 - 4.1. Text complexity levels9
 - 4.2. Validation process13
- REFERENCES17
- APPENDIX I - SET OF COMPLEXITY DESCRIPTORS TO BE VALIDATED19
- APPENDIX II - PREPARATORY TASKS24
- APPENDIX III - COMPLEXITY LEVELS DESCRIPTIONS FOR FRENCH77
- APPENDIX IV - COMPLEXITY LEVELS DESCRIPTIONS FOR PORTUGUESE85
- APPENDIX V - COMPLEXITY LEVELS DESCRIPTIONS FOR SPANISH96

0. Glossary

complexity: the inherent quality that obstructs the process of retrieving meaning.

comprehensibility: the quality of something being perceived, deciphered and understood by an individual.

feature: characteristic that is relevant for determining or inferring patterns or classes.

linguistic competency: the knowledge of a language used in communicative language activities (speaking, listening, reading, and writing) which allows its users to form grammatical utterances.

language proficiency: linguistic competence drawn upon communicative skills (linguistic, sociolinguistic and pragmatic) activating appropriate communicative strategies.

literacy: information processing skill that include oral reading fluency, reading comprehension, spelling and writing.

readability: text property that depends on factors intrinsic to the text and on factors related to the reader, and that concerns the reader ability to identify, understand, interpret, and engage with printed and written materials.

reading skills: set of abilities used in decoding and interpreting written language in texts; it is measured in terms of reading fluency and reading comprehension.

1. Introduction

In the present day, reading skills are widely recognized as essential for individuals to thrive in the world. Both UNESCO (2020) and OECD (2013b) define literacy as the process of engaging and understanding written texts in order to achieve one's goals and develop one's potential. Along the same lines, high literacy levels have been proven to be beneficial to other individual skills, showcasing its transversal aspect. According to OECD (2013a), individuals whose literacy levels are high are not only more likely to hold better-paid jobs but also to have a higher quality of life, ranging from their health habits to the active participation in their community and their country's politics.

Furthermore, the understanding of literacy levels as situated on a continuum and thus having no objective end (UNESCO, 2020) has provided space for the discussion of improving adults' literacy through the implementation of Adult Learning programs (ANQEP 2021). With the goal of providing opportunities for education in the context of vocational professional training, these programs make use of reading materials for teaching purposes. However, finding reading materials that are adequate for a low-literacy population is problematic.

In this context, the iRead4Skills project aims to contribute to the solution of this problem by devising an intelligent reading system that evaluates texts complexity and suggests appropriate readings according to the user's reading skills. To achieve this, the notion of text complexity, or what it entails when concerning low-literacy adult native speakers, as well as how it manifests/can be detected in texts needs to be attended to.

This report discusses text complexity, the existing frameworks for text complexity, and elaborates on our own framework for text complexity gauging.

Following the method undertaken for defining the complexity levels, this report is organized as follows: Section 2 discusses the relevant dimensions of text complexity and the approaches to its assessment. Section 3 presents the tools and frameworks we used to inform our proposal. Section 4 presents the method followed for identifying complexity levels. It includes i) a detailed discussion on the establishment of the relevant complexity levels, ii) the definition of the set of descriptors used for the three work languages - French (FR), Portuguese (PT) and Spanish (SP) – iii) and the process followed to validate and adjust the levels and descriptions by Adult Learning (AL) and Vocational Educational Training (VET) professionals. Results obtained through queries and

dedicated focus groups for the three languages are presented. The final and complete descriptions of the complexity levels addressed in the iRead4Skills project are attached to the present report.

2. Text complexity

Cunningham & Mesmer (2014:256) establish text complexity as features inherent and exclusive to a text that can be analysed and studied separate from the reader's difficulties with it. This understanding falls in line with Dahl's (2007) interpretation of complexity, which, in essence, considers text complexity as the amount of information needed to recreate a specific linguistic manifestation to determine how complex that specific linguistic manifestation is. An example of this is how grammatical gender is marked in many languages, thus embedding more information within the word, while in others it is not.

On a different perspective, the CCSS/ELA (*Common Core State Standards/English Language Arts*) (2010) defines the notion of text complexity as the "inherent difficulty of reading and comprehending a text combined with consideration of reader variables". This understanding addresses text complexity as a relative phenomenon, which besides depending on the text/language, it also varies among readers (Kusters, 2003). The CCSS/ELA (2010) further expands on their definition and proposes a threefold system in which texts are gauged on their complexity through i) a reader-task dimension; ii) quantitative components, and iii) qualitative components.

2.1. Reader-task dimension

The CCSS/ELA (2010) establishes the reader-task component of text complexity as the idiosyncrasies that the reader brings to the table combined with what reading tasks are demanding from the reader. In other words, this understanding is supported by the idea of a text being relatively complex, by considering not only readers' proficiency levels as diverse, but also the way texts need to be comprehended.

When it comes to the reader dimension, Flor et al. (2013) states that reading patterns become richer with age, being distinct in children and adults. This means that low-literacy adults, although ineffective in understanding specific texts, will possess implicit but rich knowledge of their language that children do not. Thus, adults' reading difficulties are not equivalent to that of children.

On the other hand, the task dimension, i.e., what the reader has to do to comprehend a text or what the purpose of the text is, can play a big role in the ability to comprehend the text (Goldman & Lee, 2014). Although these tasks are usually never explicitly formulated to the reader, further in this document we will add that a hierarchy of difficulty exists among them.

2.2. Quantitative dimension

Interest in how written texts convey their ideas and how these are perceived by the readers is something almost as old as written communication itself. However, it was only by the end of the 19th century that a formal and consistent study of this topic started to be investigated by experts (Hiebert & Pearson, 2014). At this time, the main emphasis in analysing text complexity was essentially qualitative, focusing on text features and pinpointing how these impacted the understanding of the text by the reader.

As technology brought forth new advancements, new scientific approaches arose, with the issue of comprehensibility of a text being analysed through a quantitative lens. These approaches would consist of formulae mainly contemplating word frequency and sentence length (Gunning, 1952; Flesch, 1948, Kincaid et al., 1975; McLaughlin, 1969; Dale and Chall, 1948; Spache, 1953; Lorge, 1948). Grounded in early computational models, early approaches would count the occurrences of linguistic complexity markers that were established beforehand (e.g., number of syllables in a word, number of words in a sentence, frequency of complex words, etc.) which would result in a numerical score of complexity of the text (Goldman & Lee, 2014).

These readability formulae were (and still are) the most common indexes used to evaluate the complexity level of a text, operating as the cornerstone for school curriculums like the CSSS/ELA (2010), as well as having been adapted to other languages (Hiebert & Pearson, 2014; Rabin, 1988).

According to Hiebert (2011:3), researchers have proposed more than 200 readability formulae. However, given their quantitative nature, these formulae tend to determine syntactic complexity in a text through the metric of words per sentence; and semantic complexity through the frequency of the occurrence of a word compared to a pre-established word list, resulting in the metric of word familiarity/rarity (Flor et al., 2013). This can in fact lead to an oversimplification of how complexity features can materialize in texts that, together with the fact that new readability indexes usually expanded on each other rather than proposing new metrics, has prompted the question “How precisely can readability formulae predict a reader-book match?” (Hiebert & Pearson, 2014).

By the mid-1980s, the critique was that readability indexes, while offering results that could be interpreted later, were deficient in their results (Hiebert & Pearson, 2014). As assumed by McNamara et al. (2010) and Hiebert (2011), certain components of readability formulae are certainly valid as early indicators of text complexity. However, measures like word and sentence length cannot be the only measures relied on to disclose the complexity embedded in a text.

These formulae have been criticized on the base of missing (or inflating) complexity features present in different genres of texts (Hiebert, 2011). An example of this is how narrative texts tend to make use of big sections of dialogue where short sentences are usually predominant, as they supposedly mimic the way people speak to each other in real life. Considering that a readability formula usually considers short sentences as not complex, the results of narrative texts tend to be underestimated. As a result, not only are the specificities of certain genres of texts neglected in this evaluation, but also certain texts, seem to be “digested” to a specific score when parts of these texts (e.g., a novel) would be given a “non-complex” score and others would be given a “complex” score.

Ultimately, quantitative approaches are dependent on their subsequent interpretation. Chall herself, one of the developers of one the most widely used readability formula (Chall & Dale, 1948), famously noted how there was nothing in the formula that would lead to its misuse, thus implying that the problem lied on the lack of interpretation of its results (Hiebert, 2011).

2.3. Qualitative dimension

In the 1980s, with the departure from the dominant trend of behaviourism to a more cognitive-based psychology, the debate on the topic of text complexity changed as well (Hiebert & Pearson, 2014). The focus was no longer put on the complexity features themselves but on the way they affected the comprehension of the reader. This change in how text complexity was perceived led to the subsequent change of situating the reader as the centre of text complexity and not the text.

Qualitative measures were then revived. Schema-based theories, mental model theories of reading as well as story grammars describing an inherent structure present in stories were all proposed to further understand human comprehension of a text (Harley, 2013). Equally important, they seemed to go beyond the features of text complexity proposed by the readability models used up until then which would limit themselves to surface level features of texts. Qualitative methods of text complexity analysis like text levelling, rubrics and text maps supported themselves on these theories of text comprehension and were put in practice (Pearson & Hiebert, 2014).

The first approach, text levelling, consists of creating levels of text complexity based on the knowledge and judgment of experts in the area to then relate these levels to different levels of reading proficiency. Although at first glance impressionistic, text levelling is probably the oldest qualitative measure of text complexity in existence and has been recurrently adopted through the assembly of expert language teachers, ever since the first critiques were made to readability formulae (Hiebert & Pearson, 2014).

A more developed version of text levelling is a rubrics system. In this method of qualitative assessment of

complexity, experts identify a set of features in texts that can be seen as complexity-inducing. These features are then put on a continuum where less complex manifestations of that feature rank on one end and more complex manifestations rank on the opposite end (Pearson & Hiebert, 2014). An example of a feature of the kind would be a word and its affixes: a simple word would rank as low-complexity whereas a complex or compound word would rank as high-complexity. Qualitative indexes of the kind are QATD (Chall et al., 1999) or the ACT (ACT 2006).

Turning to a more “invisible” complexity in texts, text maps shine a light on the conceptual structure of the text (Pearson & Hiebert, 2014). Expanding on the notion of story grammars, the purpose of text mapping is to reveal the conceptual structure of a text and highlight its complexity or lack thereof. Moreover, this qualitative measure of a text allows us to properly identify the type of maps associated with certain text genres. As Pearson & Hiebert (2014) refer, text maps make it possible to identify, in narrative texts, the theme, the plot, the setting, the structure, the characters, and even abstract concepts that are alluded to in the narrative but never explicitly mentioned. When it comes to nonnarrative texts, such as instructional texts, their maps tend to point out how the information is organized and its hierarchy when it comes to convey a message.

In essence, the simulation of human intelligence and its grasp on the different ways complexity manifests itself in texts is what any qualitative index strives for. By the early 2000s, digitalization gave way to more sophisticated qualitative analyses of complexity in texts. Coh-metrix is the latest (McNamara et al., 2010), although it is also based on a quantitative analysis of specific text features.

Qualitative measures of complexity are not without criticism. The holistic approach of these methods has been criticized for having the “blinding purpose” of exclusively assigning levels to texts, without specifying how it actually reaches that goal. The weight given to the different complexity features in these holistic views is never put forth (Pearson & Hiebert, 2014) and the use of qualitative indexes fails to suitably relay the information that teachers need to support and advise students through texts (Hiebert & Pearson, 2014). The attempt of qualitative indexes to break from the numeric and mathematical results of quantitative measures has led them to sometimes disregard their own methodology.

2.4. In tandem approaches

Despite the ongoing revival and application of qualitative measures, readability/quantitative measures are not to be counted out. During early 2000s, the same period in which qualitative measures were being paid close attention to, readability measures were being digitalized and becoming conveniently available to teachers as well as school publishers, thus becoming an influential part of their decisions on what texts would be part of their curriculums (Hiebert & Pearson, 2014). As a result, quantitative indexes were not replaced by qualitative ones. Quantitative measures were seen as rudimentary and did not instil the same confidence as qualitative indexes did. However, quantitative measures provided numerical and non-interpreted but factual results (Pearson & Hiebert, 2014).

To compensate for the shortcomings of both types of measures, several authors advocate for a use of both in tandem (Hiebert, 2011; Hiebert & Pearson, 2014; Cunningham & Mesmer, 2014; Douglas & Miller, 2016; Goldman & Lee, 2014): quantitative measures have no “real meaning” until they are fulfilled or analysed in a qualitative manner (Hiebert & Pearson, 2014). In other words, the results of quantitative measures are purely numbers that need to be interpreted. A good example of this was evidenced by Douglas & Miller (2016:4) when, to assess complexity at the sentence level, they employed two measures – one qualitative (AI’s Web-based L2 Syntactic Complexity Analyzer) and one quantitative (Lexile). The latter was deemed insufficient as it operated solely under the assumption that bigger sentences were more complex. Nevertheless, it provided valuable information that was then supplemented with a qualitative measure, thus yielding more trustworthy results.

This combination of measures is not only beneficial for the validity of results, but it can also be employed as an innovative way of identifying and processing certain text complexity features that have not been analysed yet (Hiebert & Pearson, 2014). In fact, some texts (namely literary and narrative) seem to “elude” both quantitative and qualitative measures when it comes to the evaluation of their complexity as they can rank low in both

measures but remain complex for a vast population (Goldman & Lee, 2014).

Ultimately, text complexity manifests itself in different ways and dimensions, so it makes sense to employ different measures that consider these different dimensions. Quantitative indexes such as QATD or ACT can be used to get a general understanding of a text's difficulty (Hiebert, 2011). However, as Hiebert (2011:8) advocates, when assessing text complexity, the use of a single measure – regardless of its nature – can only lead to the understanding of text complexity as something that manifests in a unidimensional way.

3. Tools for language proficiency assessment

Many schemas have been proposed to evaluate language proficiency. These frames of reference can be defined by associating different language tasks to the different levels of language competency. The following section briefly describes the most widely used frames of reference, corresponding to the ones used for basing our own analysis, as well as their approaches to determining reading comprehension levels for readers.

3.1. Common European Framework of Reference - CEFR

The latest CEFR Companion volume (CEFR 2020) distinguishes reading comprehension as one reception activity of language, alongside oral comprehension, and audio-visual comprehension.

When it comes to reading comprehension, CEFR determines different levels regarding the task or purpose that the text in question calls upon the reader. They advocate that there are fundamental differences between, for instance, reading as a leisure activity and reading instructions. These different purpose categories are operationalized into five constructs: reading correspondence, reading for orientation, reading for information and argument, reading instructions, reading as a leisure activity.

The proficiency levels attributed to readers are described via descriptors adapted to each category. These are organized in a hierarchy from less proficient to more proficient as follows: Pre-A1, A1, A2, B1, B2, C1, C2. Consequently, the level of proficiency of a reader is determined by their ability to successfully perform the descriptor of each level. For example, the category reading correspondence considers that a reader has an A1 level of proficiency if he can understand short, simple messages on postcards.

CEFR also proposes an overall reading comprehension schema with the same structure. As mentioned above, each descriptor of this schema proposes a task to the reader. If they are successful in achieving it, they are considered to have that level of reading comprehension.

Although not overtly, the descriptors associated to each proficiency level gauge the linguistic competencies of the readers (syntax, lexical, morphological), as well as assessing their polyvalence on the genres of texts and the subjects that they can comprehend.

3.2. Survey of Adult Skills – Program for the International Assessment of Adult Competencies (PIAAC)

The Survey of Adult skills stems from a joint effort of many Organization for Economic Cooperation and Development (OECD) member countries and was developed and implemented under the direction of PIAAC (OECD 2013a). This survey aims at the assessment of the proficiency of the adult population in 3 different information-processing skills: problem solving in technology-rich environments, numeracy, and literacy. The latter is the one relevant for our goals and it is addressed here.

PIAAC's framework differentiates itself for being created specifically for the adult population. It attempts to take into account the respondents use of literacy in an assortment of generic skills, ranging from everyday activities and leisure to work-related contexts. As a result of this, PIAAC's understanding of reading comprehension encompasses a more thorough text-based dimension, in which it also considers text support

(digital or printed) and text format (text structure). Nonetheless, a reader-based descriptor, *cognitive strategies* (the task the reader must perform in order to fully comprehend the text) is also considered, as well as a context-based one – *context* (the areas of society in which the text circulates).

With the dimensions considered, PIAAC's Survey of Adult Skills proposes a schema to assess the reading comprehension proficiency levels of adults with two distinct aspects: the features of the text and the task of the reader. Like the CEFR framework, it then relates these two together, creating five different levels of proficiency which rank on a scale of 1-500 points. To exemplify, level 1 presumes texts of this level to bear linguistic features such as small extension, simple vocabulary, and a linear structure, as well as circulating in an everyday context and having a printed support. The same level also postulates that the task of the reader is the simple task of assessing and identifying the message of the text.

3.3. Association of Language Testers in Europe (ALTE)

ALTE has consistently put forth guiding and reference materials for language-related testing since the 1990's. Because of its European roots, a major part of its materials has been advertised as supplements to the CEFR guidelines, as well as being adapted to specific European languages. ALTE's reference material that is considered to be of importance for this specific work is their Can Do project (ALTE 2020).

With the emphasis being put on making the evaluation methods more objective, the Can Do project aims to complement CEFR's levels of language proficiency by providing a framework that can improve language testing. Moreover, the multilingual perspective is also explored with ALTE attempting to make their framework as universal as possible, having been adapted to 12 different languages.

The Can Do project proposes a schema in which CEFR's levels are adopted (A1, A2, B1...) and then matched with different purposes of use in society, for instance, Social & Tourism, Work, Study. The idea is that different areas of society conjure different texts and written language, which bear unique complexity features. With the purpose of making the framework as tangible as possible, each reading comprehension level is accompanied by a sentence-long summary of the texts and of the text features appropriate for that level's readers.

4. Methodology

Considering the overall goals of the iRead4Skills project, the methodology for establishing the framework for the complexity levels addressed in the project considered three major steps:

- i) Compiling and evaluating existing tools and frameworks relevant for text readability and complexity assessment.
- ii) Determining a base proposal for three text complexity levels relevant for our target audience (adult native speakers with low reading skills) and for our goals (to promote the development of reading skills through an innovative intelligent system that evaluates texts complexity and suggests reading materials adequate to the user reading level, which can also be used by trainers in the creation or adaptation of texts with the appropriate level of complexity for their individual students).
- iii) Validating and making the necessary adjustments to the base proposal to inform a thorough and objective characterization of the text complexity levels addressed by the project.

In general terms, this meant gathering information to juxtapose the descriptors from PIAAC, ALTE and CEFR proficiency levels into three levels that would reflect different reading complexity/difficulties for Spanish, French and Portuguese texts, basing the literacy skills on the literacy model of the first three levels from PIAAC (ranging from A1 to B1 CEFR levels). These, as discussed above, are broadly characterised both internationally and cross-linguistically with norm-referenced descriptors that bring reliability and comparability to teaching environments and assessment across different societal realities. The resulting set of descriptors were then presented for independent classification to trainers from adult learning (AL) and vocational educational training

(VET) centres in Belgium, France, Portugal, and Spain. The results of this task were then discussed in focus groups to understand and adjust, whenever possible, conflicting judgements. The result is a framework for text analysis and classification, consisting of language-dependent tables with the characterization of the complexity levels addressed in the project.

The following subsections will present the work done in more detail.

4.1. Text complexity levels

For the iRead4Skills project, we decided on three text complexity levels – **Very Easy** (corresponding to A1), **Easy** (corresponding to A2), **Plain/Clear** (corresponding to B1).

For the iRead4Skills project, we decided on three text complexity levels – **Very Easy** (corresponding to A1), **Easy** (corresponding to A2), **Plain/Clear** (corresponding to B1).

The grading system was influenced by the CEFR's frame of reference (Council of Europe, 2020), namely in its labels (A1, A2, B1) and the naming system chosen taking into account two criteria: versatility and transparency.

The first aspect, versatility, stems from the fact that, in Europe, CEFR is the most widely used frame of reference for languages. All things being equal, future users of our tool of complexity analysis would be able to understand the grading system in a much easier fashion this way, as well as adapting it to other grading systems.

The second aspect, transparency, was decided considering simplicity and other already existing initiatives (e.g., Plain language). Given the potential target users of the iRead4Skills system, we consider that transparent and familiar names, characterizing the texts and not the readers, will contribute to the use of the system.

The reason for postulating these three levels of text complexity is also linked to the target audience of the project. Having considered the levels of existing frameworks of reading proficiency and their complexity descriptors (as well as our own complexity descriptors discussed below), the first three levels, starting in A1 (and not pre-A1 or A0 which are closer to the alphabetization process), were the i) relevant levels for developing and promoting reading skills and habits, on the one hand, and ii) the ones that were more problematic for trainers, since these reading materials are not easy to find, on the other hand.

The levels of complexity defined for the project can be functionally described as follows:

Very easy: Texts that are fully or almost fully understood by everyone, including people with very low schooling (i.e., that did not finish the primary school (ca. 6th year)) and almost no reading experience. It roughly corresponds to CEFR A1 level.

Easy: Texts that are fully or almost fully understood by people with low schooling (i.e., that completed the primary school but do not have more than the 9th year) and have poor reading experience. It roughly corresponds to CEFR A2 level.

Plain: Texts that are understood the first time they are read by people that completed the 9th year and have a functional-to-average reading experience. It roughly corresponds to CEFR B1 level.

4.1.1 Complexity descriptors

Once the target levels were established, we turned to the texts themselves. Beck et al. (1984) condense text complexity to three types of problems that readers might face: *surface problems* – complex syntactic structures and awkward use of connectors; *knowledge problems* – reader's lack of knowledge of the topic of the text as well as inability to understand the meaning in what is referenced in the text; *content problems* – counterintuitive or confusing relationships between the content of the text.

According to the reference frameworks consulted, text complexity is usually described considering several aspects such as the lexical complexity, related to the purpose of the texts and the topics addressed by the text, syntactic structure, and cohesion.

The CEFR framework tackles the issue of lexical complexity through their scale of “vocabulary range”. In this

schema, the CEFR considers the vocabulary extent of an individual (applied to both reading and writing tasks) as the means of grading them on the scale. More importantly, it connects this range in one's vocabulary to objective communication realms, such as one's familiarity with the topic or its concreteness. In this vein, Pearson & Hiebert (2014:178) advocate that the ideas presented in a text drive its complexity, both in structure and content domains.

In what concerns syntactic structure, across verbal and written communications, speakers of a language tend to show a synchrony between what they read/hear and what they write/say (Douglas & Miller, 2016). In other words, readers who are accustomed to reading texts in which, for example, passively constructed sentences are present will then be triggered to construct these same sentences as a means of their own communication. This mimicry between the content we read and how we use language has been evidenced to be particularly high in syntactic features (Douglas & Miller, 2016). In this perspective, a low-literacy individual would consider different syntactic structures to the one they use to be complex. However, this may not be the case for native speakers with low reading skills.

Cohesion is essential to the understanding of a text. According to Harley (2013), comprehension is a constructive process. Readers quickly forget details of surface form and constantly infer ideas, entities and their relationships in a text. Many are the authors that consider the lack of cohesion in a text as an important measure of its complexity (Harley, 2013; Fang & Pace 2013; McNamara et al., 2010; Hiebert, 2011).

Text cohesion refers to the way a text is internally organized. Comprehension problems arise when connection between ideas, sentences and parts of the text are not clear (Fang & Pace, 2013). This can manifest itself via the lack of connectors between sentences or via the use of implicit connections—referring to the text's grammatical cohesion – as well as through a deficit in the conceptual arrangement of the text, such as a lack of anaphora's antecedents or their clarity – referring to lexical/referential cohesion of the text.

Different cohesion strategies may cause the reader to infer ideas and relationships within the text, and low-proficiency readers have difficulties making these inferences (McNamara et al., 2010). Low-literacy individuals rely on a more overt or “strong” cohesion to avoid making these inferences (Hiebert, 2011). Unlike the cases of word and sentence length, in which more material increases the processing demands of a text, the more cohesion elements are present in a text, the more the reader's comprehension improves (McNamara et al., 2010).

Considering these aspects, it is possible to characterize texts of a given complexity level as exemplified in (i) below.

(i) **Very Easy:** *Short and simple texts for the purpose of performing familiar tasks, using typically simple, everyday concepts used in basic day-to-day contexts.*

However, given the goals of our project, it was necessary to densify this type of characterization, namely in what concerns the lexical-conceptual system, syntactic structures, and text cohesion.

Lexical-conceptual system

For instance, for the lexical-conceptual aspect of complexity we consider two dimensions: i) the traditional lexical rarity (word frequency & presence of simplex/complex words) as well as lexical density and tightness; and ii) the existing strong relationship between lexicon complexity and complexity of the ideas used in a text, since atypical wordings help to render a text abstract (Fang & Pace, 2013).

Moreover, figurative language is also regarded and measured as complexity-inducing. Both Fang & Pace (2013) and Harley (2008) consider figurative language to be potentially problematic, as readers may have to check for their literal meaning first and then contrast it with the context, which creates an extra load in their cognitive processes.

Obstacles in text comprehension at the lexical level can come up on account of language issues or topic and concept complexity (Fang & Pace, 2013). Nevertheless, they are the first to emerge when readers pick up a text (Harley, 2008). This framework recognizes the importance and weight of this measure, by considering it, from all the four measures, the one that potentially hinders the reader's understanding of a text the most.

One of the ways in which these ideas can be gauged on their complexity is on their level of concreteness. Concrete concepts have a material referent in the real world, whereas abstract do not. According to early literature (James, 1975), less abstract concepts are harder to process cognitively than concrete ones, especially for low-literacy readers. However, there are abstract concepts that are familiar and not hard to grasp, e.g., love, happiness.

Another way of analysing lexical-conceptual content is the reader's familiarity with the topic addressed in the text. A prior knowledge of the topic by the reader is sure to make it easier for him to understand the text in reading, whether it is by virtue of activating the appropriate concepts in a clear way or due to activating inferences more easily (Harley, 2013).

Although verbs could fall into the previous lexical-conceptual measure of complexity, due to the rich verb inflection of romance languages to which this framework is going to apply (Portuguese, French, Spanish), verbs were deemed as a measure of complexity relevant enough to be considered independently. The tenses and mood in which verbs manifest in texts are believed to be capable of generating complexity. While the presence of common and simple verb forms like the Infinitive or the Indicative Present Simple are considered uncomplicated, the use of less usual verb moods such as the Subjunctive or some compound verb forms are difficult to be recognized by the reader and are able to create temporal spaces and hypotheses which can be burdensome to process cognitively.

Syntactic structure

As mentioned before, speakers of a language tend to show a synchrony between what they read/hear and what they write/say (Douglas & Miller, 2016). However, low-literacy adult native speakers are language proficient and exposed to different language usages and contexts, meaning that the range of constructions at their disposal is expected to be much higher than in the case of children or of learners of a second language.

For this aspect, our framework considers traditional readability measures like sentence length as being able to induce complexity in a text, as well as qualitative manifestations of complicated syntax, such as a non-linear sentence structure, frequent and convoluted subordination and coordination. Moreover, structures like the passive voice, verb periphrases and modality are considered.

Text cohesion

Our framework attempts to measure complexity cohesion by considering the manifestation of referential cohesion (given by the presence of anaphora or ellipsis) as well as its scarcity and level of relationship to the antecedent.

Another aspect to consider in the domain of cohesion is how spatial-temporal information presented in the text can also help or hamper text comprehension. Readers construct a mental model of the entities in the text and their respective spatial-temporal information. Evidence shows that the spatiotemporal layout embedded in a text affects the processing of a text: the more intricate and complex it is – that is to say, the more spatiotemporal localizations there are –, the more it hinders the reader's comprehension of it (Glenberg et al., 1987).

Summary

According to the analysis of the several aspects referred above, our proposal for describing text complexity considers the following aspects:

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

Considering these aspects, and adding specific language-dependent markers, it is possible to describe a specific complexity level with some granularity, as exemplified for the Very Easy Level in (ii).

(ii) Very Easy level

GENERAL DESCRIPTION

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

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

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

Absence of figures of speech.

Basic lexicon (active lexicon) – known words and simple expressions memorized = 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 (PT *-mente*).

Short periods, with simple conjunctions and in direct 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., 'João doesn't like her because she's not nice' vs. 'João 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., PT *estar/começar/andar + a + Infinitive*; PT *deixar/acabar + de + Infinitive*; *ir + Infinitive*). No compound tenses.

Use of Indicative verb tenses, (PT except for the simple Future (compound tense - *ir + Infinitive* - or Present + adverb)). 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 and she had left').

SETS OF FEATURES

Lexical conceptual: basic lexicon (w/ absence of foreign words and less of 15% of non-basic vocabulary); mostly concrete nouns (concrete concepts); daily and familiar concepts (personal life)/ new simple concepts; rare affixation.

Verbs: simple tenses in Infinitive, Indicative, Gerund; mainly main verbs (no auxiliary); copulative verbs (PT *ser, estar, ficar, continuar, andar, parecer*); frequent multiword expression verbs in PT Indicative mode + Infinitive.

Syntactic structures: simple coordination (PT *e, ou, mas, nem, pois*); simple subordination (subject relatives, temporal - PT *antes de, depois de, quando, enquanto*; causative - PT *pois, porque*; conditional - PT *se* + Indicative).

Cohesion: few pronoun anaphora and ellipsis; one or two temporal location(s).

The idea was, thus, to achieve a more thorough, comprehensive and objective characterization of the levels of complexity to be addressed. From these characterizations, we compiled a set of 66 descriptors covering the following dimensions: text size, text purpose, communication context; conceptual diversity; concreteness; vocabulary/lexicon subset; morphology; morphosyntax; syntax; textual cohesion: anaphora and ellipsis; textual cohesion: temporal references; and text style. These descriptors, to which 15 referring to a more complex level were added (see Appendix I), were the bases for the validation process.

4.2. Validation process

The validation process consisted of gathering insights on text complexity and complexity descriptors from specific focus groups of experts for each language. The experts were recruited from AL and VET centres in Belgium, France, Portugal and Spain, where they worked as teachers in different fields: language (French, Portuguese or Spanish); natural sciences; maths; ICT. The different partners conducted these processes for each language following the previously discussed common approach:

1. Preparatory task (ca. 1 hour): before the first focus group meeting, the participants received a task (in the format of a survey, see Appendix II) concerning the descriptors in order to be able to validate them and give their opinions within 10 days upon receiving the survey.

Analyse the results before the focus groups and prepare a table with the consensual vs. problematic descriptors.

2. Focus group session I (ca. 1,5 hour): to discuss the descriptors and the results of the preparatory task. The focus groups sessions include:

- presenting a small introduction about the complexity levels, complexity in general, text complexity and readability,
- having short activities with texts that do not fit exactly the descriptors,
- discussing the results of the preparatory task and the descriptions of the levels (i.e., which descriptors go to which level), with focus on the ones reaching low agreement in the preparatory task,
- analysing the focus groups discussion to provide a final version of the complexity levels description.

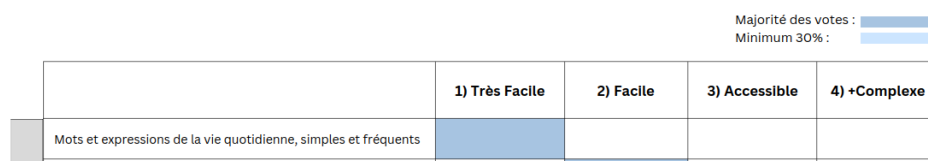
3. After session 1 & the focus group (session 2) (ca. 0,5 hour): the version of the descriptors in each level, resulting from the focus groups of the complexity level descriptions was sent to the participants, with examples of texts for each level, for validation or minor changes. This was done on a shared document, since the preparatory task and focus group discussion already produced a document on which the agreement was high.

4.2.1 French case

Two hybrid sessions (on site + Jitsi meet) were carried out: the first one took place at the Médiathèque Départementale d’Ille et Vilaine (Rennes, FR) on November 16th, from 14:30 to 16:00, and the second one at the Enseignement de Promotion et de Formation Continue (EPFC Bruxelles, BE) on November 23, from 12:30 to 14:00. For each session, there were participants on site with the hostess Alice Pintard, and participants online. 12 participants carried out the preparatory task from which 9 participated in the group discussions, from 7 different structures: Agence Nationale de Lutte Contre l’Illettrisme (ANLCI), Médiathèque Départementale d’Ille et Vilaine, Savoirs pour réussir Paris, EPFC, Welcome Babelkot (Bruxelles), Français pour adultes, and Maison des habitants de Grenoble. Most of the participants are trainers or directors of centres teaching low literacy adults, 2 participants are librarians in charge of the *Facile à Lire* program in France, one is writer of several teaching methods for low literacy adults, and one is the national coordinator for trainers professionalisation at the ANLCI.

During the focus group sessions, Alice Pintard invited all participants to present themselves, then she briefly presented the concept of text readability and the purpose of the focus group in the iRead4Skills project. She showed the results provided by the preparatory task, from which 4 types of results were defined according to the percentage of agreement. These 4 types of results were showed to the participants as follows:

- 1) Minimum 70% agreement on one level: 22 descriptors validated directly, i.e. without discussing them during the focus groups.



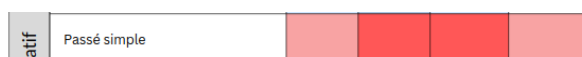
- 2) Between 50% and 70% agreement on one level + minimum 30% agreement on another level: 44 descriptors to discuss during focus group 1 (Nov 16th).



- 3) Equality between 2 levels (50/50 or 40/40): 31 descriptors to discuss during focus group 2 (Nov 23th).



- 4) Answers spread on 4 levels without clear agreement (20/30/30/20): only one descriptor to discuss during both focus groups.

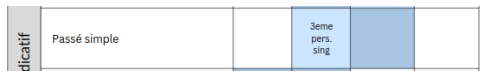


For each problematic descriptor (category 2, 3 and 4), participants were asked if the descriptor was

- a) lacking clarity so needing to be rewritten before choosing a level,
- b) not useful or not adapted to the target audience, so needing to be suppressed,

c) clear and useful but shared by two levels, so maybe needing specific examples in each level.

There was a need for more clarity and less redundancy in the descriptors, which brought the group to merge close descriptors together and differentiate them only by the type of examples, the degree of frequency of the phenomenon or the size of the phenomenon. As a result, the group settled on 79 descriptors, sometimes specific to a level, sometimes shared by two levels with small differences highlighted. Similarly, some descriptors that were shared by two levels but couldn't be rewritten in two different descriptors were differentiated with specific information in each level. For example, the "passé simple" tense on which no clear tendency appeared from the preparatory task was settled as follows:



The focus group sessions allowed participants to agree on 26 descriptors in level "Very Easy", 37 descriptors in level "Easy", 36 descriptors in level "Plain" and 18 descriptors in level "+Complex", with 33 descriptors being shared by several levels. Results of the focus group discussions are detailed in Appendix III where the full table of validated descriptors is shown.

After the focus group sessions, the group worked on a shared document to write a final description of the three levels from the descriptors table. Participants also selected 2 typical examples of text for each level through a survey. The final document is shown in Appendix III.

4.2.2 Portuguese case

The online session (via Zoom) took place on November 20th, 2023, from 15:00 to 17:00. There were 10 trainers¹ (mostly from the Odivelas e Barreiro centers) who carried out the preparatory task, one member of the PNL (Plano Nacional de Leitura), 4 researchers and the project manager from the project iRead4Skills, and the host was Susana Correia. The preparatory task was answered by 11 participants and the results were divided into 5 categories, 4 related to the complexity levels and one exclusively for the answers which had the same number of answers ("Draw"). According to the participants' answers, 12 descriptors were selected as "Very easy", 26 were "Easy", 40 were "Plain", 14 were "+complex" and 9 were "Draw". At this point, 21 descriptors related to the verb tenses were not previously classified in any of the aforementioned categories.

During the Focus Group session, Michell Moutinho started the presentation which had a small introduction and presentation of each participant, an explanation of what *readability* and *text complexity* are, and a reminder of the complexity levels present on the preparatory task. After that, Susana Correia conducted the discussion by presenting the descriptors which were from the category "Draw". Afterwards it was discussed the differences which were found in the classifications made by the project and the answers from the participants. While showing the descriptors which were considered problematic, the host asked the participants to reflect on the previous classification and their choices. According to the discussion, the greatest problem was that the trainers were thinking how their students might have considered the descriptors and examples presented in the preparatory task. Therefore, most of the discussion was based less on their opinions and feelings about the descriptors and more on whether the definitions and the examples suited the target audience with which the trainers normally deal.

After the discussion, the main changes were:

- the validated descriptors were redivided into new categories: "**Matching Descriptors**" (53%), "**Draw**" (9%) and "**Disagreements**" (38%), including the verb tenses.
- Only three descriptors changed to other levels: descriptor **number 42** (*The text has compound words whose meaning cannot easily be understood by combining the meaning of the parts*); descriptor **number 55** (*The sentences in the text frequently include relative sentences related to the subject of the sentence*); and

¹ There were two participants with problems during the meeting. One person had problems with the microphone but participated via chat. The other entered the meeting, stayed for a while, but did not interact during the discussion and only re-entered the session at the end.

descriptor **number 76** (*The text uses less common figurative senses*). The descriptors were in the “Plain” category, but they changed to the “+complex” level.

- **31 descriptors** had their texts changed to be simpler and clearer.
- The descriptors from the “Disagreement” category were maintained in the same classification slot because either the text or the example was more relevant for the participants than the descriptor itself. One participant had problems identifying the “+complex” level in the screen, so this person decided to choose the “Plain” level for some descriptors. Because of this technical problem, some answers had to be reconsidered.

Another reason for maintaining descriptors in the same classification was that the descriptors from the “Draw” and “Disagreement” categories were chosen by the majority of the trainers among either the correct level or the very next level in which the descriptor was supposed to be (e.g., if one descriptor was classified as “Very easy” and about 80% of the participants chose either “Very easy” or “Easy”, it could be considered a valid approximation). Results of the focus group discussions are detailed in Appendix IV where the full table of validated descriptors is shown.

The final document with the descriptions of each level for Portuguese is shown in Appendix IV.

4.2.3 Spanish case

The preparatory work was carried out on different dates (November 16th, 20th and 22nd) due to scheduling issues of different participants or due to time zone differences (some of the participants were attending the working sessions from Argentina). Although the project initially foresaw the participation of about 10 trainers, finally, at the request of the trainers, who suggested extending the invitation to other members of their workplaces, 50 trainers took part in the different sessions.

A summary of the different sessions is given below.

Roser Gauchola, who led the sessions as a member of the UAB research team in the project, presented the aims and purpose of the iRead4Skills project, as well as its chronology and explained the purpose of the working sessions. After presenting the survey to the attendees, he asked them their opinion about the categorisation into levels of difficulty.

About 50% of the participants (22) said that the label “plain” could be misleading in Spanish as it can be taken as a synonym for “easy”. However, after reading the glosses of the different categories, they stated that “claro” was understood as an intermediate level of difficulty.

80% of the participants stated that the items in the trainer survey were sometimes not precise enough and that an example was needed for each item discussed. When the UAB member presented them with some examples for the items, they felt that it was easier to fill in the survey.

As regards the items related to grammar (morphology and syntax) as opposed to those related to referential phenomena (textual typology, concrete/abstract notions, proper/figurative senses), the trainers overwhelmingly (80%) stated that the level of difficulty was often the result of different factors (hypotaxis, chronological sequencing, use of verb modes such as the subjunctive), making it difficult for them to apply an analytical approach here.

Among the audience, discrepancies in assessment were also expressed according to the audience they cater for in their schools. Thus, trainers of newly arrived foreign learners were more likely to assign a higher degree of difficulty to the items presented, while teachers in adult education centres with a native population tended to rate the items presented as less difficult.

Once the doubts had been clarified, the participants declared themselves willing to complete the preparatory tasks. The final document with the descriptions of each level for Spanish is shown in Appendix V.

References

- ACT. (2006). *Reading between the lines: What the ACT reveals about college readiness in reading*. Iowa City.
- ALTE. (2002). *The ALTE CAN DO PROJECT(1992–2002)*. Available at: www.alte.org/can_do/alte_cando.pdf.
- Alves, M. J. and Lameira, S. (coord.) (2021). *Referencial de Competências-chave de Educação e Formação de Adultos – Nível Básico, ANQEP, I.P.*
- Beck, I.L, McKeown, M.G, Omanson R., & Pople, M. (1984). Improving the comprehensibility of stories: The effects of revisions that improve coherence. In *Reading Research Quarterly*. 19. 263-277.
- CCSS/ELA - National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards*. Washington, DC: Authors.
- Chall, J. S., Bissex G., Conard, S., & Harris-Sharples, S. (1999). *Qualitative assessment of text difficulty*. Brookline Publishers. Brookline, MA.
- Council of Europe (2020). *Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume*. Council of Europe Publishing, Strasbourg. Available at: www.coe.int/lang-cefr.
- Cunningham, J. W., & Mesmer, H. A. (2014). Quantitative measurement of text difficulty: What's the use? In *Elementary School Journal*. 115, 255-269.
- Dahl, O. (2007) Definitions of complexity. In *Proceedings of the Colloquium on complexity, accuracy and fluency in second language use, learning & teaching*. 41-46.
- Dale, E. & Chall, J. (1948). A Formula for Predicting Readability. In *Educational Research Bulletin*. 27, 28. 11-20, 37-54.
- Douglas, Y., & Miller, S. (2016). Syntactic and Lexical Complexity of Reading Correlates with Complexity of Writing in Adults. In *International Journal of Business Administration*, 7(4).
- Fang, Z., & Pace, B. G. (2013). Teaching With Challenging Texts in the Disciplines. In *Journal of Adolescent & Adult Literacy*. 57(2). 104–108.
- Flesch, R. (1948). A new readability yardstick. In *Journal of Applied Psychology*. 32. 221-233.
- Flor, M., Klebanov, B., Sheehan, K. (2013). Lexical tightness and text complexity. In *Proceedings of the Workshop on Natural Language Processing for Improving Textual Accessibility*. 29–38. Atlanta, GA.
- Glenberg, A.M., Meyer, M., & Lindem, K. (1987). Mental models contribute to foregrounding during text comprehension. In *Journal of Memory and Language*, 26.
- Goldman, S.R., & Lee, C.D. (2014). Text complexity. State of the art and the conundrums it raises. In *The Elementary School Journal*, 115, 290–300.
- Gunning, R. (1952). *The technique of clear writing*. McGraw-Hill. New-York.
- Harley, T. A. (2013). *Psychology of Language: From Data to Theory*. Taylor and Francis.
- Hiebert, E.H. (2011). Using multiple sources of information in establishing text complexity. In *Reading Research Report*. 11.
- Hiebert, E.H., & Pearson, P.D. (2014). Understanding text complexity. In *The Elementary School Journal*, 115, 153–160.
- James, C. T. (1975). The role of semantic information in lexical decisions. In *Journal of Experimental Psychology: Human Perception and Performance*. 1(2). 130–136.

- Kincaid, J. P., Fishburne, R. P. J., Rogers, R. L. and Chissom, B. S. (1975). Derivation of new readability formulas Automated Readability Index, Fog Count and Flesch Reading Ease Formula for Navy enlisted personnel. Naval Technical Training, U. S. Naval Air Station, Memphis, TN.
- Kusters, W. (2003) Linguistic complexity: the influence of social change on verbal inflection. LOT, Utrecht.
- Lorge, I. (1948). The Lorge and Flesch Readability Formulae: A Correction. In *School and Society*. 67. 141-142.
- Mclaughlin, G. H. (1969). SMOG Grading — a New Readability Formula. In *Journal of Reading* 128. 639-646.
- McNamara, D. S., Louwrese, M. M., McCarthy, P. M., & Graesser, A. C. (2010). Coh-Metrix: Capturing Linguistic Features of Cohesion. In *Discourse Processes*. 47(4), 292–330.
- OECD (2013a). Technical report of the survey of adult skills (PIAAC). OECD Publishing. Paris.
- OECD. (2013b). The Survey of Adult Skills Reader's Companion. OECD Publishing.
- Pearson, P. D., & Hiebert, E. H. (2014). The State of the Field. In *The Elementary School Journal*. 115(2). 161–183.
- Rabin, A. T. (1988). Determining Difficulty Levels of Text Written in Languages Other than English. In *Readability: It's Past, Present, & Future*.
- Spache, G. (1953). A new readability formula for primary-grade reading materials. In *The Elementary School Journal* 537.410–413.
- UNESCO. (2020, June 22). Literacy. Uis.unesco.org. <https://uis.unesco.org/node/3079547>

APPENDIX I - Set of complexity descriptors to be validated

(with PT examples in language-dependent cases)

dimension	<i>descriptor (complexity level)</i>
text size	1. The text has up to 50 words. (<i>very easy</i>) 2. The text has up to 150 words. (<i>easy</i>) 3. The text has up to 250 words. (<i>plain</i>) 4. The text has up to 500 words. (<i>+complex</i>)
purpose	5. The text serves the purpose of performing common/familiar tasks (e.g., instruction to buy bus tickets, usage instructions on product labels) (<i>very easy</i>) 6. The text serves the purpose of presenting a new topic (e.g., initiatory text on the importance of recycling). (<i>very easy</i>) 7. The text is interesting for the reader to inform him/herself in moments of leisure (e.g., sports chronicle, daily news). (<i>easy</i>) 8. The text to learn/perform new tasks (e.g., instructions on how to use the company e-mail). (<i>easy</i>) 9. The text can be about varied topics of interest to the reader for information or leisure. (e.g., nature magazine article). (<i>plain</i>)
communication context	10. The text is used in everyday communications (e.g., transport schedules, lists, menus, general instructions, item price information, personal communication, etc.). (<i>very easy</i>) 11. The text is used in work contexts or media contexts (e.g., work communication or instructions, news, ads, sports). (<i>easy</i>) 12. The text is used in leisure contexts (e.g., literature and fiction, reportage, opinion articles). (<i>plain</i>) 13. The text is used artistic contexts (e.g., literature, poetry). (<i>plain</i>) 14. The text is used in specialized contexts (e.g., formal professional communication, online forums). (<i>plain</i>) 15. The text is used in academic or scientific contexts (e.g., essays, reports, academic papers) (<i>+complex</i>)
conceptual diversity	16. The text refers to concepts of everyday activities. (e.g., food, housekeeping, clothes). (<i>very easy</i>) 17. The text refers to concepts related to domain-specific uses (e.g., work, sports, hobbies). (<i>easy</i>) 18. The text refers to concepts that require readers to puzzle together pieces of information. (e.g., stars, volcanos). (<i>plain</i>) 19. The text refers to multifaceted or new concepts. (e.g., black matter, brain development, statistics). (<i>+complex</i>)
concreteness	20. The text is mostly composed of words that refer to concrete objects (e.g., <i>chair, apple, hammer, computer</i>). (<i>very easy</i>) 21. The text has a minor percentage of words that refer to common abstract ideas (e.g., <i>happiness, imagination, love</i>). (<i>easy</i>) 22. The text uses words that refer to abstract concepts and ideas (e.g., <i>aging, reasoning, loyalty</i>). (<i>plain</i>) 23. The text uses words that refer to multifaceted and/or domain-specific abstract concepts and ideas (e.g., <i>poetry, democracy, nuclear fusion</i>). (<i>+complex</i>)

vocabulary/lexicon subset	<p>24. The text does not use abbreviations (e.g., <i>g.</i>, <i>km</i>, <i>c.</i>, <i>V. Exa.</i>), except the frequent forms of treatment (e.g., <i>Sr.</i>, <i>Dra.</i>). (very easy)</p> <p>25. The text uses common abbreviations after explanation in the text (e.g., <i>l – litre</i>, <i>m – meter</i>). (easy)</p> <p>26. The text uses common abbreviations (e.g., <i>m</i>, <i>km</i>, <i>g.</i>, <i>l</i>, <i>Exmo.</i>, ...). (clear)</p> <p>27. The text does not use acronyms or initialisms. (very easy)</p> <p>28. The text uses acronyms and initialisms that are commonly known, i.e., that can be understood in isolation (e.g., <i>CP</i>, <i>IRS</i>, <i>IVA</i>, <i>PCP</i>, <i>PS</i> (train company, taxes, political parties...)). (easy)</p> <p>29. The text uses acronyms and initialisms after describing them (e.g., <i>Organização das Nações Unidas (ONU)</i>, <i>public institute (PI)</i>). (plain)</p> <p>30. The text uses all types of acronyms (e.g., <i>DNA</i>, <i>UNESCO</i>, <i>UNICEF</i>, <i>EU</i>) (+complex)</p> <p>31. The text has words and simple expressions of everyday matters, using frequent and commonly known nouns (e.g., <i>potatos</i>, <i>olive oil</i>, <i>vegetables</i>). (very easy)</p> <p>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 (e.g., <i>Monday</i>, <i>course</i>, <i>school year</i>). (easy)</p> <p>33. The text has words and expressions less known that refer to specific domains (e.g., <i>heart disease</i>, <i>artery</i>, <i>treatment</i>). (plain)</p> <p>34. The text has words and expressions that refer to specific domains, specialized jargon or archaic terms (e.g., <i>enfarte do miocárdio</i>, <i>artéria coronária</i>, <i>dissecção espontânea</i>, <i>genoma</i>) (+complex)</p> <p>35. The text has no foreign words, except those very common and widely known (e.g., <i>Internet</i>, <i>donut</i>) (very easy)</p> <p>36. The text has common foreign words (e.g., <i>site</i>, <i>stress</i>), but these are not the majority in the text. (easy)</p> <p>37. The text uses common foreign words frequently (e.g., <i>link</i>, <i>hobby</i>, <i>marketing</i>, <i>mail</i>, <i>gay</i>). (plain)</p> <p>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 (e.g., <i>interno</i>, <i>representar</i>, <i>forma</i>, <i>paciente</i>). (plain)</p> <p>39. The text has highly polysemic words (i.e., words that can have more than one meaning and can be ambiguous) (e.g., <i>base</i>, <i>experiência</i>, <i>leve</i>, <i>seca</i>, <i>aplicar</i>). (+complex)</p>
morphology	<p>40. The text has common compound words (e.g., <i>feijão verde</i>, <i>couve-flor</i>) or transparent compound words (words or expressions whose meaning can be easily understood by combining the meaning of the parts) (e.g., <i>tira-caricas</i>) (very easy)</p> <p>41. The text has common compound words² whose meaning can be more or less understood by combining the meaning of the parts (e.g., <i>óculos de sol</i>, <i>mal-educado</i>, <i>meia-idade</i>, <i>maus-tratos</i>). (easy)</p> <p>42. The text has compound words whose meaning cannot easily be understood by combining the meaning of the parts (e.g., <i>meio-termo</i>, <i>porta-voz</i>, <i>obra prima</i>). (plain)</p> <p>43. The text has diverse compound words whose meaning cannot easily be understood by combining the meaning of the parts (e.g., <i>pós-graduação</i>, <i>mais-valia</i>, <i>pé de meia</i>, <i>cavalo de batalha</i>) (+complex)</p> <p>44. The text has derived words, with common and familiar prefixes and suffixes (e.g., <i>infeliz</i>, <i>calmamente</i>, <i>organizar</i>, <i>trabalhador</i>, <i>casinha</i>), but these are not the majority in the text. (very easy)</p> <p>45. The text has derived words with frequent prefixes and suffixes (e.g., <i>precozinhado</i>, <i>desorganizar</i>, <i>embalagem</i>, <i>organização</i>, <i>jardineiro</i>, <i>finíssimo</i>), but these are not the majority in the text. (easy)</p> <p>46. The text has derived words, with varied suffixes and prefixes (e.g., <i>nervosismo</i>, <i>lavável</i>, <i>curiosidade</i>, <i>altitude</i>, <i>superpoder</i>, <i>preconceituoso</i>) (plain)</p> <p>47. The text has derived words³ with less common suffixes and prefixes (e.g., <i>megatonelada</i>, <i>tendinite</i>, <i>geopolítico</i>) (+complex)</p>

² 'compound word' is here used as an umbrella term for both compound words at the morphological level or others multi-element words formed by lexicalization processes. It is intended to be more accessible to non-experts and not to reflect specific theoretical frameworks.

³ 'derived word' is here used as an umbrella term for words formed by any of the available affixes, including classical ones such as 'geo' or 'mega', that can be considered stems. It is intended to be more accessible to non-experts and not to reflect specific theoretical frameworks.

48. Please put the structures in the level(s) you think they occur the most/they are more typical.		very	easy	plain	more
		easy			complex
morphosyntax	verbs in the Infinitive (e.g., <i>gostar, comer, fugir</i>)				
	verbs in the Indicative Present (e.g., <i>gosto, comem, foge</i>)				
	verbs in the Indicative Pretérito Imperfeito (e.g., <i>gostavam, comia, fugias</i>)				
	verbs in the Indicative Pretérito Perfeito (e.g., <i>gostou, comi, fugiram</i>)				
	verbs in the Indicative Pretérito Mais-que-perfeito (e.g., <i>gostara, comeras, fugira</i>)				
	verbs in the Future Indicative simple (e.g., <i>esquecerá, gostará</i>)				
	verbs in the Gerund (e.g., <i>gostando, comendo, fugindo</i>)				
	verbs in the Conditional (e.g., <i>gostaria, comerias, fugiriam</i>)				
	verbs in the Subjunctive Present (e.g., <i>gosta, coma, fujam</i>)				
	verbs in the Subjunctive Pretérito Imperfeito (e.g., <i>gostasses, comessem, fugisse</i>)				
	verbs in the Subjunctive Future (e.g., <i>se eu gostar, se tu comeres, se eles fugirem</i>)				
	verbs in the Pretérito Perfeito Composto Indicative (e.g., <i>tenho gostado, tem comido, têm fugido</i>)				
	verbs in the Pretérito Mais-que-perfeito Composto Indicative (e.g., <i>tinhas gostado, tinham comido, tinha fugido</i>)				
	verbs in the Futuro Composto Indicative (e.g., <i>terei gostado, terão comido, terão fugido</i>)				
	verbs in the Pretérito Perfeito Composto Subjunctive (e.g., <i>tenha gostado, tenha comido, tenham fugido</i>)				
	verbs in the Pretérito Mais-que-perfeito Composto Subjunctive (e.g., <i>tivesses gostado, tivessem comido, tivesse fugido</i>)				
	verbs in the Futuro Composto Subjunctive (e.g., <i>tiverem gostado, tiver comido, tiveres fugido</i>)				
	auxiliar verbs <i>poder, ir, estar</i> in the Indicative tenses (e.g., <i>posso gostar, vai comer, estão a fugir</i>)				
	expressions with the verbs <i>permanecer, querer, poder, conseguir</i> in simple tenses (e.g., <i>permaneça calmo, queria almoçar, consigo ir</i>)				
	expressions with the verbs <i>estar, ser, permanecer, querer, poder, dever</i> in compound tenses (e.g., <i>tenham estado ligados; foram feitas; temos de permanecer calmos; devíamos ir</i>)				

syntax	<p>49. The majority of the sentences in the text are short and direct or combined through frequent and familiar connectors (e.g., <i>A criança fugiu ontem. O pai trabalha na fábrica <u>mas</u> a mãe fica em casa.</i>) (very easy)</p> <p>50. The text has sentences that are combined through coordination and subordination connectors (e.g., <i>Os músicos foram surpreendidos pela polícia <u>e</u> estão concentrados em frente ao centro comercial. <u>Se</u> ainda não pagou a conta, pode pagar aqui.</i>) (easy)</p> <p>51. The sentences in the text are frequently combined through coordination and subordination connectors, and sometimes with less common connectors (e.g., <i>Apesar da confusão provocada pelos animais à solta, a situação foi controlada <u>sem</u> os animais serem atropelados.</i>) (plain)</p> <p>52. The sentences in the text are long and composed of several other sentences combined through coordination and subordination connectors (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>) (+complex)</p> <p>53. The text has relative sentences related to the subject of the sentence (e.g., <i>O senhor <u>que vive na casa ao lado</u> ouviu o barulho.</i>), but these are not the majority of the text. (very easy)</p> <p>54. The text has relative sentences related to the subject of the sentence (e.g., <i>O senhor <u>que vive na casa ao lado</u> ouviu o barulho.</i>). (easy)</p> <p>55. The sentences in the text frequently include relative sentences related to the subject of the sentence (e.g., <i>Francisca, <u>a quem os amigos tratam por «Kika»</u>, começou a surfar muito pequena.</i>). (plain)</p> <p>56. The sentences in the text frequently include relative sentences related to the subject or the object of the sentence (e.g., <i>O rapaz <u>que abraçou a mãe</u> dirigiu-se à entrada; O rapaz <u>que a mãe abraçou</u> dirigiu-se à entrada.</i>). (+complex)</p> <p>57. The majority of the sentences in the text are in the active voice (e.g., <i>O cão comeu o bolo</i> (ativa) vs. <i>O bolo foi comido pelo cão.</i> (passiva)) (very easy)</p> <p>58. The text has sentences in the passive voice (e.g., <i>A visita foi paga pelo patrão.</i>), but these are not the majority of the text. (easy)</p> <p>59. The text has sentences in the passive voice (e.g., <i>O almoço foi pago pelo empresa.</i>), including common impersonal and passive constructions with <i>-se</i> (e.g., <i>O trabalho faz-se bem.</i>) (plain)</p> <p>60. The text has simple sentences and in the basic linear order (subject-verb-complements) (e.g., <i>Eu ajudei os meus filhos a comprar casa.</i>). (very easy)</p> <p>61. The text has simple sentences not in the basic linear order (subject-verb-complements) (e.g., <i>À escola, chegaram os autocarros da visita de estudo.</i>), but these are not the majority of the text (easy)</p> <p>62. The sentences of the text are not necessarily in the basic linear order (subject-verb-complements) (e.g., <i>De ar puro, todos precisamos.</i>). (plain)</p> <p>63. The sentences of the text have diverse linear orders (subject-verb-complements; verb-complements-subject, etc.) (e.g., <i>Até final de março, porque antes seria impossível encontrar todos os ingredientes, haverá na ementa vários pratos de caça.</i>) (+complex)</p>
textual cohesion: pronouns, anaphora and ellipsis	<p>64. The text has few or none omitted elements (e.g., <i>O acidente causou dois feridos leves. Os bombeiros levaram os feridos para o hospital.</i>). (very easy)</p> <p>65. There are repeated elements instead of pronouns (e.g., 'João doesn't like <u>Maria</u> because <u>Maria</u> is not nice'). (very easy)</p> <p>66. The text has some omitted elements (e.g., <i>[] Reparo desde pequena nos pássaros. vs. <u>Eu</u> reparo desde pequena nos pássaros.</i>). (easy)</p> <p>67. The text uses pronouns (e.g., 'João doesn't like <u>Maria</u> because <u>she</u> is not nice'). (easy)</p> <p>68. The text has omitted elements (e.g., <i>[] Mostrámos em estudos anteriores que a gordura solidifica quando [] arrefece. vs. <u>Nós</u> mostrámos em estudos anteriores que a gordura solidifica quando <u>ela</u> arrefece.</i>) (plain)</p> <p>69. The text uses pronouns and related words to indicate what is being referred (e.g., 'In the book, we give a complete definition of blizzard and explain how <u>it</u> is formed. We also indicate where the <u>phenomenon</u> can be seen.') (plain)</p>

textual cohesion: temporal references	<p>70. The text has simple and linear temporal sequence or location, expressed by common temporal adverbs or connectors such as <i>hoje, amanhã, antes de, depois de, quando, enquanto</i> (e.g., <i>Os bombeiros chegaram depois da polícia.</i>)(very easy)</p> <p>71. The text has a linear temporal sequence or location, expressed by temporal adverbs or connectors such as <i>após, assim que, antes de</i> and by verb tenses (e.g., <i>Ele está a caminho e a festa já começou. A banda começou a tocar antes de todos saírem.</i>) (easy)</p> <p>72. The text can have a non-linear temporal sequence or location (e.g., <i>Após analisarem as respostas dos doentes que tinham sofrido um ataque cardíaco, os médicos confirmaram que a gordura é o alimento mais perigoso.</i>). (plain)</p> <p>73. The text has diverse and non-linear temporal sequences or locations (e.g., <i>Será preciso fazer alguma coisa aos cadáveres e, por isso, é bom que haja gente que se debruce sobre o tema. Eu seria incapaz, porque nem sequer consigo ir a funerais, como já fui, onde repousam caixões abertos com os indivíduos em descanso eterno.</i>) (+complex)</p>
style	<p>74. The text does not use figurative senses (e.g., <i>Este verão os preços subiram vs. Este verão os preços dispararam</i> (fig.)) (very easy)</p> <p>75. The text uses common figurative senses (e.g., <i>Este verão os preços dispararam.</i>) (easy)</p> <p>76. The text uses less common figurative senses (e.g., <i>El Barça barrió al Madrid; O Estoril esmagou o Santa Clara.</i>) (plain)</p> <p>77. The text uses uncommon figurative senses (e.g., <i>Os primeiros raios de sol aconchegaram-me as faces.</i>) (+complex)</p> <p>78. The text uses only comparison, personification, hyperbole (e.g., <i>Gostar de alguém é como fazer prédios; As doenças escondem-se em todo o lado; Eu estava gelada; A tua mala pesa uma tonelada.</i>)(very easy)</p> <p>79. The text uses metonymy (e.g., <i>tomar se una copa; beber um copo; prendre un verre</i>), familiar metaphors (e.g., <i>O meu caminho é outro. Não quero estudar mais.</i> (Mine is another path, I don't want to study more.) (easy)</p> <p>80. The text uses non common metaphors (e.g., <i>Marcelo e Costa, um casamento de convergência. Os dois coabitam desde 9 de março de 2016, quando Marcelo assumiu funções.</i>) (plain)</p> <p>81. The text uses diverse figures of speech (e.g., <i>O cão da fábrica de curtumes acendia latidos fosforescentes nas noites de julho, quando o pólem da acácia me cobria as pálpebras, eu, morto de amores pela mulher de Sandokan, descobria-me unicórnio trancado na retrete da escola...</i>) (+complex)</p>

APPENDIX II - Preparatory tasks

Présentation

Bienvenue dans le groupe de travail du projet [iRead4Skills](#)

Ce groupe de travail a pour but de définir et caractériser trois niveaux de complexité textuelle adaptés au public adulte en cours d'alphabétisation ou en situation d'illettrisme. Ces niveaux serviront ensuite à classer des textes sur une échelle de difficulté et à analyser puis modéliser la difficulté textuelle pour notre public cible.

L'équipe du projet iRead4Skills vous remercie d'avance pour votre participation !

Titre du projet : iRead4Skills - Intelligent Reading Improvement System for Fundamental and Transversal Skills Development

Ce projet est financé par la Commission européenne, HORIZON-CL2-2022-TRANSFORMATIONS-01-07 Program (ref. 101094837)

Responsable pour la recherche en français : Thomas François (Université Catholique de Louvain, Belgique)

Contacts : thomas.francois@uclouvain.be, alice.pintard@uclouvain.be

Informations nécessaires à la préparation des groupes de discussion

Prénom

Nom

Structure dans laquelle vous travaillez avec les adultes

Groupe de discussion choisi

16 novembre de 14h30 à 16h, format hybride (Rennes / en ligne)

23 novembre de 12h30 à 14h, format hybride (Bruxelles / en ligne)

Tâche préparatoire aux groupes de discussion

L'objectif de cette tâche est de vous faire prendre connaissance des descripteurs de niveaux proposés par les équipes du projet IRead4Skills. Votre expérience du public et du terrain est capitale pour ce travail et nous attirons votre attention sur le fait que ces descripteurs ne sont qu'une première étape dans la définition de niveaux de complexité textuelle. Les groupes de discussion puis un travail de rédaction de la part des équipes de recherche viendront compléter cette tâche.

Nous proposons trois niveaux de complexité textuelle pour ce projet :

Très facile : Textes entièrement ou presque entièrement compris par tous les adultes, y compris les personnes très peu scolarisées (c'est-à-dire qui n'ont pas terminé l'école primaire) et qui n'ont pratiquement aucune expérience de la lecture.

Facile : Textes entièrement ou presque entièrement compris par des personnes peu scolarisées (c'est-à-dire qui ont terminé l'école primaire mais n'ont pas dépassé la troisième secondaire (Belgique) ou la troisième (France)) et qui ont une faible expérience de la lecture.

Accessible : Textes compris dès la première lecture par des personnes ayant suivi la troisième secondaire (Belgique) ou la troisième (France) et ayant une expérience moyenne de la lecture.

Tous les textes sortant de ce cadre sont au-delà de la limite haute de notre échelle de difficulté. Nous proposons le niveau **+Complexe** pour faire référence à ces textes qui ne constituent pas le cœur du projet.

Les questions suivantes vous invitent à classer les descripteurs dans les niveaux de complexité. Chaque descripteur est associé à une zone de texte vous permettant d'écrire un commentaire si besoin (ex: descripteur trop flou, inutile, idée de reformulation du descripteur...).

Si vous ne pouvez faire un choix entre deux niveaux adjacents, vous avez la possibilité de cocher plusieurs cases mais merci de ne réserver cette option qu'aux descripteurs qui vous semblent vraiment à cheval entre deux niveaux.

Questions

Placez les propositions suivantes dans le(s) niveau(x) qui vous semble(nt) approprié(s).

1 - Très Facile

2 - Facile

3 - Accessible

+Complexe

Le texte est utilisé dans des contextes académiques ou scientifiques (ex: essais, rapports, documents académiques).

Le texte fait référence à des concepts multiformes ou nouveaux. (ex: la matière noire, le développement du cerveau, les statistiques).

Le texte compte jusqu'à 250 mots.

Le texte sert à accomplir des tâches habituelles/quotidiennes (ex: instructions pour acheter des tickets de bus, étiquettes de produits).

Le texte fait référence à des concepts liés à un domaine spécifique (ex: travail, sport, loisirs).

1 - Très Facile

2 - Facile

3 - Accessible

+Complexe

Le texte utilise des mots qui se réfèrent à des idées et des concepts abstraits à multiples facettes et/ou spécifiques à un domaine (ex: la poésie, la démocratie, la fusion nucléaire).

Le texte peut porter sur des sujets variés qui intéressent le lecteur à titre d'information ou pour une lecture plaisir. (ex: article de magazine sur la nature).

Le texte compte jusqu'à 50 mots.

Le texte est principalement composé de mots qui font référence à des objets concrets (ex: chaise, pomme, ordinateur).

Le texte fait référence à des concepts du quotidien. (ex: nourriture, ménage, vêtements).

Le texte est utilisé dans des contextes spécialisés (ex: communication professionnelle formelle, forums en ligne).

Le texte est intéressant pour le lecteur et lui sert à s'informer pendant son temps libre (ex: chronique sportive, nouvelles quotidiennes).

1 - Très Facile

2 - Facile

3 - Accessible

+Complexe

Le texte sert à apprendre/effectuer de nouvelles tâches (ex: instructions sur l'utilisation du courrier électronique de l'entreprise).

Le texte est utilisé dans des contextes artistiques (ex: littérature, poésie).

Le texte fait référence à des concepts qui exigent du lecteur qu'il rassemble différents éléments d'information (ex: les étoiles, les volcans).

Le texte contient un faible pourcentage de mots faisant référence à des idées abstraites courantes (ex: le bonheur, l'imagination, l'amour).

Le texte est utilisé dans les communications quotidiennes (ex: horaires de transport, listes, menus, informations sur le prix des articles, communications personnelles, etc.).

Le texte utilise des mots qui font référence à des idées et des concepts abstraits (ex: vieillissement, raisonnement, loyauté).

Le texte compte jusqu'à 150 mots.

1 - Très Facile

2 - Facile

3 - Accessible

+Complexe

Le texte sert à présenter un nouveau sujet (ex: texte d'introduction sur l'importance du recyclage).

Le texte compte jusqu'à 500 mots.

Le texte est utilisé dans des contextes de loisirs (ex: fiction, reportages, articles d'opinion).

Le texte est utilisé dans des contextes professionnels ou médiatiques (ex: communications professionnelles, nouvelles, publicités, sports).

Avez-vous des commentaires sur cette première partie relative à la taille du texte, à l'objectif communicatif, au domaine de communication et à la diversité conceptuelle ? (ex : descripteurs à ajouter, commentaires généraux)

Selon vous, quel(s) niveau(x) de complexité ces propositions décrivent-elles ?

1 - Très Facile

2 - Facile

3 - Accessible

+Complexe

Le texte contient des expressions et mots composés fréquents (ex: grand-mère, peut-être, pomme de terre).

Le texte contient des mots et des expressions moins connus qui se réfèrent à des domaines spécifiques (ex: maladie cardiaque, traitement).

1 - Très Facile

2 - Facile

3 - Accessible

+Complexe

Le texte ne contient pas de mots étrangers, à l'exception de ceux qui sont très courants et largement connus (ex: Internet, weekend).

Le texte utilise des abréviations courantes (ex: km, g., rdv).

Le texte comporte des mots dérivés avec des préfixes et des suffixes courants (ex: dé-, in-, -age, -ité).

Le texte contient peu de mots étrangers et ces mots sont courants (ex: email, parking).

Le texte utilise des acronymes et des sigles courants, c'est-à-dire qui peuvent être compris isolément (ex: SNCF, EDF, CDD, PDG, SIDA, ...).

Le texte contient peu de mots dérivés, avec des préfixes et suffixes très fréquents (ex: re-, -ment, -tion).

Le texte contient des mots et expressions fréquents qui peuvent se référer à des sujets quotidiens mais aussi à d'autres domaines spécifiques ou au contexte professionnel (ex: lundi, réunion, sortie).

1 - Très Facile

2 - Facile

3 - Accessible

+Complexe

Le texte n'utilise pas d'abréviations, à l'exception des formes très fréquentes (ex: Mme, Prof.).

Le texte contient des mots et des expressions simples de la vie quotidienne, avec des noms fréquents et connus (ex: table, train, pantalon).

Le texte utilise des abréviations après explication dans le texte (ex: millilitre (mL), mètre (m)).

Le texte n'utilise pas d'acronymes ou d'initiales.

Le texte contient des mots et des expressions qui font référence à des domaines spécifiques, à un jargon spécialisé ou à des termes archaïques (ex: génome, présentiel, moult)

Le texte contient des mots dérivés, avec des affixes variés (ex: auto-, anti-, co-, ex-, inter-, mini-, pré-, sou(s)-, sur-, -able, -aire, -esse, -eur, ien, -iste).

Le texte utilise tous les types d'acronymes (ex: ADN, UNESCO, UNICEF, UE)

1 - Très Facile

2 - Facile

3 - Accessible

+Complexe

Le texte utilise des acronymes et des sigles après les avoir décrits (ex: Organisation des Nations Unies (ONU), Parti socialiste (PS)).

Le texte contient des mots composés peu fréquents et dont le sens ne peut pas être facilement compris en combinant le sens des parties (ex: ça et là, année-lumière, à tout bout de champ, tendre l'oreille).

Le texte comporte des mots dérivés avec des affixes peu courants (ex: mega-, geo-, -(i)fier).

Le texte contient des expressions et mots composés dont le sens ne peut pas être facilement compris en combinant le sens des parties (ex: troisième âge, rez-de-chaussée).

Le texte contient des expressions et mots composés dont le sens peut être plus ou moins compris en combinant le sens des parties (ex: non-fumeur, sac à dos).

Le texte utilise fréquemment des mots étrangers (ex: marketing, gay, show, feedback, low-cost).

Avez-vous des commentaires sur cette deuxième partie, relative au lexique/vocabulaire et à la morphologie ? (ex: descripteurs à ajouter, commentaires

généraux)

Placez ces structures dans le(s) niveau(x) où vous pensez qu'on les trouve le plus souvent ou qu'elles sont les plus typiques.

	1 - Très Facile	2 - Facile	3 - Accessible	4 - +Complexe
Verbes à l'impératif passé (ex: aie marché, ayez fini)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Verbes au passé simple (ex: je marchai, tu finis)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Verbes au participe passé (ex: ayant marché, ayant fini)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Verbes au passé antérieur (ex: j'eus marché, tu eus fini)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Verbes à l'imparfait de l'indicatif (ex: je marchais, tu finissais)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Verbes au passé composé (ex: j'ai marché, tu as fini)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Verbes à l'impératif présent (ex: marche, finissez)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Verbes au futur antérieur (ex: j'aurai marché, tu aurais fini)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Verbes au conditionnel passé (ex: j'aurais marché, tu aurais fini)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Verbes au conditionnel présent (ex: je marcherais, tu finirais)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Verbes au participe présent (ex: marchant, finissant)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Verbes à l'imparfait du subjonctif (ex: (que) je marchasse, (que) tu finisses)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

	1 - Très Facile	2 - Facile	3 - Accessible	4 - +Complexe
Verbes à l'infinitif présent (ex: marcher, finir)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Verbes au gérondif passé (ex: en ayant marché, en ayant fini)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Verbes au futur proche, passé récent et présent continu : "aller", "venir de", "être en train de" + infinitif (ex: je vais marcher, tu viens de finir)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Verbes au présent du subjonctif (ex: (que) je marche, (que) tu finisses)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Verbes au gérondif présent (ex: en marchant, en finissant)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Verbes au passé du subjonctif (ex: (que) j'aie marché, (que) tu aies fini)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Verbes au futur simple (ex: je marcherai, tu finiras)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Verbes au plus que parfait (ex: j'avais marché, tu avais fini)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Verbes à l'infinitif passé (ex: avoir marché, avoir fini)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Verbes au présent de l'indicatif (ex: je marche, tu finis)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Avez-vous des commentaires sur cette troisième partie, relative aux temps et modes des verbes ? (ex: descripteurs à ajouter, commentaires généraux)

Selon vous, quel(s) niveau(x) de complexité ces propositions décrivent-elles ?

1 - Très Facile

2 - Facile

3 - Accessible

+Complexe

Le texte comporte très peu ou pas d'ellipses (ex: "Jean préfère la voile et sa sœur préfère le kayak." vs "Jean préfère la voile; sa sœur, le kayak.").

Le texte utilise des sens figurés courants (ex: "La nuit tombe tôt en hiver.")

Le texte peut comporter des comparaisons, personnifications et hyperboles (ex: "Il est grand comme une montagne", "Les maladies se cachent partout", "Ta valise pèse une tonne").

Les phrases du texte sont longues et composées de plusieurs propositions combinées par des connecteurs de coordination et de subordination peu fréquents (ex: Étant donné qu'une grande partie de ce livre traite de la gestion des aspects difficiles liés au télétravail, je tiens d'abord à préciser que ce mode de travail peut être fantastique : c'est pourquoi tant de gens le considèrent comme une option).

Le texte contient des phrases à la voix passive et des structures passives avec "se faire" (ex: Je me suis fait couper les cheveux hier.)

1 - Très Facile

2 - Facile

3 - Accessible

+Complexe

Le texte comporte des éléments répétés à la place de pronoms (ex: "Je ne comprends pas Maria parce que Maria parle trop vite").

Le texte présente une séquence temporelle linéaire, exprimée par des adverbes, des connecteurs et les temps verbaux (ex : "Piero était parti quand la police est arrivée.").

Les phrases du texte comportent fréquemment des propositions subordonnées introduites par une grande variété de connecteurs.

Le texte contient des phrases à la voix passive à l'indicatif.

Le texte contient des phrases simples qui ne sont pas dans l'ordre de base sujet-verbe-complément (ex: Très lentement, il prend son sac.).

Le texte contient des phrases simples et composées avec une variété d'ordres syntaxiques.

1 - Très Facile

2 - Facile

3 - Accessible

+Complexe

Le texte comporte des propositions combinées à l'aide de connecteurs de coordination et de subordination (ex : Ma fille arrive demain donc elle ne verra pas le spectacle ce soir. Si vous n'avez pas encore payé votre facture, vous pouvez le faire ici.)

Le texte contient très peu ou pas de phrases à la voix passive.

Le texte comporte quelques ellipses (ex: "Jean préfère la voile; sa sœur, le kayak.").

Le texte est principalement composé de phrases simples dans l'ordre de base sujet-verbe-complément (ex: Nadia aime le sport.).

Le texte contient peu de propositions subordonnées mais elles peuvent être introduites par des connecteurs variés : relatives (qui, que, dont), temporelles (quand, lorsque, avant que, après que), causatives (comme, parce que, à cause de, alors, pour) et conditionnelles (si).

Les ellipses sont fréquentes (ex: "Jean préfère la voile; sa sœur, le kayak.").

1 - Très Facile

2 - Facile

3 - Accessible

+Complexe

Le texte n'utilise pas de mots au sens figuré (ex: "Cet été, les prix ont augmenté." vs "Cet été, les prix ont grimpé.")

Le texte présente un cadre spatio-temporel simple et linéaire, la cohérence temporelle est donnée au moyen d'adverbes ou de connecteurs courants (ex: aujourd'hui, demain, avant, après...) et non par les temps verbaux. (ex : "Piero est parti avant l'arrivée de la police" vs "La police est arrivée et Piero était parti").

Les propositions du texte sont fréquemment combinées à l'aide de connecteurs de coordination et de subordination, parfois peu courants (ex: Malgré la confusion causée par les animaux en liberté, la situation a été maîtrisée.)

La majorité des phrases/propositions sont courtes et directes ou combinées à l'aide de connecteurs fréquents et familiers (ex: Il est parti hier. J'aime les fruits mais ils sont chers.)

Le texte utilise tout type de figures de style.

Le texte utilise des sens figurés peu courants (ex: "Un manteau de neige recouvre la plaine".)

1 - Très Facile

2 - Facile

3 - Accessible

+Complexe

Le texte comporte des anaphores pronominales (ex: "Je ne comprends pas Maria parce qu'elle parle trop vite").

Le texte peut contenir des propositions subordonnées introduites par des connecteurs fréquents : relatives (qui, que), temporelles (quand, lorsque) et causatives (comme, parce que).

Le texte comporte des métaphores non courantes (ex: Parti socialiste et Modem: un mariage de convenance).

Le texte peut avoir une séquence temporelle ou spatiale non linéaire (ex: 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).

Le texte présente des séquences temporelles ou spatiales non linéaires.

1 - Très Facile

2 - Facile

3 - Accessible

+Complexe

Le texte utilise des pronoms et des mots apparentés pour indiquer ce à quoi il est fait référence (ex: "Dans cet article, nous donnons une définition complète du blizzard et expliquons comment il se forme. Nous indiquons également les endroits où l'on peut observer le phénomène".)

Le texte peut comporter des comparaisons, personnifications, hyperboles, métonymies (ex: prendre un verre) et des métaphores familière (ex: Je veux suivre un chemin différent et arrêter d'étudier).

Avez-vous des commentaires sur cette dernière partie, relative à la syntaxe, à la cohésion et au style ? (ex: descripteurs à ajouter, commentaires généraux)

Copiez-collez ci-après des textes qui représentent pour vous les niveaux de complexité (un texte par niveau, max. 250 mots par texte).

1 - Très Facile

2 - Facile



3 - Accessible



Ce travail est terminé, merci et bravo !

Si nous avons omis des aspects importants pour définir la complexité des textes au niveau linguistique, merci de nous en faire part ici.



Powered by Qualtrics

Boas-vindas

Bem-vindo/a ao projeto [iRead4Skills!](#)

Este grupo de discussão tem como objetivo definir e caracterizar três níveis de complexidade textual adaptados a um público adulto em processo de alfabetização ou com baixos níveis de literacia. Estes níveis servirão posteriormente para classificar textos relativamente a uma escala de dificuldade e para analisar e modelizar a dificuldade textual para o nosso público-alvo.

A equipa do projeto iRead4Skills agradece antecipadamente a sua participação! As suas respostas são muito importantes para nós.

Título do projeto: iRead4Skills - Intelligent Reading Improvement System for Fundamental and Transversal Skills Development [101094837]

Investigadora principal: Raquel Amaro (raquelamaro@fcsb.unl.pt)

Tarefa preparatória para o grupo de discussão

O objetivo desta tarefa é dar a conhecer as características dos níveis propostos pelas equipas do projeto iRead4Skills. A sua experiência no terreno com o público-alvo do projeto é, assim, essencial. As características listadas abaixo são uma primeira etapa da definição dos níveis de complexidade textual. O grupo de discussão e o posterior trabalho das equipas de investigação completarão esta tarefa.

Para este projeto, propomos os seguintes três níveis de complexidade textual:

Muito fácil: Textos que são totalmente ou quase totalmente compreendidos por todas as pessoas, incluindo pessoas com escolaridade muito baixa (i.e., que não chegaram ao 6º ano do ensino básico) e que não têm quase nenhuma prática de leitura.

Fácil: Textos que são totalmente ou quase totalmente compreendidos por pessoas com escolaridade baixa (i.e., que completaram o 6º ano do ensino básico mas que não têm mais que o 9º ano) e que têm pouca prática de leitura.

Claro: Textos que são compreendidos à primeira leitura por pessoas que completaram o 9º ano e que têm uma prática de leitura média.

Todos os textos para além dos considerados nestes níveis estão acima da nossa escala de dificuldade. Propomos a classificação não claro/+complexo para nos referirmos a esses textos que estão fora do escopo do projeto.

Por favor, reflita antes de responder às perguntas abaixo.

É-lhe pedido que classifique as características relativamente aos níveis de complexidade a que dizem respeito. A cada característica está associada uma caixa de texto onde poderá escrever os comentários que considere relevantes (por exemplo: característica muito vaga, proposta de reformulação da característica, etc).

Pode parar e retomar este inquérito em qualquer altura.

Não há perguntas iguais!

Block 1

Níveis de complexidade do texto

Muito fácil: Textos que são totalmente ou quase totalmente compreendidos por todas as pessoas, incluindo pessoas com escolaridade muito baixa (i.e., que não chegaram ao 6º ano do ensino básico) e que não têm quase nenhuma prática de leitura.

Fácil: Textos que são totalmente ou quase totalmente compreendidos por pessoas com escolaridade baixa (i.e., que completaram o 6º ano do ensino básico mas que não têm mais que o 9º ano) e que têm pouca prática de leitura.

Claro: Textos que são compreendidos à primeira leitura por pessoas que completaram o 9º ano e que têm uma prática de leitura média.

Qual o nível de complexidade que associa às características listadas abaixo?

**muito
fácil** **fácil** **claro** **não é
claro/mais
complexo**

O texto tem até 50 palavras.

O texto usa palavras que referem conceitos e ideias abstratas em contextos multifacetados e/ou âmbitos especializados (e.g., poesia, democracia, fusão nuclear).

O texto emprega todos os tipos de siglas (e.g., ADN, UNESCO, EU, UNICEF).

A maioria das frases no texto estão na voz ativa (e.g., O cão comeu o bolo (ativa) vs. O bolo foi comido pelo cão. (passiva)).

O texto não faz uso de estrangeirismos, exceto aqueles que são muito frequentes e conhecidos (e.g., internet, donut).

**muito
fácil** **fácil** **claro** **não é
claro/mais
complexo**

O texto tem o propósito de desempenhar tarefas familiares/quotidianas (e.g, instruções para comprar bilhetes de autocarro, instruções de utilização no rótulo de um produto).

O texto tem uma ligeira percentagem de palavras que referem ideias abstratas (e.g., felicidade, imaginação, amor).

O texto combina frases através de conjunções coordenativas e subordinativas (e.g., Os músicos foram surpreendidos pela polícia e estão concentrados em frente ao centro comercial. Se ainda não pagou a conta, pode pagar aqui.)

O texto tem elementos repetidos em vez de pronomes (e.g., O João não gosta da Maria porque a Maria não é simpática.)

O texto contém frases simples apresentadas não numa ordem linear (sujeito-verbo-complementos) (e.g., À escola, chegaram os autocarros da visita de estudo), no entanto, estas não estão em maioria no texto.

Block 2

Níveis de complexidade do texto

Muito fácil: Textos que são totalmente ou quase totalmente compreendidos por todas as pessoas, incluindo pessoas com escolaridade muito baixa (i.e., que não chegaram ao 6º ano do ensino básico) e que não têm quase nenhuma prática de leitura.

Fácil: Textos que são totalmente ou quase totalmente compreendidos por pessoas com escolaridade baixa (i.e., que completaram o 6º ano do ensino básico mas que não têm mais que o 9º ano) e que têm pouca prática de leitura.

Claro: Textos que são compreendidos à primeira leitura por pessoas que completaram o 9º ano e que têm uma prática de leitura média.

muito fácil **fácil** **claro** **não é claro/mais complexo**

O texto apresenta uma sequência ou localizações temporais simples e lineares, dadas através de advérbios de tempo e conjunções como hoje, amanhã, antes de, depois de, quando, enquanto (e.g., Os bombeiros chegaram depois da polícia.)

O texto apresenta pronomes e palavras relacionadas para indicar o que é referido (e.g., 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.)

O texto tem uma sequência ou localizações temporais lineares, dadas através de advérbios de tempo ou conjunções como após, assim que, antes de e também por tempos verbais (e.g., Ele está a caminho e a festa já começou. A banda começou a tocar antes de todos saírem.)

O texto utiliza palavras compostas comuns, cujo sentido pode ser mais ou menos compreendido através da combinação das duas partes (e.g., óculos de sol, mal-educado, meia-idade, maus-tratos).

O texto é utilizado em contextos científicos e académicos (e.g., ensaios, relatórios, artigos científicos).

muito fácil **fácil** **claro** **não é claro/mais complexo**

O texto refere conceitos que os leitores precisam de interligar para compreender totalmente (e.g., estrelas, galáxias, vulcões).

muito fácil
fácil
claro
não é claro/mais complexo

O texto utiliza palavras e expressões de âmbitos específicos, jargão ou termos arcaicos (e.g., enfarte do miocárdio, artéria coronária, dissecação espontânea, genoma).

O texto menciona conceitos presentes em âmbitos específicos (e.g., profissional, desporto, hobbies).

O texto tem até 150 palavras.

O texto apresenta um tópico novo (e.g., um texto introdutório sobre a importância da reciclagem).

Block 3

Níveis de complexidade do texto

Muito fácil: Textos que são totalmente ou quase totalmente compreendidos por todas as pessoas, incluindo pessoas com escolaridade muito baixa (i.e., que não chegaram ao 6º ano do ensino básico) e que não têm quase nenhuma prática de leitura.

Fácil: Textos que são totalmente ou quase totalmente compreendidos por pessoas com escolaridade baixa (i.e., que completaram o 6º ano do ensino básico mas que não têm mais que o 9º ano) e que têm pouca prática de leitura.

Claro: Textos que são compreendidos à primeira leitura por pessoas que completaram o 9º ano e que têm uma prática de leitura média.

muito fácil
fácil
claro
não é claro/mais complexo

O texto utiliza palavras derivadas com sufixos e prefixos menos comuns (e.g., megatonelada, tendinite, geopolítico, neoliberal).

O texto tem o propósito do leitor se entreter em momentos de lazer (e.g., desporto, crónica, notícias).

O texto é empregue em contextos especializados (e.g., comunicação profissional formal, fóruns *online*).

muito fácil **fácil** **claro** **não é claro/muito complexo**

O texto utiliza siglas conhecidas, isto é, que podem ser compreendidas tendo sido utilizadas sem explicação previa (e.g, empresas de transportes – CP, impostos – IRS, IVA, partidos políticos – PCP, PS).

O texto emprega palavras que remetem para conceitos e ideias abstratas (e.g., envelhecimento, razão, fidelidade).

muito fácil **fácil** **claro** **não é claro/muito complexo**

As frases no texto incluem frequentemente orações relativas correspondentes ao sujeito ou ao objeto da frase (e.g., O rapaz que abraçou a mãe dirigiu-se à entrada; O rapaz que a mãe abraçou dirigiu-se à entrada).

O texto tem orações relativas correspondentes ao sujeito da frase (e.g., O senhor que vive na casa ao lado ouviu o barulho), contudo, estas não estão em maioria no texto.

O texto circula em contextos artísticos (e.g., literatura, poesia).

O texto tem palavras derivadas com prefixos e sufixos comuns e familiares (e.g., infeliz, calmamente, organizar, trabalhador, casinha), mas estas estão em maioria no texto.

O texto é utilizado em contextos profissionais ou de comunicação dos media (e.g., comunicação ou instrução profissional, notícias, anúncios, desporto).

Block 4

Níveis de complexidade do texto

Muito fácil: Textos que são totalmente ou quase totalmente compreendidos por todas as pessoas, incluindo pessoas com escolaridade muito baixa (i.e., que não chegaram ao 6º ano do ensino básico) e que não têm quase nenhuma prática de leitura.

Fácil: Textos que são totalmente ou quase totalmente compreendidos por pessoas com escolaridade baixa (i.e., que completaram o 6º ano do ensino básico mas que não têm mais que o 9º ano) e que têm pouca prática de leitura.

Claro: Textos que são compreendidos à primeira leitura por pessoas que completaram o 9º ano e que têm uma prática de leitura média.

	muito fácil	fácil	claro	nã claro com
O texto tem frases simples apresentadas numa ordem básica e linear (sujeito-verbo-complementos) (e.g., Eu ajudei os meus filhos a comprar casa.) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
O texto possui várias palavras compostas cujo sentido não é facilmente compreendido através da combinação das duas partes (e.g., pós-graduação, mais-valia, pé de meia, cavalo de batalha). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
O texto apresenta sentidos figurativos comuns (e.g., Este verão os preços dispararam.) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
O texto utiliza palavras derivadas com sufixos e prefixos variados (e.g., nervosismo, lavável, curiosidade, altitude, superpoder, preconceituoso). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
O texto tem frases na voz passiva (e.g., A visita foi paga pelo patrão), no entanto, não estão em maioria. <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	muito fácil	fácil	claro	nã claro com
As frases do texto não estão necessariamente numa ordem básica e linear (sujeito-verbo-complementos) (e.g., De ar puro, todos precisamos.) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
O texto contém palavras polissémicas, ou seja, palavras que possuem mais do que um significado e podem ser ambíguas, bastante frequentes (e.g., base, experiência, leve, seca, aplicar, banco). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
O texto tem até 500 palavras. <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
O texto tem alguns elementos omitidos (e.g., Reparo desde pequena nos pássaros. vs. Eu reparo desde pequena nos pássaros.) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
O texto circula em comunicações do dia a dia (e.g., horários de transportes, ementas, instruções, preços de produtos, comunicações pessoais, etc.). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Block 5

Níveis de complexidade do texto

Muito fácil: Textos que são totalmente ou quase totalmente compreendidos por todas as pessoas, incluindo pessoas com escolaridade muito baixa (i.e., que não chegaram ao 6º ano do ensino básico) e que não têm quase nenhuma prática de leitura.

Fácil: Textos que são totalmente ou quase totalmente compreendidos por pessoas com escolaridade baixa (i.e., que completaram o 6º ano do ensino básico mas que não têm mais que o 9º ano) e que têm pouca prática de leitura.

Claro: Textos que são compreendidos à primeira leitura por pessoas que completaram o 9º ano e que têm uma prática de leitura média.

	muito fácil	fácil	claro	nã claro/ comp
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., segunda-feira, curso, ano escolar). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
O texto faz uso de estrangeirismos comuns frequentemente (e.g. <i>link, hobby, marketing, mail, gay</i>). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
O texto tem estrangeirismos comuns (e.g., <i>site, stress</i>), mas são uma percentagem reduzida. <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
A maioria das frases no texto são curtas e diretas ou combinadas através de conjunções familiares (e.g., <i>A criança fugiu ontem. O pai trabalha na fábrica, mas a mãe fica em casa</i>). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
O texto tem até 250 palavras. <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	muito fácil	fácil	claro	nã claro/ comp
O texto faz uso da metonímia (e.g. <i>beber um copo</i>);, assim como metáforas familiares (e.g., <i>O meu caminho é outro. Não quero estudar mais.</i>) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
O texto utiliza apenas comparações, personificações e hipérboles (e.g., <i>Gostar de alguém é como fazer prédios; As doenças escondem-se em todo o lado; Eu estava gelada; A tua mala pesa uma tonelada.</i>) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	muito fácil	fácil	claro	não claro/ comp
O texto utiliza siglas depois de as ter explicado (e.g., Organização das Nações Unidas (ONU), Instituto Público (IP)). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
O texto usa pronomes (e.g., O João não gosta da Maria porque ela não é simpática). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
O texto é utilizado em contextos de lazer (e.g., literatura e ficção, reportagens, artigos de opinião). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Block 6

Níveis de complexidade do texto

Muito fácil: Textos que são totalmente ou quase totalmente compreendidos por todas as pessoas, incluindo pessoas com escolaridade muito baixa (i.e., que não chegaram ao 6º ano do ensino básico) e que não têm quase nenhuma prática de leitura.

Fácil: Textos que são totalmente ou quase totalmente compreendidos por pessoas com escolaridade baixa (i.e., que completaram o 6º ano do ensino básico mas que não têm mais que o 9º ano) e que têm pouca prática de leitura.

Claro: Textos que são compreendidos à primeira leitura por pessoas que completaram o 9º ano e que têm uma prática de leitura média.

	muito fácil	fácil	claro	não é claro/m comple
O texto menciona conceitos relacionados com atividades do dia a dia (e.g., comida, lida da casa, roupa). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
O texto apresenta sentidos figurativos menos comuns (e.g., O Barcelona varreu o Real Madrid; O Estoril esmagou o Santa Clara). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
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). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

muito fácil **fácil** **claro** **não é claro/muito complexo**

O texto contém palavras polissémicas, ou seja, palavras que possuem mais do que um sentido e podem ser ambíguas, pouco frequentes. (e.g., interno, representar, forma, paciente).

O texto apresenta diversas figuras de estilo (e.g., 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...).

muito fácil **fácil** **claro** **não é claro/muito complexo**

As frases no texto são apresentadas em diversas ordens não lineares (sujeito-verbo-complementos; verbo-complementos-sujeito, etc.) (e.g., Até final de março, porque antes seria impossível encontrar todos os ingredientes, haverá na ementa vários pratos de caça.)

As frases no texto são longas e compostas por diversas outras frases combinadas através de conjunções coordenativas e subordinativas (e.g., 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.)

O texto tem palavras e expressões menos conhecidas que fazem referência a âmbitos específicos (e.g., doença cardíaca, artéria, tratamento).

O texto tem orações relativas correspondentes ao sujeito da frase (e.g., O senhor que vive na casa ao lado ouviu o barulho.)

As frases no texto incluem frequentemente orações relativas correspondentes ao sujeito da frase (e.g., Francisca, a quem os amigos tratam por «Kika», começou a surfar muito pequena.)

Block 7

Níveis de complexidade do texto

Muito fácil: Textos que são totalmente ou quase totalmente compreendidos por todas as pessoas, incluindo pessoas com escolaridade muito baixa (i.e., que não chegaram ao 6º ano do ensino

básico) e que não têm quase nenhuma prática de leitura.

Fácil: Textos que são totalmente ou quase totalmente compreendidos por pessoas com escolaridade baixa (i.e., que completaram o 6º ano do ensino básico mas que não têm mais que o 9º ano) e que têm pouca prática de leitura.

Claro: Textos que são compreendidos à primeira leitura por pessoas que completaram o 9º ano e que têm uma prática de leitura média.

	muito fácil	fácil	claro	não claro/n compl
O texto apresenta sentidos figurativos menos incomuns (e.g., Os primeiros raios de sol aconchegaram-me as faces). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
O texto faz uso de metáforas incomuns (e.g., Marcelo e Costa, um casamento de convergência. Os dois coabitam desde 9 de março de 2016, quando Marcelo assumiu funções.) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
O texto não utiliza siglas. <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
O texto é majoritariamente composto por palavras que referem objetos concretos (e.g., cadeira, maçã, martelo, computador). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
O texto utiliza palavras compostas comuns (e.g., feijão verde, couve-flor) ou palavras compostas transparentes, ou seja, palavras ou expressões que podem ser facilmente compreendidas através da combinação das duas partes (e.g., tira-caricás). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	muito fácil	fácil	claro	não claro/n compl
O texto utiliza abreviações comuns (e.g., m, km, g., l., Exmo.). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
As frases no texto são frequentemente combinadas através de conjunções coordenativas e subordinativas e, por vezes, com conjunções menos comuns (e.g., Apesar da confusão provocada pelos animais à solta, a situação foi controlada sem os animais serem atropelados.) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
O texto não faz uso de sentidos figurativos (e.g., Este verão os preços subiram vs. Este verão os preços dispararam (fig.)) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
O texto faz referência a conceitos novos e multifacetados (e.g., matéria negra, desenvolvimento do cérebro, estatísticas). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	muito fácil	fácil	claro	não claro/n compl
O texto tem nenhum ou poucos elementos omitidos (e.g., O acidente causou dois feridos leves. Os bombeiros levaram os feridos para o hospital.) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Block 8

Níveis de complexidade do texto

Muito fácil: Textos que são totalmente ou quase totalmente compreendidos por todas as pessoas, incluindo pessoas com escolaridade muito baixa (i.e., que não chegaram ao 6º ano do ensino básico) e que não têm quase nenhuma prática de leitura.

Fácil: Textos que são totalmente ou quase totalmente compreendidos por pessoas com escolaridade baixa (i.e., que completaram o 6º ano do ensino básico mas que não têm mais que o 9º ano) e que têm pouca prática de leitura.

Claro: Textos que são compreendidos à primeira leitura por pessoas que completaram o 9º ano e que têm uma prática de leitura média.

	muito fácil	fácil	claro	não é claro/m comple
O texto apresenta frases na voz passiva (e.g., O almoço foi pago pela empresa.), incluindo construções impessoais comuns e passivas de -se (e.g., O trabalho faz-se bem.) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
O texto utiliza palavras e expressões simples relacionadas com assuntos do quotidiano, fazendo uso de substantivos frequentes (e.g. batatas, azeite, legumes). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
O texto serve para aprender/realizar tarefas novas (e.g., instruções sobre como utilizar o e-mail da empresa). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
O texto tem diversas sequências ou localizações temporais não lineares (e.g., <u>Será preciso fazer</u> alguma coisa aos cadáveres e, por isso, é bom que <u>haja gente que se debruce</u> sobre o tema. Eu <u>seria</u> incapaz, porque nem sequer <u>consigo ir</u> a funerais, como já <u>fui</u> , onde <u>repousam</u> caixões abertos com os indivíduos em descanso eterno.) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	muito fácil	fácil	claro	não é claro/m comple
O texto não emprega abreviações (e.g., km, c., V. Exa), exceto formas frequentes de tratamento (e.g., Sr. Dra.). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
O texto tem palavras compostas cujo sentido não é facilmente compreendido através da combinação das duas partes (e.g., meio-termo, porta-voz, obra prima). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
O texto apresenta elementos omitidos (e.g., Mostrámos em estudos anteriores que a gordura solidifica quando arrefece. vs. Nós mostrámos em estudos anteriores que a gordura solidifica quando ela arrefece). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
O texto inclui uma sequência ou localizações temporais não lineares (e.g., Após analisarem as respostas dos doentes que tinham sofrido um ataque cardíaco, os médicos confirmaram que a gordura é o alimento mais perigoso.) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
O texto utiliza abreviações comuns após ter explicado o que significam (e.g., l – litro, m – metro). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
O texto tem palavras derivadas com prefixos e sufixos frequentes (e.g. pré-cozinhado, desorganizar, embalagem, organização, jardineiro, finíssimo), mas estas não estão em maioria no texto. <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Block 9

Níveis de complexidade do texto

Muito fácil: Textos que são totalmente ou quase totalmente compreendidos por todas as pessoas, incluindo pessoas com escolaridade muito baixa (i.e., que não chegaram ao 6º ano do ensino básico) e que não têm quase nenhuma prática de leitura.

Fácil: Textos que são totalmente ou quase totalmente compreendidos por pessoas com escolaridade baixa (i.e., que completaram o 6º ano do ensino básico mas que não têm mais que o 9º ano) e que têm pouca prática de leitura.

Claro: Textos que são compreendidos à primeira leitura por pessoas que completaram o 9º ano e que têm uma prática de leitura média.

	muito fácil	fácil	claro	nã claro, comj
Verbos no infinitivo(e.g., gostar, comer, fugir) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Verbos no Presente do Indicativo (e.g., gosto, comem, foge) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Verbos no Pretérito Imperfeito do Indicativo (e.g., gostavam, comia, fugias) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Verbos no Pretérito Perfeito do Indicativo (e.g., gostou, comi, fugiram) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Verbos no Pretérito Mais-que-perfeito do Indicativo (e.g., gostara, comeras, fugira) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Verbos no Futuro Simples do Indicativo (e.g., esquecerá, gostará) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	muito fácil	fácil	claro	nã claro, comj
Verbos no Gerúndio (e.g., gostando, comendo, fugindo) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Verbos no Imperativo (e.g., come!, falai!; não comas!; não falem!) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Verbos no Condicional (e.g., gostaria, comerias, fugiriam) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Verbos no Presente do Conjuntivo (e.g., goste, coma, fujam) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Verbos no Pretérito Imperfeito do Conjuntivo (e.g., gostasses, comessem, fugisse) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Verbos no Futuro do Conjuntivo (e.g., se eu gostar, se tu comeres, se eles fugirem) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	muito fácil	fácil	claro	nã claro, comj
Verbos no Pretérito Perfeito Composto do Indicativo (e.g., tenho gostado, tem comido, têm fugido) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	muito fácil	fácil	claro	nã claro, comj
Verbos no Pretérito Mais-que-perfeito Composto do Indicativo (e.g., tinhas gostado, tinham comido, tinha fugido) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Verbos no Futuro Composto do Indicativo (e.g., terei gostado, terás comido, terão fugido) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Verbos no Pretérito Perfeito Composto do Conjuntivo (e.g., tenha gostado, tenha comido, tenham fugido) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Verbos no Pretérito Mais-que-perfeito Composto do Conjuntivo (e.g., tivesses gostado, tivessem comido, tivesse fugido) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Verbos no Futuro Composto do Conjuntivo (e.g., tiverem gostado, tiver comido, tiveres fugido) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	muito fácil	fácil	claro	nã claro, comj
Verbos auxiliares poder, ir, estar no Indicativo (e.g., posso gostar, vai comer, estão a fugir) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Expressões com os verbos permanecer, querer, poder, conseguir em tempos simples (e.g., permaneça calmo, queria almoçar, consigo ir). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Expressões com os verbos estar, ser, permanecer, querer, poder, dever em tempos compostos (e.g., tenham estado ligados; foram feitas; temos de permanecer calmos; devíamos ir) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Block 10

Esta parte do trabalho está finalizada. Muito obrigado!

Se omitimos algum aspeto importante para definir a complexidade dos textos a nível linguístico, por favor, indique aqui.

Boas-vindas

Bienvenido al proyecto iRead4Skills

Este grupo de discusión se propone definir y caracterizar tres niveles de complejidad textual adaptados al público adulto en proceso de alfabetización o con bajos niveles de alfabetización. Los niveles servirán posteriormente para clasificar textos en una escala de dificultad y para analizar y modelizar la dificultad textual para nuestro público meta.

¡El equipo del proyecto iRead4Skills le agradece de antemano su participación! Sus respuestas son muy importantes para nosotros.

Título del proyecto: iRead4Skills - Intelligent Reading Improvement System for Fundamental and Transversal Skills Development

Este proyecto está financiado por la Comisión Europea, HORIZON-CL2-2022-TRANSFORMATIONS-01-07 Program (ref. 101094837)

Responsable para la investigación sobre el español: Xavier Blanco Escoda

Contacto: xavier.blanco@uab.cat

Tarea preparatoria para los grupos de discusión

El objetivo de esta tarea es presentarle descriptores de niveles propuestos por los equipos del proyecto iReadk4Skills. La experiencia en el terreno y con el público meta del proyecto que usted tiene resulta crucial para la buena marcha del proyecto. Esos descriptores son sólo una primera etapa en la definición de niveles de complejidad textual, que completarán los grupos de discusión y un posterior trabajo de los equipos de investigación.

Se proponen **tres niveles de complejidad textual** para este proyecto:

Muy fácil: textos que comprenden por entero o casi totalmente todos los adultos, incluso personas con escasa escolarización (esto es, que no han terminado la enseñanza primaria) y que apenas practican la lectura.

Fácil: textos que comprenden por entero o casi totalmente personas con escasa

escolarización (esto es, personas que han terminado la enseñanza primaria, pero que no han continuado su formación) y que practican poco la lectura.

Accesible: textos que entienden ya en primera lectura personas que han proseguido su formación más allá de la enseñanza primaria y que practican moderadamente la lectura.

Todos los textos no contemplados en esos niveles sobrepasan nuestra escala de dificultad. Para esos textos, que no forman parte del objetivo de nuestro proyecto, proponemos el nivel "+ Complejo".

A partir de las preguntas siguientes le rogamos que clasifique los descriptores según los niveles de complejidad. Cada descriptor está asociado a un cuadro de texto que le permitirá escribir un comentario al respecto si lo estima oportuno (por ejemplo, descriptor demasiado vago, propuesta de reformulación del descriptor...).

Descriptores

Niveles de complejidad

Muy fácil: Los textos los entiende prácticamente todo el mundo, incluso personas con pocos años de escolaridad (esto es, las que han completado la enseñanza primaria) y con ninguna práctica de lectura.

Fácil: Los textos los entienden por completo casi por completo personas con pocos años de escolaridad (esto es, las que han finalizado la enseñanza primaria, pero no la enseñanza secundaria obligatoria) y con poca práctica de lectura.

Claro: Los textos los entienden en primera lectura personas que han completado la enseñanza secundaria obligatoria y cuya práctica de lectura corresponde a la media.

muy fácil fácil claro más complejo

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

El texto tiene hasta 500 palabras.

muy
fácil fácil claro más
complejo

En el texto se utilizan palabras que se refieren a conceptos e ideas abstractos (por ej., *envejecimiento, razonamiento, lealtad*).

En el texto se utilizan acrónimos y siglas que son comúnmente conocidos, esto es, se pueden entender de manera aislada (e.g., *RENFE, IVA, PSOE, PP* (compañía ferroviaria, impuestos, partidos políticos...)).

Las oraciones del texto siguen distintos órdenes de palabras (sujeto-verbo-complementos; verbo-complementos-sujeto, etc.) (por ej., *Compraron los niños azúcar para hacer la tarta; Perdieron la guerra los que la habían iniciado.*)

muy
fácil fácil claro más
complejo

El texto se compone sobre todo de palabras que se refieren a conceptos concretos (por ej., *silla, manzana, martillo, ordenador*).

El texto presenta una secuenciación temporal lineal simple, expresada por medio de adverbios temporales o conectores como *desde que, en cuanto, después de que* y por medio de tiempos verbales (por ej., *Cuando llegó María, la fiesta ya había comenzado.*)

En el texto hay sentidos figurados (por ej., *Este verano los precios se dispararán.*)

El texto contiene oraciones con el pronombre “se” incluso pasivas reflejas (por ej., *En Barcelona se venden muchos pisos*).

La mayoría de las oraciones del texto son cortas y directas o se articulan por medio de conectores frecuentes y comunes (por ej., *La casa es grande, bonita y además es barata.*)

Block 2

Niveles de complejidad

Muy fácil: Los textos los entiende prácticamente todo el mundo, incluso personas con pocos años de escolaridad (esto es, las que han completado la enseñanza primaria) y con ninguna

práctica de lectura.

Fácil: Los textos los entienden por completo casi por completo personas con pocos años de escolaridad (esto es, las que han finalizado la enseñanza primaria, pero no la enseñanza secundaria obligatoria) y con poca práctica de lectura.

Claro: Los textos los entienden en primera lectura personas que han completado la enseñanza secundaria obligatoria y cuya práctica de lectura corresponde a la media.

muy
fácil fácil claro más
complejo

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.).

Las oraciones del texto con frecuencia 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.*)

En el texto se utiliza la metonimia (por ej., *tomarse una copa*), metáforas corrientes (por ej., *No voy a estudiar más. Mi camino es otro.*)

En el texto se utilizan abreviaciones corrientes (por ej., *m., km., gr., l., Excmo., ...*).

En el texto se usan pronombres y distintos tipos de palabras para indicar a qué se refiere (por ej., *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 estas 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.*)

muy
fácil fácil claro más
complejo

El texto se refiere a actividades cotidianas. (por ej., comida, cuidado del hogar, ropa).

muy fácil fácil claro más complejo

El texto tiene pocos o ningún elemento omitido (por ej., *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*).

El texto contiene oraciones de relativo que complementan el sujeto de la oración (por ej., *El señor que vive en la casa de al lado escuchó el ruido*), pero no son mayoría en el texto.

El texto tiene hasta 250 palabras.

El texto se utiliza en contextos de ocio (por ej., literatura y ficción, reportaje, artículos de opinión).

Block 3

Niveles de complejidad

Muy fácil: Los textos los entiende prácticamente todo el mundo, incluso personas con pocos años de escolaridad (esto es, las que han completado la enseñanza primaria) y con ninguna práctica de lectura.

Fácil: Los textos los entienden por completo casi por completo personas con pocos años de escolaridad (esto es, las que han finalizado la enseñanza primaria, pero no la enseñanza secundaria obligatoria) y con poca práctica de lectura.

Claro: Los textos los entienden en primera lectura personas que han completado la enseñanza secundaria obligatoria y cuya práctica de lectura corresponde a la media.

muy fácil fácil claro más complej

En el texto se utilizan abreviaciones corrientes tras explicarlas (por ej., *l. – litro, m. – metro*).

El texto se propone informar al lector en momentos de ocio (por ej., crónica deportiva, noticias).

	muy fácil	fácil	claro	más complej
Hay elementos repetidos en lugar de pronombres (por ej., <i>A Juan no le gusta María porque María no es amable</i>). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
El texto se utiliza en contextos especializados (por ej., comunicación profesional formal, foros en Internet). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
En el texto no se utilizan ni acrónimos ni siglas. <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	muy fácil	fácil	claro	más complej
El texto presenta una secuenciación temporal lineal simple, expresada por medio de adverbios o conectores como hoy, mañana, antes de, después de, cuando, mientras (por ej., <i>Pablo se levanta a las 7, después se ducha y toma el desayuno</i>). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
El texto tiene algunos elementos omitidos (por ej., <i>Juan se ha comprado una casa. Es muy bonita y le ha salido muy barata</i>). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
El texto tiene palabras derivadas con sufijos y prefijos muy variados (e.g., <i>nerviosismo, lavable, curiosidad, altitud, superpoder, preconcebido</i>). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
En el texto se utiliza todo tipo de acrónimos o siglas (por ej., <i>ADN, UNESCO, UNICEF, EE UU</i>). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
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). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Block 4

Niveles de complejidad

Muy fácil: Los textos los entiende prácticamente todo el mundo, incluso personas con pocos años de escolaridad (esto es, las que han completado la enseñanza primaria) y con ninguna práctica de lectura.

Fácil: Los textos los entienden por completo casi por completo personas con pocos años de escolaridad (esto es, las que han finalizado la enseñanza primaria, pero no la enseñanza secundaria obligatoria) y con poca práctica de lectura.

Claro: Los textos los entienden en primera lectura personas que han completado la enseñanza secundaria obligatoria y cuya práctica de lectura corresponde a la media.

	muy fácil	fácil	claro	m com
El texto tiene un porcentaje muy pequeño de palabras que se refieren a ideas abstractas corrientes (por ej., <i>felicidad. Imaginación, amor</i>). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
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, portavoz, ópera prima</i>). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
El texto tiene hasta 50 palabras. <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
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>). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
El texto tiene palabras y expresiones menos conocidas que se refieren a ámbitos específicos (por ej., <i>enfermedad cardíaca, arteria, tratamiento</i>). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	muy fácil	fácil	claro	m com
El texto se utiliza en contextos académicos o científicos (por ej., ensayos, informes, artículos académicos). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
El texto tiene palabras y expresiones que se refieren a ámbitos específicos, jergas especializadas o términos arcaicos (por ej., <i>infarto de miocardio, arteria coronaria, disección espontánea, genoma</i>). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
El texto enseña a llevar a cabo tareas nuevas (por ej., instrucciones sobre cómo usar el correo electrónico de la empresa). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
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, caballo de batalla, en pie de guerra</i>). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
El texto se refiere a conceptos nuevos o con múltiples facetas. (por ej., materia oscura, desarrollo cerebral, estadística). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Block 5

Niveles de complejidad

Muy fácil: Los textos los entiende prácticamente todo el mundo, incluso personas con pocos años de escolaridad (esto es, las que han completado la enseñanza primaria) y con ninguna práctica de lectura.

Fácil: Los textos los entienden por completo casi por completo personas con pocos años de escolaridad (esto es, las que han finalizado la enseñanza primaria, pero no la enseñanza secundaria obligatoria) y con poca práctica de lectura.

Claro: Los textos los entienden en primera lectura personas que han completado la enseñanza secundaria obligatoria y cuya práctica de lectura corresponde a la media.

	muy fácil	fácil	claro	má comp
El texto tiene hasta 150 palabras. <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
El texto contiene oraciones de relativo que complementan el sujeto de la oración (por ej., <i>El señor que vive en la casa de al lado escuchó el ruido.</i>) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
En el texto se utilizan metáforas poco comunes (por ej., <i>Se destejen los días. Las noches se consumen antes de darnos cuenta.</i>) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
El texto se refiere a conceptos que requieren juntar distintas informaciones. (por ej., estrellas, volcanes). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
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>) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	muy fácil	fácil	claro	má comp
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>) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
El texto presenta un tema nuevo (por ej., texto de iniciación sobre la importancia del reciclaje). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

muy fácil fácil claro más complejo

El texto tiene palabras compuestas cuyo significado se puede entender más o menos combinando el significado de las partes (por ej., *gafas de sol, mediana edad, maltrato*).

El texto tiene palabras derivadas con prefijos y sufijos corrientes y frecuentes (por ej., *infeliz, tranquilamente, organizar, trabajador, casita*), pero no son mayoritarias en el texto.

El texto puede tener una secuenciación temporal no lineal (por ej., *Pablo llegó tarde a casa; antes había pasado por el supermercado para comprar pan; su mujer se lo había pedido porque ella no tenía tiempo de ir*).

Block 6

Niveles de complejidad

Muy fácil: Los textos los entiende prácticamente todo el mundo, incluso personas con pocos años de escolaridad (esto es, las que han completado la enseñanza primaria) y con ninguna práctica de lectura.

Fácil: Los textos los entienden por completo casi por completo personas con pocos años de escolaridad (esto es, las que han finalizado la enseñanza primaria, pero no la enseñanza secundaria obligatoria) y con poca práctica de lectura.

Claro: Los textos los entienden en primera lectura personas que han completado la enseñanza secundaria obligatoria y cuya práctica de lectura corresponde a la media.

muy fácil fácil claro más complejo

En el texto se utilizan distintas figuras de estilo (por ej., *La luna me sonríe desde lo alto del cielo, el reloj me grita la hora; yo me esfuerzo por olvidarte y sin querer te recuerdo*).

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).

	muy fácil	fácil	claro	más complej
En el texto se utiliza solo la comparación, la personificación, la hipérbole (por ej., <i>Duerme como un tronco. Los virus se esconden por todas partes. Esa maleta pesa una tonelada.</i>) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
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). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
En el texto no hay sentidos figurados (por ej., <i>Este verano los precios subirán vs. Este verano se dispararán</i> (fig.)). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	muy fácil	fácil	claro	más complej
El texto tiene palabras frecuentes y expresiones que se pueden referir a cuestiones cotidianas o laborales / profesionales (por ej., <i>lunes, clase, año académico</i>). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
El texto tiene palabras derivadas con sufijos y prefijos muy poco comunes (e.g., <i>megafiesta, tendinitis, geopolítico</i>). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
El texto se utiliza en ámbitos artísticos (por ej., literatura, poesía). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
El texto contiene con frecuencia oraciones de relativo que complementan el sujeto de la oración (por ej., <i>Francisca, a quien sus amigos llaman «Kika», comenzó a surfear de muy pequeña.</i>). <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
El texto contiene oraciones con el pronombre “se” (por ej., <i>Pablo se lava la cara</i>), pero no son la mayoría en el texto. <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Block 7

Niveles de complejidad

Muy fácil: Los textos los entiende prácticamente todo el mundo, incluso personas con pocos años de escolaridad (esto es, las que han completado la enseñanza primaria) y con ninguna práctica de lectura.

Fácil: Los textos los entienden por completo casi por completo personas con pocos años de escolaridad (esto es, las que han finalizado la enseñanza primaria, pero no la enseñanza secundaria obligatoria) y con poca práctica de lectura.

Claro: Los textos los entienden en primera lectura personas que han completado la enseñanza secundaria obligatoria y cuya práctica de lectura corresponde a la media.

	muy fácil	fácil	claro	más complejo
<p>En el texto hay sentidos figurados inhabituales (por ej., <i>el latir del reloj</i>).</p> <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<p>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>).</p> <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<p>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>).</p> <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<p>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>).</p> <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<p>El texto se refiere a conceptos relacionados con usos de ámbitos específicos (por ej., trabajo, deportes, aficiones).</p> <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	muy fácil	fácil	claro	más complejo
<p>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.</p> <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<p>El texto no tiene palabras extranjeras, salvo las que son muy comunes o son muy conocidas (por ej., <i>Internet, donut</i>).</p> <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<p>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>).</p> <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<p>El texto tiene palabras extranjeras comunes (por ej., <i>estrés, rock</i>), pero no son la mayoría en el texto.</p> <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

muy más
fácil fácil claro comple

Las oraciones del texto no siguen necesariamente el orden básico de palabras (sujeto-verbo-complementos (por ej., *A París voy, no a Londres.*).

Block 8

Niveles de complejidad

Muy fácil: Los textos los entiende prácticamente todo el mundo, incluso personas con pocos años de escolaridad (esto es, las que han completado la enseñanza primaria) y con ninguna práctica de lectura.

Fácil: Los textos los entienden por completo casi por completo personas con pocos años de escolaridad (esto es, las que han finalizado la enseñanza primaria, pero no la enseñanza secundaria obligatoria) y con poca práctica de lectura.

Claro: Los textos los entienden en primera lectura personas que han completado la enseñanza secundaria obligatoria y cuya práctica de lectura corresponde a la media.

muy más
fácil fácil claro comple

En el texto se utilizan palabras que se refieren a conceptos e ideas con muchas facetas y/o de ámbitos específicos (por ej., *poesía, democracia, fusión nuclear.*

El texto tiene oraciones que se articulan por medio de conectores de coordinación y subordinación (por ej., *La casa es barata, pero no me gusta. Mi hermano me dijo que compraría la casa).*

El texto contiene oraciones con el pronombre “se” incluso pasivas impersonales (por ej., *Se detuvo a los ladrones en dos horas.*

El texto tiene palabras polisémicas (esto es, palabras que pueden tener más de un significado y pueden ser ambiguas), pero no son la mayoría en el texto (por ej., *interno, representar, forma, paciente.*

muy más
fácil fácil claro comple

muy fácil fácil claro más comple

El texto tiene muchas palabras polisémicas (esto es, palabras que pueden tener más de un significado y pueden ser ambiguas) (por ej., *base, experiencia, leve, aplicar*).

El texto contiene oraciones simples que no siguen el orden básico de palabras (sujeto-verbo-complementos) (por ej., *Al niño le han regalado una bicicleta sus padres*), pero no son mayoritarias en el texto.

En el texto se utilizan acrónimos y siglas después de describirlos (por ej., *Organización de las Naciones Unidas (ONU), instituto de enseñanza secundaria (IES)*).

El texto tiene palabras y expresiones simples relativas a cuestiones cotidianas y usa nombres frecuentes y comúnmente conocidos (por ej., *patatas, aceite de oliva, verduras*).

muy fácil fácil claro más comple

La mayoría de las oraciones del texto están en voz activa (por ej., *El gato ha cazado el ratón* (activa) vs. *El ratón ha sido cazado por el gato*. (pasiva)).

En el texto se utilizan palabras extranjeras comunes con frecuencia (por ej., *link, hobby, marketing, mail, gay*).

El texto utiliza pronombres (por ej., *Pablo contrató a María porque se la recomendó un amigo*).

Block 9

Niveles de complejidad

Muy fácil: Los textos los entiende prácticamente todo el mundo, incluso personas con pocos años de escolaridad (esto es, las que han completado la enseñanza primaria) y con ninguna práctica de lectura.

Fácil: Los textos los entienden por completo casi por completo personas con pocos años de escolaridad (esto es, las que han finalizado la enseñanza primaria, pero no la enseñanza secundaria obligatoria) y con poca práctica de lectura.

Claro: Los textos los entienden en primera lectura personas que han completado la enseñanza secundaria obligatoria y cuya práctica de lectura corresponde a la media.

	muy fácil	fácil	claro	m: comi
Verbos en infinitivo (por ej., <i>gustar, comer, huir</i>) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Verbos en presente de indicativo (por ej., <i>gusta, come, huye</i>) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Verbos en imperfecto de indicativo (por ej., <i>gustaba, comía, huías</i>) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Verbos en pretérito indefinido (por ej., <i>gustó, comí, huyeron</i>) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Verbos en pretérito perfecto (por ej., <i>ha gustado, has comido, han huido</i>) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Verbos en futuro simple (por ej., <i>comerán, gustará</i>) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	muy fácil	fácil	claro	m: comi
Verbos en gerundio (por ej., <i>gustando, comiendo, huyendo</i>) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Verbos en condicional simple (por ej., <i>gustaría, comerías, huirían</i>) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Verbos en condicional compuesto (por ej., <i>habría gustado, habrías comido, habrían huido</i>) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Verbos en presente del subjuntivo (por ej., <i>guste, coma, huyan</i>) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Verbos en imperfecto de subjuntivo (por ej., <i>gustara, comiesen, huyera</i>) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Verbos en imperativo (por ej., <i>mira, oíd, estudiad</i>) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	muy fácil	fácil	claro	m: comi
Verbos en pretérito perfecto del subjuntivo (por ej., <i>haya ido, hayas oído, hayan estudiado</i>) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Verbos en pluscuamperfecto de subjuntivo (por ej., <i>hubiera gustado, hubiese comido, hubiera huido</i>) <input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

muy fácil fácil claro muy difícil

Verbos en futuro compuesto del indicativo (por ej., *habrá gustado, habrás comido, habré huido*)

Verbos auxiliares *poder, ir, estar* en los tiempos del indicativo (por ej., *puede gustar, voy a comer, está huyendo*)

Expresiones con verbos como *permanecer, querer, poder, conseguir* en tiempos simples (por ej., *permanezca callado, quería almorzar, consigo ganar*)

Expresiones con verbos como *estar, ser, permanecer, querer, poder, deber* en tiempos compuestos (por ej., *habían querido estudiar; han estado huyendo, habíamos debido huir*)

Block 10

Se ha terminado el trabajo. ¡Muchas gracias!

Si hemos omitido algún aspecto importante para definir la complejidad de textos a nivel lingüístico, le agradeceríamos que nos lo dijera aquí.

Powered by Qualtrics

APPENDIX III - Complexity Levels Descriptions for French

Descripteurs de complexité textuelle

Groupe de travail - Novembre 2023

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 soeur,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: vieillissement, 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 causatives 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 imparfait. 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) Résultats du groupe de travail (tâche préparatoire + groupes de discussion)

		1) Très Facile	2) Facile	3) Accessible	4) +Complexe
Taille	Textes courts	< 50 mots	< 150 mots		
	Textes de tailles variées			< 250 mots	< 500 mots
But	Accomplir des tâches habituelles/quotidiennes	ex: cuisine, horaires de bus, infos pratiques	ex: notices d'utilisation, tutoriel		
	S'informer pendant son temps libre		ex: presse quotidienne, information de proximité, météo locale	ex: infos plus généraliste, reportage	
	Présenter un nouveau sujet				
	Apprendre/effectuer de nouvelles tâches				
	S'informer sur des sujets variés + lecture plaisir				

		1) Très Facile	2) Facile	3) Accessible	4) +Complexe
Domaine / Contexte	Communications quotidiennes (horaires de transport, menus)				
	Contextes de loisirs (fiction, reportages)				
	Contextes professionnels, médiatiques et artistiques				
	Contextes spécialisés et académiques				
Concepts	Du quotidien (nourriture, ménage, vêtements)				
	Liés à un domaine spécifique (travail, sport, loisirs)		proche du lecteur		
	Qui exigent de rassembler différents éléments d'information			ex: concepts complexes, texte introductif (volcans, étoiles)	ex: concepts complexes traités en profondeur
	A multiples facettes ou nouveaux (la matière noire, les statistiques)				
Concrétude	Principalement des objets concrets (pomme, ordinateur)				
	Quelques idées abstraites courantes (bonheur, amour, imagination)		ex: affectif, émotions, partagées par tous		
	Idées et concepts abstraits (vieillesse, loyauté)				
	Idées et concepts abstraits à multiples facettes et/ou spécifiques à un domaine (poésie, démocratie, fusion nucléaire)				

		1) Très Facile	2) Facile	3) Accessible	4) +Complexe
Lexique / Vocabulaire	Mots et expressions de la vie quotidienne, simples et fréquents				
	Mots et expressions fréquents qui se réfèrent au contexte professionnel (lundi, réunion, sortie)				
	Mots et expressions qui se réfèrent à des domaines spécifiques				
	Jargon spécialisé et termes archaïques (génom, mout)				
	Pas de mots étrangers, sauf mots largement connus (Internet)				
	Peu de mots étrangers, et ces mots sont courants (email)				
	Nombreux mots étrangers (marketing, feedback, low-cost)				

Lexique / Vocabulaire	Pas d'abréviations, sauf formes très fréquentes (Mme, Prof.)				
	Abréviations courantes (km, g., rdv, svp)				
	Abréviations après explications dans le texte (millilitre (mL))				
	Pas d'acronymes				
	Acronymes et sigles passés dans le langage courant (SNCF, EDF, CDD, PDG)				
	Acronymes et sigles après explications (Parti socialiste (PS))				
	Tous types d'acronymes et sigles (ADN, UNESCO, UNICEF)				

		1) Très Facile	2) Facile	3) Accessible	4) +Complexe
Morphologie	Pas ou peu de mots dérivés	affixes très fréquents (re-, -ment, -tion)			
	Présence de mots dérivés		affixes courants (dé-, in-, -age, -ité, -eur)	affixes variés (auto-, anti-, co-, ex-, inter-, mini-, pré-, sou(s)-, sur-, -able, -aire, -esse, -ien, -iste, -rie)	affixes peu courants (mega-, geo-, -(il)fier)
	Expressions et mots composés fréquents (grand-mère, pomme de terre)				
	Expressions et mots composés dont le sens peut être compris en combinant le sens des parties (sac à dos, non-fumeur)				
	Expressions et mots composés dont le sens ne peut pas être facilement compris en combinant le sens des parties			expressions courantes (troisième âge, rez-de-chaussée)	expressions peu fréquentes (ça et là, année-lumière, à tout bout de champ)

		1) Très Facile	2) Facile	3) Accessible	4) +Complexe
Ordre syntaxique	Phrases simples, dans l'ordre de base sujet-verbe-complément				
	Quelques phrases simples ne suivant pas l'ordre de base	compléments de temps et de lieu			
	Phrases simples et composées variété d'ordre syntaxique			variété d'ordre syntaxique	
Phrases/propositions	Phrases/propositions courtes et directes ou combinées à l'aide de connecteurs fréquents et familiers (ex: et, mais)				
	Peu de propositions subordonnées relatives, temporelles et causatives		connecteurs fréquents (qui, que, quand, comme, parce que)	connecteurs variés (dont, avant que, après que, à cause de, alors)	
	Propositions combinées à l'aide de connecteurs de coordination et subordination (ex: donc, si)		connecteurs courants		
	Propositions combinées à l'aide de connecteurs parfois peu fréquents (ex: malgré)				
	Majorité de phrases longues et complexes + présence de connecteurs peu fréquents (ex: étant donné que)				
Voix passive	Pas ou très peu de phrases à la voix passive				
	Phrases à la voix passive et structures passives avec "se faire"				
	Phrases à la voix passive à l'indicatif				
Cohésion textuelle	Présence d'éléments répétés à la place des pronoms				
	Pas ou très peu d'ellipses (Marc aime la voile, sa soeur le kayak)	pas d'ellipses	très peu		
	Anaphores pronominales	pronoms pers sujet			
	Ellipses fréquentes				
	Anaphores pronominales et nominales (Jean est boulanger. Il a repris l'échoppe de son père. Cet enfant du pays est la fierté du village.)				
Cohésion spatio-temporelle	Cadre spatio-temporel linéaire	adverbes et connecteurs courants (aujourd'hui, demain, avant)	adverbes, connecteurs et temps verbaux (Piero était parti quand la police est arrivée)		
	Séquences spatio-temporel non linéaires			distinction en fonction du nb de temps	
Style	Pas de mots au sens figuré				
	Utilisation de mots au sens figuré		courant (la nuit tombe en hiver)	peu courant (un manteau de neige)	
	Figures de style		Comparaison + métonimies et métaphores courantes	Personnifications, hyperboles et métaphores	Tout type de figures de style + Métaphores non courantes

		1) Très Facile	2) Facile	3) Accessible	4) +Complexe
Indicatif	Présent				
	Imparfait				
	Futur simple				
	Passé simple		3eme pers. sing		
	Passé composé				
	Plus que parfait				
	Futur antérieur (j'aurais marché)				
	Passé antérieur (j'eus marché)				
Subjonctif	Présent				
	Passé (que j'aie marché)				
	Imparfait (que je marchasse)				
Conditionnel	Présent		politesse + pers du sing		
	Passé				
Impératif	Présent				
	Passé (aie marché)				
Infinitif	Présent				
	Passé (avoir marché)				
Participe	Présent + gérondif (en) marchant	gérondif	gérondif	participe	
	Passé + gérondif (en) ayant marché				
	Futur proche, passé récent, présent continu				



APPENDIX IV - Complexity Levels Descriptions for Portuguese

Descritores de complexidade textual

1) Descrição dos níveis

Nível 1 – Muito fácil

Textos curtos e simples até cerca de 50 palavras, com o objetivo de desempenhar tarefas do cotidiano 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 abreviações, 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.: *permanença 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 pelo empresa.*), incluindo construções impessoais comuns e passivas de *-se* (ex.: *O trabalho faz-se bem.*).

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

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

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

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

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

2) Resultados do *Focus Group* (tarefa preparatória + grupos de discussão):

	Descritores		Muito fácil	Fácil	Claro	+ complexo
<i>tamanho</i>	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 "A reciclagem é quando transformamos o lixo em coisas novas e úteis.")				
	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, ementas, instruções, preços de produtos, comunicações pessoais, etc.).				
	11	O texto é utilizado em contextos profissionais ou de comunicação dos <i>media</i> (e.g., comunicação ou instrução profissional, notícias, anúncios, desporto).				
	12	O texto é utilizado em contextos de lazer (e.g., literatura e ficção, reportagens, artigos de opinião).				
	13	O texto circula em contextos artísticos (e.g., literatura, poesia).				
	14	O texto é empregue em contextos especializados (e.g., comunicação profissional formal, fóruns <i>online</i>).				
	15	O texto é utilizado em contextos científicos e académicos (e.g., ensaios, relatórios, artigos científicos).				

		Descritores (PT)	Muito fácil	Fácil	Claro	+ complexo
<i>diversidade conceitual</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 (PT)	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 (PT)	Muito fácil	Fácil	Claro	+ complexo
<i>vocabulário/ conjunto de léxico</i>	24	O texto não emprega abreviações.				
	25	O texto utiliza abreviações comuns após ter explicado o que significam (e.g., <i>l – litro, m – metro</i>).				
	26	O texto utiliza abreviações comuns (e.g., <i>m, km, g., l, Exmo.</i>).				
	27	O texto não utiliza siglas.				
	28	O texto apenas utiliza siglas e acrónimos conhecidos (e.g., <i>CP, IVA, RTP, SIC, TVI</i>).				
	29	O texto utiliza siglas e acrónimos não conhecidos juntamente com o nome a que se referem (e.g., <i>Organização das Nações Unidas (ONU), Instituto Público (IP)</i>).				
	30	O texto emprega todos os tipos de siglas (e.g., <i>ADN, UNESCO, EU, UNICEF</i>).				
	31	O texto utiliza palavras e expressões simples relacionadas com assuntos do quotidiano, fazendo uso de substantivos frequentes (e.g., <i>batatas, azeite, legumes</i>).				
32	O texto tem palavras e expressões frequentes que fazem referência a assuntos do dia a dia, mas também a outros âmbitos específicos ou profissionais (e.g., <i>segunda-feira, curso, ano escolar</i>).					

	33	O texto tem palavras e expressões um pouco menos conhecidas e frequentes que fazem referência a âmbitos específicos (e.g., <i>doença cardíaca, artéria, tratamento</i>).				
	34	O texto utiliza palavras e expressões de âmbitos específicos, jargão ou termos arcaicos (e.g., <i>enfarte do miocárdio, artéria coronária, dissecação espontânea, genoma</i>).				
	35	O texto não faz uso de estrangeirismos ou faz uso de estrangeirismos muito frequentes e conhecidos (e.g., <i>internet, donut</i>).				
	36	O texto tem poucos estrangeirismos comuns (e.g., <i>site, stress</i>).				
	37	O texto faz uso de estrangeirismos comuns frequentemente (e.g. <i>link, hobby, marketing, mail, gay</i>).				
	38	O texto contém palavras polissêmicas (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 polissêmicas (palavras que possuem mais do que um significado) ambíguas e/ou bastante frequentes (e.g., <i>base, experiência, leve, seca, aplicar, banco</i>).				

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

		Descritores (PT)	Muito fácil	Fácil	Claro	+ complexo
sintaxe	49	No texto, há mais frases curtas e diretas ou combinadas através de conjunções familiares (e.g., <i>A criança fugiu ontem. O pai trabalha na fábrica, mas a mãe fica em casa.</i>).				
	50	O texto combina frases através de conjunções coordenadas e subordinadas (e.g., <i>Os músicos estavam tocando, <u>mas</u> poucos clientes ouviam. <u>Quando</u> pararam de tocar, apenas um senhor se manifestou.</i>).				
	51	As frases no texto são frequentemente combinadas através de conjunções coordenadas e subordinadas e, por vezes, com conjunções menos comuns (e.g., <i><u>Apesar da</u> confusão provocada pelos animais à solta, a situação foi controlada <u>sem</u> os animais serem atropelados.</i>).				
	52	As frases no texto são longas e compostas por diversas outras frases combinadas através de conjunções coordenadas e subordinadas (e.g., <i><u>Dado que</u> grande parte deste livro fala na gestão das partes difíceis do trabalho a solo e analisa modos de as corrigir ou de as transcender, primeiro quero deixar bem claro que trabalhar sozinho é fantástico: é <u>por isso que</u> tantas pessoas o consideram uma opção.</i>).				
	53	O texto tem poucas orações relativas referentes ao sujeito da frase (e.g., <i>O senhor 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, <u>a quem os amigos tratam por «Kika»</u>, 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 <u>que abraçou a mãe</u> dirigiu-se à entrada; O rapaz <u>que a mãe abraçou</u> dirigiu-se à entrada.</i>).				
	57	No texto, há mais frases 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 (PT)	Muito fácil	Fácil	Claro	+ complexo
coesão textual: pronomes, anáforas e elipses	64	O texto tem nenhum ou poucos elementos omitidos (e.g., <i>O acidente causou dois feridos leves. Os bombeiros levaram os feridos para o hospital</i>).				
	65	O texto tem elementos repetidos em vez de pronomes (e.g., <i>O João não gosta da <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 <u>fenómeno</u> pode ser visto</i>).				

		Descritores (PT)	Muito fácil	Fácil	Claro	+ complexo
coesão textual: referências temporais	70	O texto apresenta uma sequência ou localizações temporais simples e lineares feitas por advérbios de tempo e conjunções como hoje, amanhã, antes de, depois de, quando, enquanto (e.g., <i>Os bombeiros chegaram depois da polícia</i>).				
	71	O texto tem uma sequência ou localizações temporais lineares, dadas através de advérbios de tempo ou conjunções como após, assim que, antes de e também por tempos verbais (e.g., <i>Ele <u>está</u> a caminho e a festa <u>já</u> começou. A banda começou a tocar <u>antes de</u> todos saírem</i>).				
	72	O texto inclui uma sequência ou localizações temporais não lineares (e.g., <i><u>Após</u> analisarem as respostas dos doentes que <u>tinham sofrido</u> um ataque cardíaco, os médicos <u>confirmaram</u> que a gordura é o alimento mais perigoso</i>).				
	73	O texto tem diversas sequências ou localizações temporais não lineares (e.g., <i><u>Será preciso fazer</u> alguma coisa aos cadáveres e, por isso, <u>é bom que haja gente que se debruce</u> sobre o tema. Eu <u>seria</u> incapaz, porque nem sequer <u>consigo ir</u> a funerais, como já <u>fui</u>, onde <u>repousam</u> caixões abertos com os indivíduos em descanso eterno</i>).				

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



APPENDIX V - Complexity Levels Descriptions for Spanish

NIVELES DE DIFICULTAD

De los resultados de la encuesta se desprenden las tendencias siguientes en lo relativo a la determinación de los niveles de dificultad:

Muy fácil:

1. Longitud: el texto tiene hasta 50 palabras
2. Morfología verbal: los verbos están en infinitivo.
3. Léxico: El texto se refiere a conceptos concretos o remite a actividades cotidianas; para ello, se utilizan palabras y expresiones frecuentes
4. Sintaxis: el texto presenta oraciones cortas con conectores comunes, mayoritariamente paratácticos.

Fácil:

1. Longitud: el texto tiene hasta 500 palabras
2. Morfología verbal: los verbos están conjugados en imperativo.
3. Morfología nominal: en el texto se utilizan palabras compuestas corrientes.
4. Léxico: El texto remite a tareas corrientes y también a algunas ideas abstractas; se utilizan pocos sentidos figurados, acrónimos, siglas y abreviaciones
5. Sintaxis: es fundamentalmente paratáctica; se utilizan construcciones con el pronombre "se" incluso pasivas reflejas o impersonales y oraciones de relativo complemento del sujeto.
6. Cohesión textual: en el texto hay pocos elementos omitidos y se repiten con frecuencia elementos en lugar de pronombres; incluso se omiten algunos pronombres.
7. Estilística: en el texto hay algunos casos de metáfora y metonimia.

Claro:

1. Morfología verbal: Los verbos están conjugados en subjuntivo (presente, pretérito perfecto o pluscuamperfecto) o en futuro compuesto
2. Morfología nominal: en el texto hay alguna palabra compuesta con significado no decodificable analíticamente o palabras derivadas con morfemas (prefijos y sufijos) corrientes.
3. Léxico: el texto remite a conceptos o ideas abstractos, a tareas nuevas o aborda temáticas variadas; se usan palabras polisémicas; se utilizan abreviaciones, siglas y acrónimos corrientes, así como palabras extranjeras corrientes.
4. Sintaxis: el texto presenta una estructura mayoritariamente paratáctica; no se sigue siempre el orden básico de palabras.
5. Cohesión: el texto presenta un uso frecuente de la pronominalización.
6. Coherencia textual: el texto presenta secuenciaciones cronológicas no lineales.

Más complejo:

1. Morfología verbal: en el texto hay verbos en imperfecto de subjuntivo.
2. Morfología nominal: en el texto se utiliza la derivación con morfema (prefijos y sufijos) variados y poco corrientes
3. Léxico: el texto se utiliza en ámbitos específicos y remite a conceptos nuevos con un vocabulario especializado (textos académicos). Se utilizan todos tipo de acrónimos y siglas.
4. Sintaxis: en el texto se recurre con frecuencia a oraciones largas y a la hipotaxis.
5. Estilística: se recurre con frecuencia a figuras de estilo: comparación, hipérbole, sentidos figurados poco habituales.