



UNIVERSIDAD METROPOLITANA DE CIENCIAS DE LA EDUCACION
DEPARTAMENTO DE INGLÉS

A TEXT GRAMMAR-BASED APPROACH FOR TEXT COMPLEXITY MEASUREMENT
ACROSS CEFR LEVELS

TESIS PARA OPTAR AL GRADO DE MAGISTER EN ENSEÑANZA-APRENDIZAJE DEL
INGLÉS COMO LENGUA EXTRANJERA

AUTOR: GABRIEL MORALES MORALES

PROFESOR GUÍA: PABLO CORVALÁN REYES

SANTIAGO DE CHILE, SEPTIEMBRE DE 2021

2021, Gabriel Morales Morales

Se autoriza la reproducción total o parcial de este material, con fines académicos, por cualquier medio o procedimiento, siempre que se haga la referencia bibliográfica que acredite el presente trabajo y su autor.

Acknowledgements

Este trabajo es mi forma de agradecer a todas y todos los maestros que me han guiado por este proceso, tanto quienes son maestros en el corazón, como quienes lo son en el aula; tanto quienes se están formando, como quienes ya salieron del sistema; tanto a quienes aún no les convence esta bonita forma de vida como quienes viven por y para esto. Gracias a las maestras de mi corazón: Mami, Mamá, Tía Mallita, Muchacha; al consejo docente supremo familiar: Tío Carlos y Tía Erika, Mónica y Axa; a mis mentores: Profe Carlos, Miss Coty, Profe Rosa; a mis colegas y amigos: Rosita, Ayleen, Rubén, Alexis, Marquitos, Erick y Gianfranco. Gracias por ser y por estar en este camino.

También quiero agradecer especialmente al Profesor Pablo Corvalán, por confiar en mi e instarme a pensar en grande, a hacer un trabajo riguroso y por mostrarme que el legado de nuestros maestros sobrevive a través de nosotros.

Gabriel Jesús Morales Morales

Table of Contents

1. Introduction	1
2. Literature review	3
2.1.Foreign language learning/acquisition and input processing.....	3
2.2.Reading comprehension.....	6
2.3.Text complexity, readability and modification.....	8
2.4.Text grammar.....	10
3. Rationale and scope	12
3.1.Research questions.....	13
4. Methodology	14
4.1.Instruments.....	14
4.2.Participants.....	16
4.3.Procedure	17
5. Analysis of results	19
5.1.Expert instrument quantitative results and analysis.....	19
5.2.Expert instrument qualitative results and analysis.....	21
5.3.Student instrument quantitative results and analysis	26
5.4.Student instrument qualitative results and analysis	28
6. Conclusion	34
7. References	37
8. Appendixes	44
8.1.Multi-level text with TCi	44

8.2. Answers and labels for fragment 1 (experts' instrument).....	46
8.3. Answers and labels for fragment 2 (experts' instrument).....	47
8.4. Answers and labels for fragment 3 (experts' instrument).....	48
8.5. Answers and labels for fragment 4 (experts' instrument).....	49
8.6. Answers and labels for fragment 1 (students' instrument).....	50
8.7. Answers and labels for fragment 2 (students' instrument).....	51
8.8. Answers and labels for fragment 3 (students' instrument).....	52
8.9. Answers and labels for fragment 4 (students' instrument).....	53

Abstract

This research aims to present a novel approach to text complexity and readability measurement. The study tested 4 first-year undergraduate students, divided into 4 groups based on their linguistic performance according to the Common European Framework of Reference, and 9 Teachers of English as a Foreign Language who meet the characteristics to be labelled as experts. The data was collected through a series of exercises based on an adapted variable-complexity text, such as grammaticality judgements and comprehension questions, plus a retrospective report which allowed to collect quantitative and qualitative information. Based on the results obtained, this approach to text complexity is partially validated, as it correlates with different CEFR levels, as well as with the criteria observed by the expert teachers at the moment of labelling materials due to their complexity. Results also allowed to identify that vocabulary and syntax are the sub-aspects of complexity that are more predominant for students at different levels, and vocabulary and reference are the most observed aspects by teachers. The final objective of this research is to provide EFL teachers with a comprehensive framework which facilitates the measurement of texts' complexity, which will help teachers provide their students with readings which meet their language needs.

KEY WORDS: text complexity, text grammar, text adaptation, reading comprehension, CEFR

1. Introduction

Text selection, adaptation and creation are significant challenges for English as a Foreign Language (EFL) teachers. The various theoretical stands towards language learning, unclear measurements for readability and the different elements contained in these measurements force teachers to modify texts from their perspective and knowledge, or trust in text editors who offer levelled readings according to age-group or linguistic competence level (Pearson & Hiebert, 2013). Nevertheless, despite the considerable number of proposals for material adaptation at different educational levels, these have not been able to achieve their objective (Oh, 2001). This failure resides because available adaptations focus on atomic elements of language inside the text, which hinder student processing of information and do not aid them achieve successful reading comprehension.

The following research addresses the issue of text complexity from the perspective of text grammar, a holistic approach towards texts as contextually bound, purposeful units of meaning. From this perspective, the research aims to establish the main criteria for measuring the complexity of a text, with four key aspects to be considered: lexical density, syntactical complexity, potential reference instances and potential unit of information display. These aspects serve as constituents of what is introduced as the text complexity index, a measurement of the interaction of old and new information in a text, which addresses learners' knowledge at the atomic, lexical-grammatical level, and also at the holistic, information processing level.

The purpose of this research is to explore an alternative view of complexity, which includes grammatical, lexical, semantic and pragmatic information at the text level. As a means to achieve this purpose, this research aims to measure teachers' material judgement and students' reading comprehension of an increasing-difficulty text from a quantitative and qualitative perspective. Moreover, a secondary objective of the proposal is to assess how the index and its sub-indexes interact with different stages of development in relation to the Common European Framework of Reference (CEFR) as students were previously diagnosed through an evaluation aligned with this international referent.

Results obtained from this study indicate that both teachers and students make use of some of the components of the proposed approach, and that more comprehensive use of components

increases reading comprehension scores in students, and complexity identification accuracy in teachers. Moreover, there is a degree of correlation between CEFR levels and the components of the framework used by learners, which allows to partially validate text complexity index as a functioning tool for measuring texts' features.

2. Literature review

As a means to study text complexity and the impact it may have on readers at different performance levels, it is necessary to establish a theoretical perspective. Regarding this research project, language acquisition will be addressed from a *cognitive perspective*, namely, the processing of input from different sources and with different levels and types of complexities. Moreover, the implementation context will be defined as a *foreign language acquisition (FLA)* context, and its features will be described. As a means to align this proposal with its pedagogical aims, *texts* will be defined as the essential materials for reading comprehension tasks; these will be seen from the perspective of text grammar, namely, as contextually and purposeful units. Finally, *complexity* will be understood as one of the features of texts, which is usually affected by adaptation and modification processes for pedagogical purposes.

2.1. Foreign language learning/acquisition and input processing

The process of developing any linguistic system different from the first language (L1) of an individual is labelled as second language acquisition and/or learning. This area has been a significant subject of research for linguists, psychologists, neuroscientists and educators, among others. Most authors agree on the fact that there is still not a final, comprehensive answer regarding how second languages are learned (Mitchell, Myles and Marsden, 2013). Nevertheless, this multidisciplinary research field has come to various theories, models and hypotheses that shed light on possible explanations for the phenomena that have been observed in the process of trying to have a clearer understanding of how language acquisition/learning occurs.

One of the contemporary theories for L2 learning/acquisition has its origins in psychology and tries to explain it from a cognitive perspective. According to this theory, language acquisition is similar to learning any complex skill; thus it requires mental processes such as input perception, organization and re-organization of mental representations and retrieval for output production (Saville-Troike, 2012). From this perspective, humans have limited resources for learning, such as attention and memory; in this manner, it is necessary to practice the comprehension and use of simpler language (namely vocabulary, structures, etc.) so as to achieve automatization of it, which in turns allows addressing tasks which require the use (comprehension and/or production) of complex language. Van Patten's Input Processing theory (Van Patten, 2002; 2004; 2013) addresses the process of connecting form and meaning as a critical aspect required for acquisition to take

place. IP follows a series of principles that underlie critical statements about how learners' brains process language when they encounter it. Some of these principles are the predominance of content over function words, the relevance of sentence structure and identifying subjects, and the importance of contextual information for making sense of what is being expressed. These principles allow researchers and educators to identify those elements that must be highlighted in the input presented to students in order to boost comprehension, and provide the necessary conditions for acquisition.

Van Patten's theory can be connected with Pienneman's Processability theory (Pienemann, 2005; 2010; 2015), which explains that acquisition occurs through the processing of information at increasing levels of difficulty, or developmental trajectory. According to this theory, learners are constantly exchanging information between elements they can process, increasing in its complexity through at least five different stages, which are exemplified in Figure 1 (Yamaguchi & Usami, 2017, p. 225), which start from the lemma level and end at the sentence level. Each of these levels may represent a barrier, if not mastered to be able to process successfully the next level. Thus, researchers and educators must focus on their students' learning stage if they want acquisition to be successful since stages cannot be skipped (Lightbown and Spada, 2013). Based on the interaction between these two theories, it is possible to ascertain that language (in this case input) needs to be adapted depending on the acquisition stage and the processing capacity of learners' linguistic mechanisms as a means to assure an effective acquisition process.

Table 1.*Processability Theory stages for English L2 morphology.*

Stage / Procedure	Morphological outcome	Structure	Example
4. Sentence procedure	Interphrasal morphology	3rd person singular -s	<i>Peter loves rice</i>
3. Phrasal procedure	VP morphology	AUX + V: have + V -ed MODAL + V be + V -ing	<i>they have walked</i> <i>you can go</i> <i>I am going</i>
		NP morphology	phrasal plural -s (with quantifiers) <i>these girls</i> <i>many dogs</i> <i>three cats</i>
2. Category procedure	Lexical morphology	past -ed verb -ing (without AUX) plural -s (without quantifiers)	<i>Mary jumped</i> <i>he working</i> <i>I miss my friends</i>
1. Lemma access	Invariant forms	Invariant forms	<i>dog, girl</i>

Note. From “Plural Marking in Spoken and Written Narratives: A Corpus-based Study of Japanese Learners of English” by Y. Yamaguchi & H. Usami, 2017, *International Journal of Applied Linguistics & English Literature*, 6 (5), p. 225 (<http://dx.doi.org/10.7575/aiac.ijalel.v.6n.5p.224>). Copyright 2017 by Australian International Academic Centre.

Apart from the relevance of information processing for language learning/acquisition, the learning and acquisition dichotomy has caused debate among researchers regarding the specific characteristics of these processes, and how to address them in the most effective manner (Ellis, 2015). This discussion has been usually connected with the contrast between second and foreign language contexts (SL and FL respectively), as acquisition has mostly been associated with SL, while learning has been connected to FL. According to authors like Donato (2000); Ortega (2008) and Cook (2016), among others, there are significant differences between FL learning and SL acquisition, which have an impact on the methods selected for teaching the L2. Second language

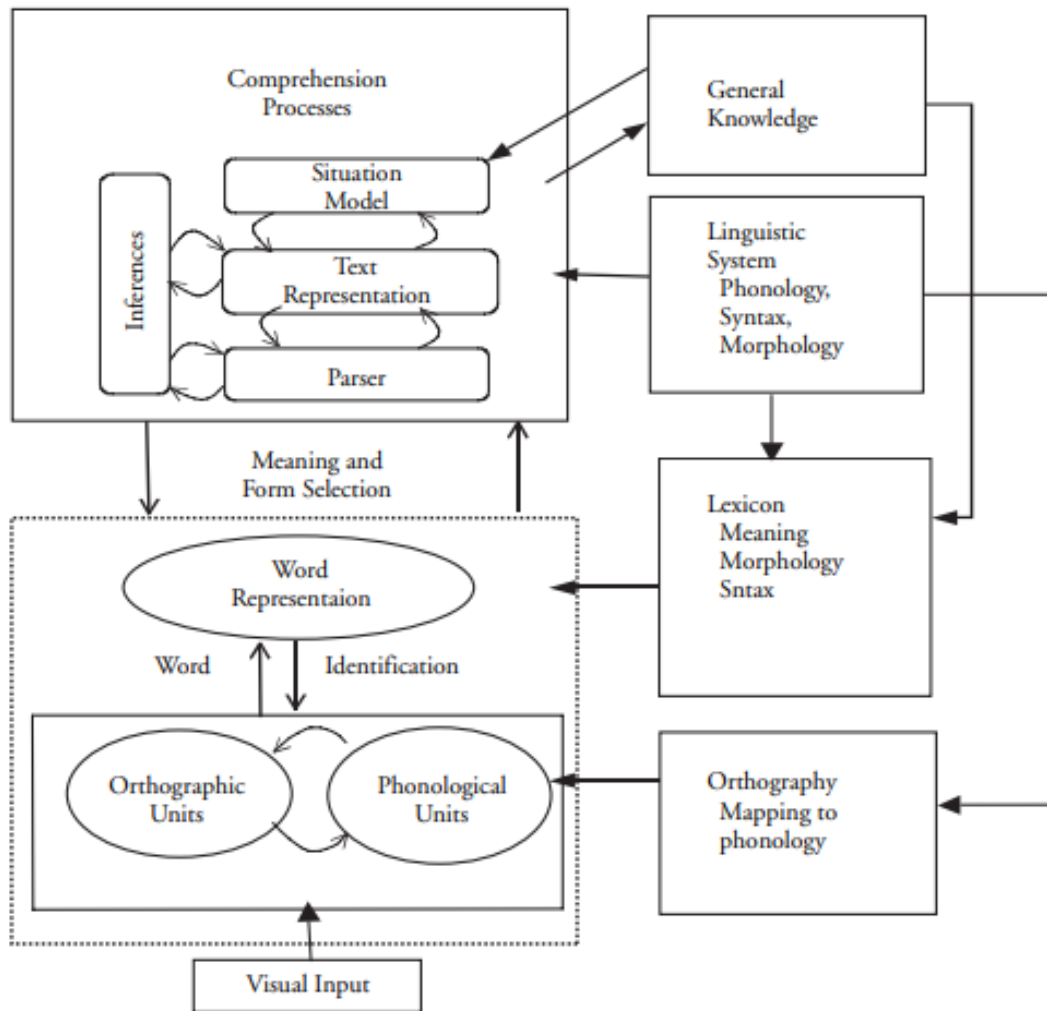
acquisition (SLA) occurs in those contexts where there are opportunities for the use of L2 in various communicative situations (day to day conversation, the media, etc.), facilitating learning through exposure or acquisition. Foreign Language Learning, on the other hand, occurs in those contexts where there are limited or no opportunities for communication in the L2, thus learning the target language through explicit instruction is necessary. In spite of the fact that these two processes and contexts seem opposite, learning and acquisition are separate processes that can take place in both contexts. Thus, it is possible to subconsciously acquire traits from L2 communication instances in English as a foreign language (EFL) contexts, as well as it is possible to consciously learn language aspects in English as a Second Language (ESL) contexts (Ellis, 2015). As a means to address the context where the present study was implemented, learning and acquisition will be referred to through acquisition as an umbrella concept, as this work adheres to the thought that both processes co-occur in determined contexts. In terms of the study setting, this will be defined as an FLA context, since English in Chile is not used as a primary nor secondary means of communication, and most of the language exposure occurs during classes.

2.2. Reading comprehension

Reading comprehension has been an area of major interest in SLA, as it involves an individual and challenging task for most learners (Alderson, 1984). It is an activity which merges strategies, attentional resources, skills and knowledge (Grabe & Stoller, 2011) into a series of processes that involve not only comprehension, but also learning, evaluation, interaction, linguistic analysis, among others. Reading comprehension, understood as building a mental representation of a text message, requires to make use of limited processing capacity in order to make sense of what is being read (Reed & Kershaw-Herrera, 2016), thus working memory is demanded to carry lower-level processes, such as access to lexical meaning, parsing of syntactic structures and organization into semantic propositions, as well as higher level processes, such as the use of background knowledge, formulating a situation model, among others, shown in Figure 2. The consistent interplay between new and previously known information in reading comprehension, and the amount of elements that play a role in the process require the individual to carry out these tasks in a short period of time (Pollastek & Treiman, 2015) as working memory is not only limited in terms of load, but also in terms of retention.

Figure 2.

The components of reading comprehension from identifying words to comprehending texts.



Note. From “The acquisition of Reading Comprehension Skill” by C. Perfetti, N. Landi and J. Oakkill, 2005, *The Science of Reading: A Handbook*, p. 227. Copyright 2005 by Blackwell Publishing Ltd.

Successful comprehensive reading needs to be developed, and is defined by authors like Papola-Ellis (2014) and Williamson et al. (2013) as an accurate, fluent and comprehensive experience, which requires texts to be paired to the needs of the readers, and levelled according to their increasing levels of language development. The idea is that learners constantly face materials which are challenging enough to them to use different strategies and mechanisms that aid in their

skill development process. For this reason, text complexity and readability have become a major concern for curriculum and materials developers, as they expect to provide teachers with specifications that will boost the acquisition of reading comprehension in the L1 for literacy purposes, which if well developed, can boost the process of L2 reading comprehension through the transfer of key sub-skills and mechanisms (Proctor et al, 2010).

2.3.Text complexity, readability and modification

The insight obtained from cognitive information processing-related theories diverted researchers' attention from lexical accounts of text complexity to text structures at the micro and macro levels, to predict success in the processing of specific texts (Pearson & Hiebert, 2013; Wray & Janan, 2013; Dascalu, 2014) and maximize text comprehension in language acquisition processes, whether first or second. From this change of paradigm, two currents of thought appeared. On the one hand, there were educators focused on teaching comprehension strategies to their students to boost their performance, such as Rosenblatt's (2005) experience/transaction approach. On the other hand, there were educators interested in the constituents of the text, and how to adapt them to make them more readable for their students, such as Merlini-Barbaresi's (2003) textcentric approach. Authors such as DuBay (2007); Williamson et al (2013); Janan (2011); and Sanden (2014) among others, state that it is necessary to have a balance between individual performance, which is expressed through prior knowledge, reading skill, interest and motivation; and textual quality, which is present in the content, style, design, and organisation of the text. Based on these two groups of aspects, the author states that by the end of the 80s, there were two hundred formulas, such as Flesch Reading Ease (Flesch, 1948), Fog Index (Gunning, 1952) and SMOG Grading (McLaughlin, 1969). and over a thousand studies published which addressed readability from different perspectives. None of these were able to ascertain one measurement that could predict student failure or success before engaging a reading comprehension task.

Given the high criticism towards readability formulas, educators decided to address their reading comprehension tasks through modified texts, or directly creating their own texts (Vajjala & Meurers, 2012) which usually requires a significant amount of effort and time. Authors like Oh (2001), Spiaggiari et al (2005), O'Donnell (2009) and Temnikova (2012) define modification as the changes made to texts with the aim of achieving higher levels of readability, comprehensibility

and/or translatability. Modification is divided into two categories: simplification is the process by which complex and lengthy vocabulary and structures are replaced by more straightforward and shorter elements, thus reducing the difficulty of the formal elements in the text. Elaboration is the process by which complex thematic relations are aided by the inclusion of signals which boost student's opportunities for using peripheral information to achieve comprehension. After a comprehensive analysis of previous research, authors state that both processes are not entirely correlated with students' needs at different levels of linguistic development, an example of this is that simplification promoted fossilisation in entry-level learners of English. At the same time, elaboration did not prove a statistically significant difference in contrast with unmodified texts in terms of student comprehension rates. According to authors such as Pearson & Hiebert (2013), this happens because these processes rely mainly on teacher assumptions about students' potential difficulties and expert knowledge, which has not been systematised nor tested by empirical means.

The solution to the problem of expert knowledge in text adaptation appeared in the early nineties, with the concept of L2 complexity (Bulté & Housen, 2012). The focus on the linguistic aspects of the text lies on the need of analysing only those aspects which are not highly variable, labelling them as absolute complexity, while reader-related characteristics were labelled as relative complexities (Merlini 2014; Miestano, 2006; 2009). These complexities are independent but can be mutually correlated. This approach serves as a framework for empirical and systematic evaluation of lexical and grammatical density, diversity and sophistication. Each of these constructs is related to particular formulae, which allows rationalizing these measurements and contrasting them with other instances. These formulas allow to account for the knowledge that is necessary for learners to master in order to comprehend a particular text (Sanden, 2014); thus, different approaches have linked various linguistic aspects in elaborate formulas which automatically calculate, using Natural Language Processing (NLP)-based models, how difficult a text is. Some examples for this formulae are Lexile Framework (Stenner et al., 1989), Coh-Metrix (Graesser et al., 2004) and TextEvaluator (Sheehan et al., 2013), among others. Nevertheless, the analysis through these means is still a matter of debate, as each of them has a strong component of whether lexical complexity (Flor et al., 2013), syntactic complexity (Frantz et al., 2015) or cohesion/coherence (Dascalu, 2014), which leads to non-conclusive results on the effectivity of these tools. Finally, the use of these tools requires advanced linguistic and computer-science knowledge, depending on the approach it follows (manual, semi-automatic or automatic)

(Temnikova, 2012) and does not provide clear guidelines about how to carry out text modification based on these models, in spite of the fact of their various applications.

2.4. Text Grammar

Opposite to the formal complexity approach to L2 performance is text grammar (Van Dijk, 1977; De Beaugrande & Dressler, 1981; Corvalán & Zenteno, 2009). This approach to grammar focuses on the communicative and contextual aspect of units intra and inter sentence in a text (Corvalán & Zenteno, 2009). From this model, macro-structures are in charge of modelling text units, giving the internal sentential structure a sense of dependence on these macro-structures. As this analytical approach focuses on the level of text, it provides with a comprehensive account of language use for productive and receptive perspectives, considering oral and written samples from various lengths and purposes. Moreover, it considers pragmatic aspects, which are left aside by other models, such as intentionality (purpose), acceptability (reception) and situationality (coherence with the context), among others. All these elements were intentionally taken into account for the model, as it was designed over theoretical-practical knowledge for improving teaching-learning processes of English as an L2.

The main components of Corvalán and Zenteno's text grammar are cohesion, attitudinal markers, pragmatic markers, information structures and discourse-bound grammatical realisations. Each aspect has its purpose and units, and interacts with one another at different levels to accomplish text unity. For the purpose of this study, two concepts will be defined more specifically; the first of them will be *information units* (IUs), defined by Lambrecht (In Corvalán, 2007) as "the functions that sentences in the text may perform" (p. 11) these information units will portray how the coding of information transits in a continuum, from strong to weak, and is represented by certain specific grammatical constructions which will depend on the co(n)text in which they are used. These units are the pragmatic realisations/interpretations of grammatical forms in a text, which are the lexico-grammatical realisation of semantic units, or propositions. Thus, meaning is shaped into sentences, and these will be structurally bound to the contextual factors which affect its production and interpretation. Units of information, according to Lambrecht, can be catalogued into four different categories, depending on the relation between referents, propositions and degree of activation: topic/comment (which answers to "What did X do?"), identificational (which answers to "Who did X?"), presentational (which answers to "What

happened?”), and scene-setting (which provides a context for the other IUs). Each of these IUs is associated with specific sentence structures, overlapping in some cases the correspondence between structure and IU type.

The second concept to be addressed for the purpose of this article will be *cohesion*, which is defined by Halliday & Hassan (1976) and Corvalán (2007) as “the set of features that contribute to the unity of the text” (p. 19) cohesion in a text is achieved through the use of cohesive devices, which are; reference, substitution, conjunction and lexical cohesion. This proposal will consider only reference as a practical aspect to be addressed due to the magnitude that would imply including all the other devices in this research project. Reference corresponds to the relational connection between an entity and the expressions that can be used to maintain it active within the text, referring expressions can be grouped into personal (i.e. expressions related to human referents), demonstrative (i.e. expressions related to previously activated referents), and comparative (i.e. expressions related to two elements and their relationships in a text). This proposal will only consider personal and demonstrative references.

The various elements of linguistic and extra-linguistic aspects considered by this model of text grammar provide a perspective that escapes from previous approaches to text complexity since it allows for a comprehensive approach to texts, as it combines grammatical, semantic and pragmatic information and relations present within texts. Its relation with contextual information and thematic progression articulate positively with the cognitive theories of Input Processing and Processability Theory, which is why this approach was selected as the theoretical basis for the text complexity index to be proposed.

3. Rationale and scope

Despite the large number of text complexity indexes and systems available in the literature, the adaptation of reading materials from the results obtained by those systems is an intricate process. In most cases, this leads teachers to rely on the expert system (Hiebert & Pearson, 2014) or search for pre-levelled readings, which may be significantly biased based on the authors' context, intentions and perceptions. Teacher choices regarding material selection and adaptation have a considerable impact on students' performance in the English classroom, affecting their motivation and perception of complexity towards the task and the subject. From this perspective, authors such as Kumaravadivelu (2007) recommend teachers to first create their own materials, then to adapt them, and finally rely on authentic materials, to provide a transition which would be context-sensitive and will allow students' construction of meaning.

This proposal aims to test a novel approach towards measuring the complexity of texts, which considers not only the lexical aspect but also the text dimension, which accounts for syntax, semantics and pragmatics. This approach was tested, first, regarding teachers' expert judgement about text complexity, and second, across CEFR levels, based on an international, adaptive and highly accurate (Pollitt, 2018) placement test, such as Oxford Placement Test. As a means to achieve a high level of validity for this measurement, two research questions (RQ) were formulated, one per each group of participants. These research questions are sub-divided into sub-questions that aim to address the quantitative and the qualitative nature of the data collected by the instruments formulated for this purpose. The second group of sub-questions, which addresses the group of students, considers a question that aims at the reading comprehension aspect of this instrument, which corresponds to the real-life application of this framework.

3.1. Research questions

The primary and secondary research questions this research proposal aims to answer are stated below:

1. Is there any correlation between the text complexity approach proposed and teachers' perceptions of complex texts?
 - 1.1. How relatable are teachers' judgements of text complexity and texts framed in this approach??
 - 1.2. Which of the aspects in this complexity approach are more relevant on teachers' perspective at the time of identifying a complex text?
2. Is there any correlation between text complexity and student comprehension of adapted texts?
 - 2.1. How successfully do students from different CEFR levels comprehend a text with different levels of text complexity?
 - 2.2. Which of the aspects in this complexity approach are more relevant on students' reading comprehension?
 - 2.3. How do the four aspects proposed for text complexity interact with participants' reading comprehension across CEFR levels?

4. Methodology

This proposal corresponds to a mixed-methods quasi-experimental design, where a group of intentionally selected individuals will be tested with an instrument which combines tasks that can obtain numerical and textual information. data will be analysed separately, numerical quantitative information will be processed using IBM SPSS software, looking for correlations. Textual information, on the other hand, will be processed using ATLAS.ti software, to label and look for common information among testimonies.

4.1. Instruments

The concept of text complexity is introduced through the instrument designed for this research. The instrument to be used corresponds to an adapted text with variable levels of text complexity, this means that the text (Appendix 1) contains different levels of complexity in each paragraph, levels of complexity do not follow any type of organization pattern, as they were randomized using an online tool designed for this purpose (www.randomizer.org). The text presents comprehension checking questions, which work as acceptability judgement exercises at the end of each paragraph since levels of complexity vary in each paragraph. The topic of the text is religious beliefs in India, as it is a topic which is not of immediate access to most participants, as a means to reduce variables such as background knowledge and motivation towards the topic. The format of the instrument corresponds to a Google Forms form, as it facilitates information collection and processing.

As a means to provide a clear definition of what will be understood by text complexity, this will consist of the addition of four aspects. *Lexical diversity* is the first of them; this is represented by a type-token ratio measurement and calculated by dividing the number of unique words (or types) by the square root of two times the number of words (tokens) in the text, as seen in (1). This measurement is discussed by authors such as Kuiken and Vedder (2007), Nation and Webb (2011), and Vajjala & Meurers (2012), among others. This formula corresponds to an adapted version of the original formula, which corrects variability of results caused by texts of different lengths.

$$\text{Word Types}/\sqrt{2\text{Word tokens}}=\text{Lexical diversity} \quad (1)$$

The second aspect to be considered in the text will correspond to the *text-reference index*, proposed by the author, based on observations made by authors such as Crossley et al (2007), Sheehan, Kostin & Futagi (2008) and Dascalu (2014), among others. This corresponds to the number of potential references (referents activated by the use of referring expressions) throughout the text, which is calculated by dividing the multiplication of referring expressions times the types of reference present in the text by the number of referents present, as seen in (2).

$$(\text{Referring expressions} * \text{Types of Reference}) / \text{Referents} = \text{Reference index} \quad (2)$$

The third measurement corresponds to the *text-information structure index*, which addresses the potential number of information structures throughout a text. This index is based on the works by Mesmer, Cunningham & Hiebert (2012); Sheehan, Flor & Napolitano (2013), among others, and it is obtained by dividing the number of unique grammatical structures times the total number of sentences in the text, which is then multiplied times two eights, as it corresponds to the proportion of sentence types that approximately correspond to the four information structure types, as expressed in (3).

$$(\text{Sentence types} / \text{number of sentences}) * 2/8 \quad (3)$$

The fourth and final measurement corresponds to syntactic complexity, which is currently addressed from different perspectives depending on the author, as it has been observed in analyses such as Ravid, Dromi & Kotler (2010); Frantz, Starr & Bailey (2015); among others. In spite of the fact that there are more accurate measurements, since syntax is not the central feature of this measurement, it will be calculated using the total amount of clauses in the text, divided the total amount of sentences.

$$\text{Number of clauses} / \text{number of sentences} \quad (4)$$

These formulas will allow the processing and adaptation of the different aspects previously described into different amounts of significant complexity from the perspective of the index proposed, as a final operation, the numbers obtained per each of these measurements need to be added, so as to obtain a final number, which will be referred to from this point onwards as *text complexity index* (TCi). Each paragraph in the text has a different variance, which oscillates between 0.1 and 0.5 and are distributed in an uneven manner. Paragraph distribution in terms of complexity index is described in Table 1.

Table 2.

Complexity index paragraph distribution

Paragraph	Text complexity index (TCi)
1	2.0
2	2.5
3	1.8
4	1.7

Note. TCi is the addition of the previously described indexes per each paragraph.

4.2.Participants

Participants in this research project correspond to two groups, the first of them corresponds to a group of 9 teachers of English as a Foreign Language, which will be labelled as experts. According to Goodwyn (2011), the concept of expert is highly controversial and is composed of many factors, which have led to various models of competence, some of which differ in terms of the aspects to consider. As a means to characterize and validate the label of expert for the purpose of this study, teachers selected for the study accomplish the following conditions, retrieved from Palmer et al. (2005):

1. Teachers must be a trained Teacher of English as a Second/Foreign language.
2. Teachers must have taught English for more than 5 years.
3. Teachers must have a C1 (or superior) command of English.

The conditions previously stated account for the concepts of training, experience and competence. Each of these concepts on its own is not enough to label a teacher as an expert; thus, they must be addressed together with norm-based (to stand out within a group of others) and criterion-based (to stand out frequently and overtime) performances (Arkbari & Yazdanmehr, 2014). As a means to account for these two aspects, the group of teachers participating in this study was reduced to teachers of the English Institutional Programme at Universidad de O'Higgins, as this programme has been evaluating their teachers with a single instrument over the last five years.

After obtaining the results from Evaluación Docente (applied twice throughout each semester), it was possible to establish two more conditions for participants:

1. Teachers must have a 6.0 or superior evaluation in Evaluación Docente.
2. Teachers' results must have been consistent over 6.0 in the last 3 years.

The second group of participants is composed of 4 students from undergraduate programmes at Universidad de O'Higgins. These will be selected based on their results in the Oxford placement test (OPT), a standardised placement instrument designed, validated, and provided by Oxford University Press. This test measures individual use of language and listening skills in English and provides results aligned with the Common European Framework of Reference (CEFR) (Purpura, 2007). Participants were selected and distributed based on their linguistic competence, covering levels A1 to B2.

The criteria to select these participants corresponds firstly to their age, which varies between 18 and 19 years old since it corresponds to the natural transition from secondary to university education. This criterion ensured that students had not been exposed to university-level English lessons, which focus mainly on the development of skill-related strategies. Secondly, participants were selected based on their results in the use of the language aspect of OPT. This allows to accurately select students based on their performance regarding vocabulary, grammatical knowledge, and reading comprehension.

Every teacher in the programme who met this conditions was invited to participate in the study, as well as students who were within the description of characteristics. To assure the study's reliability, both teachers and students were asked about their prior knowledge about India and its religion. Thus familiarity and motivation towards the topic of the text would not affect their performance in the application of the instrument.

4.3.Procedure

Information was collected by two different procedures, each of which allowed qualitative and quantitative data to be obtained. Participants were asked to read the text while their reactions were being recorded. Instructions were to stop at the end of each paragraph and answer the

comprehension checking questions, which served as acceptability-judgement exercises, to test participant comprehension of the information contained in each paragraph. Furthermore, participants carried out a retrospective analysis of each fragment, expressing all their reactions, and more specifically, when they face an excerpt in the text, which is difficult for them, or where they were not able to understand an idea. This task served as a truth-value judgement exercise, which allowed to point out which of the elements considered as components of text complexity cause disruption in terms of reading comprehension.

After all the groups interacted with the text, the information was processed separately. For students, scores in comprehension questions were correlated with participants' scores in the Oxford Placement Test. Information obtained from this analysis shed light on RQ 2.1, as text-related scores were contrasted with general score ranges. Regarding teachers, overall and specific accuracies were measured so as to detect how often the metrics matched the criteria of the expert teachers, in terms of which fragment is more complex than the others, thus answering RQ 1.1.

Regarding qualitative information, this was processed and labelled for each participant using ATLAS.ti software. Regarding students, labels were contrasted with CEFR level-groups across participants, and they were analysed inside each level based on the four aspects of this text complexity measurement. Results from this analysis related to RQs 2.2. and 2.3, since the co-occurrence of labels allows the identification of aspects in text complexity, which cause significant difficulty for participants' comprehension in each CEFR level. Regarding teachers, labels obtained were checked for co-occurrence across fragments of the text, so as to identify the most notorious aspects for each complexity level. This information shed light on RQ 1.2.

As a final step, and as a means to be able to answer to RQ 1 and 2, quantitative and qualitative analyses were interpreted in order to validate this approach to text complexity from both students' and teachers' perspectives, crossing information from both qualitative and quantitative analyses.

5. Analysis of Results

In this section, significant discoveries made from the analysis of the results of both questionnaires are analysed, quantitative and qualitative results are presented and discussed first from the teacher-oriented instrument and posteriorly, from the student-oriented instrument. Moreover, these discussions will aim at answering the research questions, and sub-questions (RQs) stated in the rationale and scope section of the study.

5.1. Expert instrument quantitative results and analysis

Regarding the answers provided by the 9 expert teachers who read the text, their results show high levels of inaccuracy at the time of distributing the levels of textual complexity in the different fragments. With an average of 28%, only two participants achieved an accuracy equal to or over 50%, and one participant catalogued correctly 100% of the fragments. Moreover, four participants achieved a 25% and three participants scored a 0% in this task, as can be observed in Table 3.

Table 3.

Complexity measurement accuracy per participant.

	Fragment 1	Fragment 2	Fragment 3	Fragment 4	Accuracy
P1	4	3	1	2	0%
P2	3	1	4	1	50%
P3	4	2	3	1	25%
P4	2	3	1	4	0%
P5	2	3	4	1	25%
P6	1	3	4	2	25%
P7	3	2	4	1	25%
P8	3	4	2	1	100%
P9	4	3	1	2	0%
Average					28%

Note. The correct order (read from left to right) for the degree of textual complexity is 3, 4, 2, 1.

Some relevant tendencies to observe are that two of the participants who obtained a 0% of accuracy identified the paragraphs in the same manner, this is by switching the position of the fragments that had the highest and lowest complexity indexes, which may mean that participants

had difficulties to identify subtle variation in terms of complexity, and would require higher index difference so as to make accurate assertions regarding fragments. Something similar happens with participants who obtained a 25% accuracy, who were able to label in most of the cases the fragments with the highest or the lowest complexity indexes. Based on these preliminary results, it is possible to theorize that complexity indexes superior to 0.8 generate a difference for experts at the time of labelling fragments of text according to their perception of complexity. Along the same line, the participant who labelled correctly 50% of the texts was able to identify the least complex fragment accurately, giving this label to two fragments. However, this participant did not follow the instructions for the task correctly, and it was pertinent to add their results as it supports the idea that experts have a more accurate perception of highly complex fragments, as well as their opposites, but fail to identify subtle complexity increases or decreases.

Regarding a fragment by fragment accuracy analysis, the stretch with the highest score was fragment 4, which had the lowest index for complexity, and was correctly labelled 55% of the time (4 out of 9 participants). The fragments with the lowest score for accuracy were fragments 2 and 3, correctly labelled 11% of the time. It is possible to observe from these results that, despite the considerable difference between fragments with lower and higher scores for accuracy, the latter are statistically not significant, as they correspond to nearly half of the participants tested only, as it can be observed in Table 4.

Table 4.

Complexity measurement accuracy per fragment.

	Fragment 1 (TCi 2.0)	Fragment 2 (TCi 2.5)	Fragment 3 (TCi 1.8)	Fragment 4 (TCi 1.7)
Correct	3	1	1	5
Incorrect	6	8	8	4
%	33%	11%	11%	55%

Note. These results were obtained from a sample of 9 participants

Similar to the previous results analysed, both the fragment with the highest and with the lowest indexes are the ones that the participants most frequently recognised. This serves as a support for the previous idea about sharper differences among fragments as a means to allow more accurate

identification being necessary. Nevertheless, the second-most accurate fragment identified was not the text with the highest TCi, but the second-highest, in spite of the fact that these two fragments had the most significant difference in the sample. It is possible to argue on the one hand that probably participants are not sensitive to texts above certain TCi; and on the other hand, that a higher TCi on its own is not enough to facilitate participants' identification of complex texts. Another observation that would allow shedding light on the difficulties faced by participants is that fragments 3 and 4 had a low difference in terms of TCi, yet participants were able to identify the fragment with the lowest, though the second-lowest fragment was mostly labelled as to whether the most complex or the least complex. This opens a discussion about the aspects that were considered by participants at the time of labelling fragments.

Some of the reasons for this distribution of results may be due to the fact that, first, the difference between fragments 3 and 4 in terms of TCi is considerably low (0.2); thus, nearly half of the participants scored Fragment 3 as the least complex, as they had difficulties identifying a significant difference between both fragments. Regarding Fragments 1 and 2, there was a significant distance in terms of textual complexity between them. Nevertheless, more than 50% of the participants labelled Fragment 2 as the most complex over the two, and nearly half of the participants labelled Fragment 1 as the most complex. One of the possible reasons for this may be the prevalence of one of the components considered in this model over others for evaluating this aspect since, despite having an overall higher score for text complexity, some fragments have lower aspect-by-aspect rates than the ones in fragments with overall lower scores.

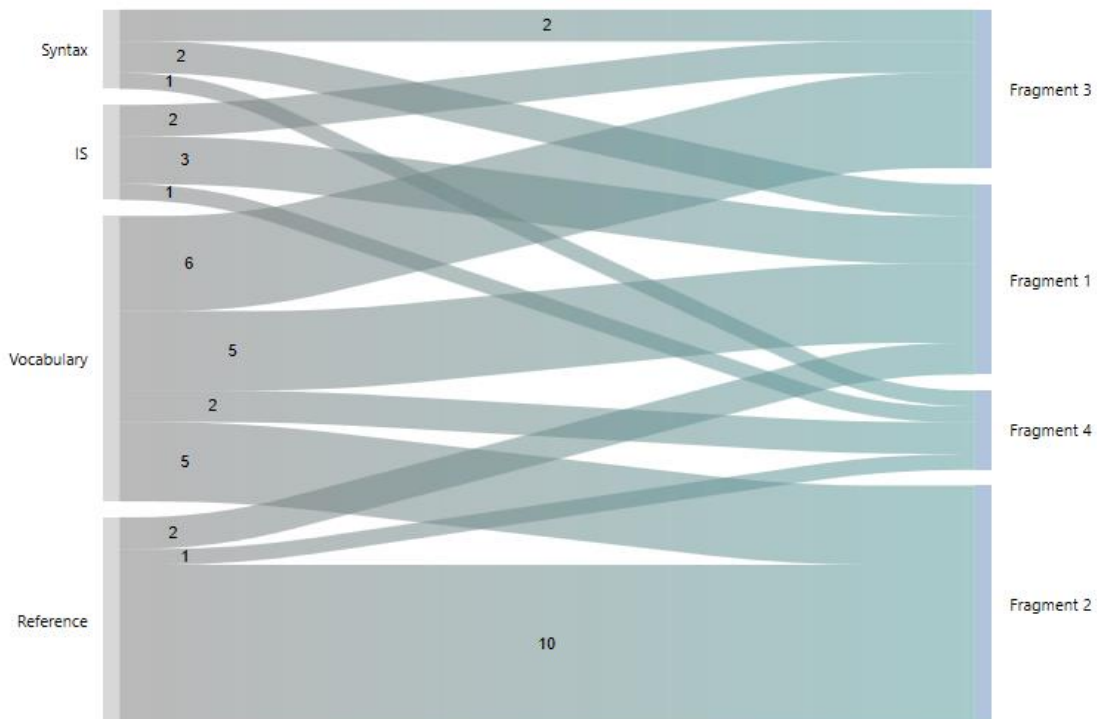
The overall low degree of accuracy at the time of cataloguing the fragments allows concluding that, on the one hand, expert teachers' judgement is only partially relatable to this measurement for complexity (RQ 1.1) as it is possible to identify a tendency towards identifying highly complex texts as well as their simpler counterpart. Higher distances in terms of TCi, component measurement and expert rationale at the time of making a decision are aspects that should be studied further in order to re-test and check for higher or lower levels of relatability. Qualitative analysis in the next section sheds light on some other relevant pieces of information that can complement the already achieved conclusions.

5.2. Expert instrument qualitative results and analysis

Regarding the qualitative information retrieved from the questionnaires, each fragment was commented on by the participants, and 5 major labels were identified, corresponding to the 4 components of the text complexity measurement that is part of this study and 1 extra label for other relevant comments which could be catalogued as incidental findings. As it is possible to observe in Figure 2, participants provided different amounts of relevance to different features of the construction in the fragments analysed. The most commonly found label was vocabulary, which is common for the 4 fragments of the text, and was mentioned 18 times as a factor that may affect reading comprehension in a positive or in a negative manner. Syntax, information structure and reference are labels that appeared only in the comments for 3 of the 4 fragments; from these three, the label “reference” was the most frequent, with 13 occurrences.

Figure 2.

Label occurrence per fragment across participants.



Note. Numbers correspond to the number times the label identified in each fragment.

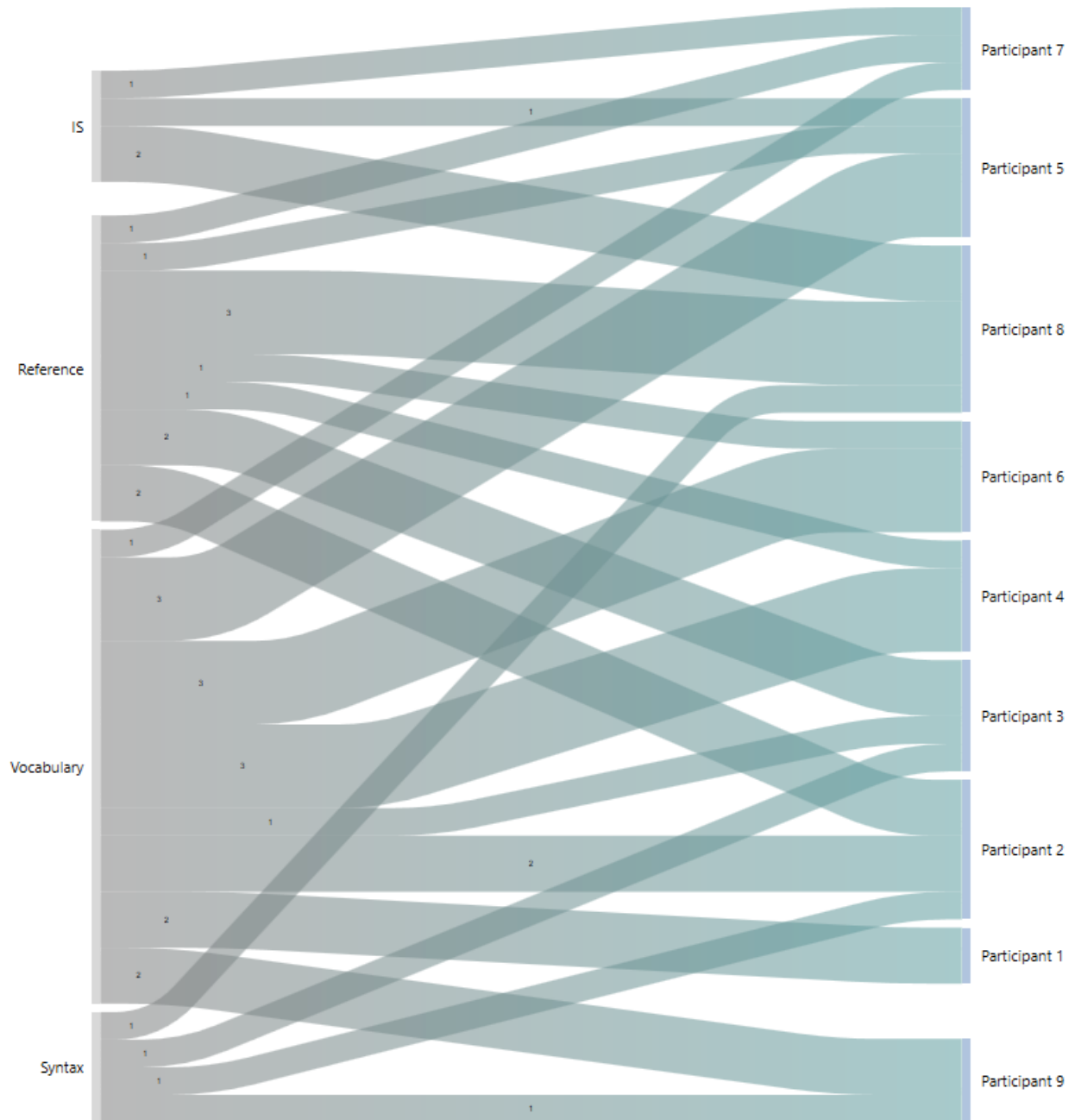
In terms of the distribution of labels in the different fragments, it is possible to observe that these are distributed unevenly, which means that participants did not identify all the aspects as constituents of complexity for each fragment, and, at the same time, some fragments have condensed most of the labels for certain components. For example, fragment 2 for participants is complex due to the lack of reference instances, which make the text redundant, as well as the repetition of keywords; this leaves aside aspects such as information structure or syntax. Similar to this example, fragment 3 is commented mainly in terms of its vocabulary, information structure and syntax, making no mention of referential expressions as aspects that contribute to complexity, thus, affecting students' reading comprehension processes. Another observation that is relevant for this analysis is that for some participants, no label was retrievable. For others more than one per fragment, moreover, the labels of vocabulary and syntax are obtained from the explicit mention to these concepts, while information structure and reference were obtained from the interpretation of the answers given by participants. Taking into account this first approach to the data obtained, it is possible to complement the points made in the quantitative analysis, as well as to address RQ1.2 by identifying vocabulary as the most relevant component of this text complexity measurement at the time of deciding how complex a text is. Similar to this are referential expressions, which, though not mentioned explicitly, are explained through phrases such as "repetition of information" or "redundancy". Finally, information structure and syntax are the least mentioned components of the measurement, as the first is referred to as "disorganisation of ideas" and the latter as "sentences that are too long", but comparatively speaking, these were not as relevant as the vocabulary and reference for experts in the creation and selection of materials for reading comprehension.

As a means to complement the discussion started in the quantitative analysis regarding the components of the text complexity measurement that are taken into account by teachers at the time of analysing materials. Figure 3 shows the labels from the answers in each of the participants, which allow identifying a general tendency of participants to observe vocabulary and reference, as they are the most repeated label which also appeared in all the analyses made by the participants. Among the different combinations of aspects to address by participants, it is relevant to mention that no participant made reference to all the components of the measurement at the moment of analysing complexity in the exercise; some tendencies are to consider reference and vocabulary, or information structure and vocabulary. The most frequent combination of components is

reference, syntax and vocabulary, and the two components that are not together in any of the analyses are syntax and information structure.

Figure 3.

Label occurrence per participant across fragments.



Note. Numbers correspond to the number times the label identified in each fragment.

These observations allow us to confirm the idea from the quantitative analysis, that, as participants are not analysing each fragment according to the four different aspects, they are not going to be accurate in the identification of the level of complexity for each of them, irrespective of the difference these fragments have in terms of TCi. Similarly, it is possible to assert that while vocabulary and reference are standard components for experts to analyse in texts, syntax and information structures are not as expected. This may be due to various reasons; for syntax, this may be due to the considerable amount of perspectives from which this concept can be addressed for reading comprehension (number of words per sentence, number of clauses, structures used, etc.). This directly affects information structures, as they are directly connected to certain grammatical patterns for the accomplishment of specific communicative purposes. Current tendencies that advocate for a more communicative approach for the use of language tend to give syntax a secondary role over the weight of vocabulary (Frantz, Starr & Bailey, 2015), thus affecting teachers' perception about the relevance of these factors and the different levels of complexity they provide to texts beyond the length of sentences.

Some other relevant remarks are that those participants who considered more aspects for their answer got a higher degree of accuracy in their identification of the more complex fragments; similarly, those participants who took into account elements beyond syntax and vocabulary obtained higher scores, such as the case of participant 8, who achieved a 100% of accuracy by considering information structure, reference and syntax. Opposite to this is the case of participant 1, who obtain a 0% by only considering vocabulary on his argument for choosing the most complex fragments. Along the same line, those answers that showed more evident theory-based arguments are the ones that obtained higher scores for accuracy, as can be observed in appendixes 2 to 5.

As a conclusion for the analyses of this instrument, and as a means to give an answer to RQ1, there is a correlation between experts' judgement and text complexity measurements, though both the measurement and experts' foci must be developed further in order to increase correlation among them. It is possible to observe that the more comprehensive the analysis from the expert, the more accurate the identification of complex fragments is, and this is boosted if the expert has prior knowledge about referential expressions and information structure. Finally, this needs to be studied further, with adaptations to the formulae as well as with fragments with higher differences among each other in order to obtain the expected results.

5.3. Student instrument quantitative results and analysis

Regarding the scores obtained by the 4 students that participated in this study, Table 4 shows a clear correlation between CEFR level of linguistic skills and comprehension tasks for each of the fragments, as students obtained higher scores in the tasks based on the CEFR level they have. This allows us to infer that participants with higher levels of linguistic competence comprehend texts in a better manner with higher levels of text complexity. Meanwhile, students with lower levels of complexity have difficulties understanding all the information from texts with higher TCi.

Table 4.

Comprehension tasks scores per participant.

	Fragment 1 (TCi 2.5)	Fragment 2 (TCi 2.0)	Fragment 3 (TCi 1.8)	Fragment 4 (TCi 1.7)	Total Score
A1	0	1	2	3	6
A2	1	1	3	2	7
B1	2	2	3	2	9
B2	2	3	2	3	10

Note. These scores were obtained based on 3 reading comprehension questions per fragment, each of which was scored with 1 point per correct answer.

These results allow RQ2.1 to be answered, as there seems to be a correlation between CEFR levels and TCi for fragments at the time of testing the reading comprehension of learners who have been catalogued in these different levels of language command. Nevertheless, it is essential to argue that, despite of the direct relation between the scores obtained and the TCi for each fragment, these results are not conclusive, as the tested sample is not mathematically significant. Moreover, participants at the lower levels of competence (A1 and A2) seem to have a more direct relation with TCi measurements for fragments, as students with higher levels of competence (B1 and B2) obtained similar scores among the 4 levels that were established for this study. Similar to the results obtained from the expert sample, it is possible that texts with higher differences in terms of TCi would allow identifying if there are any significant differences and at which range of difficulty.

Regarding accuracy, it is possible to observe a contrast between the participants' accuracy at the time of determining the complexity of the fragments read. As it can be seen in Table 4, it is possible to observe a tendency where students with higher CEFR levels for language skills have higher scores for accuracy than their counterparts. Similar to the teachers' scores, the least and most complex fragments were identified, while stretches of text for intermediate levels of complexity were not identified.

Table 5. Complexity measurement accuracy per participant.

	Fragment 1	Fragment 2	Fragment 3	Fragment 4	Accuracy
B2	4	3	2	1	50%
B1	2	4	3	1	50%
A2	1	4	3	2	25%
A1	1	3	4	2	0%
				Average	38%

Note. The correct order (read from left to right) for the degree of textual complexity is 3, 4, 2, 1.

Some other relevant pieces of information are the fact that students in the B level both scored the same amount of accuracy, while there seems to be a more significant difference between students in levels A1 and A2, as the participant in the A1 level scored 0% of accuracy. This may be due to the lack of exposition of the participant in the A1 level to reading tasks in English, as well as metalinguistic tasks, thus achieving the lowest score. Regarding the scores for B1 and B2 participants, though they are the same, the participant in the B2 level is implicitly more accurate than their counterpart, as they were able to label both complex fragments as well as both less complex fragments, but with mixed labels. In contrast, the participant from the B1 level was only correct in their evaluation of fragments 2 and 4, which are the extremes for complex and less complex.

In terms of fragment accuracy (shown in Table 6), the two stretches with the highest scores are 2 and 4, which are the fragments with the lowest and highest TCis accordingly, fragments 1 and 3, on the other hand, obtained 0% and 25% of accuracy, which means participants presented more difficulties at the time of identifying them. If contrasted with the scores obtained by the first group of participants (experts) this distribution shows a tendency towards identifying extreme

TCis, while it shows a clear disadvantage when identifying the levels in among these extremes. Notwithstanding, it is possible to connect the 25% of accuracy for fragment 3 with the low difference between TCis for fragments 3 and 4, indicating that participants at higher levels of language competence (B2) are more sensitive towards subtle TCi differences rather than significant differences such as the 0.5 TCi points between fragments 1 and 2. Another explanation for this, along the lines of the previous group of participants, can be the prevalence of some components of text complexity over others, which allows some participants to benefit from unbalanced internal scores.

Table 6.

Complexity measurement accuracy per fragment.

	Fragment 1 (TCi 2.0)	Fragment 2 (TCi 2.5)	Fragment 3 (TCi 1.8)	Fragment 4 (TCi 1.7)
Correct	0	2	1	2
Incorrect	4	2	3	2
%	0%	50%	25%	50%

Note. These results were obtained from a sample of 4 participants.

Similar to the previous observation made, the sample is not significant enough so as to make conclusive remarks on the validity of this measurement. Qualitative information analysed in the upcoming section sheds light upon some of the phenomena observed in this analysis of results, complementing some ideas about accuracy and reading comprehension.

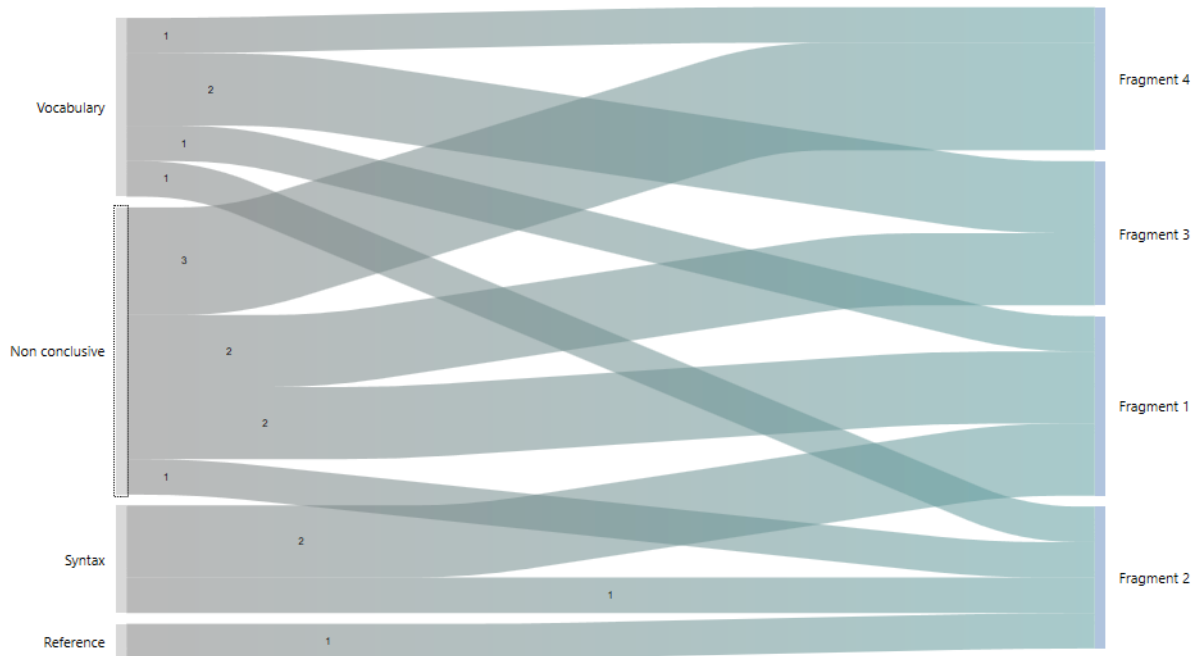
5.4. Student instrument qualitative results and analysis

Qualitative information collected for this group of participants allowed the identification of labels that share similar characteristics to those of the first group of participants. While vocabulary and syntax are consistent and play a significant role in the labelling process of text fragments, reference has a low appearance (1) and information structure is not mentioned nor referred to in the analyses made by this group of participants. Moreover, a new label emerged, which corresponds to non-conclusive observations, this label makes reference to difficulty perceptions stated by the participants without further argument, with examples such as “the paragraph was

easy/difficult to read”, among others. The occurrences of each of these labels per each fragment of the text are characterised in Figure 4.

Figure 4.

Label occurrence per fragment across participants.



Note. Numbers correspond to the number times the label identified in each fragment.

As it can be observed, all the fragments were polluted by non-conclusive observations, which cannot be labelled as any of the components of the complexity framework designed for this study, the second most common label is vocabulary, and it appears in all the fragments, in this case as an element that causes comprehension problems due to unknown concepts. Along this lines, reference is only present in one case, making reference to fragment 2, which is one of the fragments that has a high amount of accuracy. Similar to what happened in participant group 1, fragment 2 is an example of various components of text complexity contributing to the correct labelling of a fragments' complexity which is coherent to its TCi. Opposite to this example is fragment 4, which obtained the same accuracy score, and its identification is based only on one of the framework's components (vocabulary). One of the conclusions that can possibly be achieved from this example

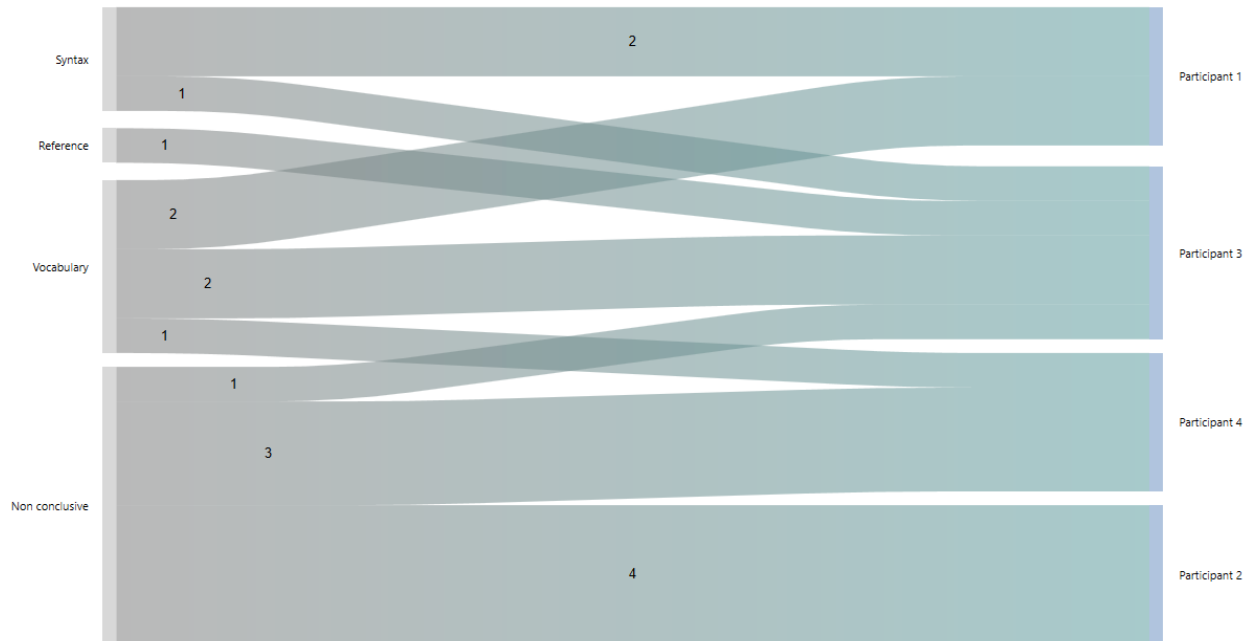
is the prevalence of vocabulary for both teachers and students as a significant aspect for both to consider, which at the same time is a predictor for text complexity, which can be effectively complemented by considering other aspects, such as syntax and reference, if possible.

Finally, information structure does not appear as a label in this case; this may be due to the fact that the metalinguistic knowledge of students from General English courses is developed only for the purpose of interaction and is probably not addressed with the purpose of achieving higher-level skills, such as evaluating a text's complexity. As a consequence of this observation, it is possible to raise the question if this should be a type of knowledge that would increase students' performance in comprehension tasks, as it may facilitate understanding the communicative intentions behind specific sentence patterns.

Regarding participants' use of labels, a varied range of cases emerged and are presented in Figure 5. On the one hand, participant 2 (A1) makes only use of non-conclusive labels, which may explain their lower accuracy (0%) at the time of identifying the complexity levels of each fragment. The best explanation for this situation is that A1 is an entry-level of language competence, which implicates the lowest language domain and metalinguistic knowledge in the group; this, in turn, was also reflected in the comprehension scores for this participant, which were the lowest for the task.

Figure 5.

Label occurrence per participant across fragments.



Note. Numbers correspond to the number times the label identified in each fragment.

Similar to participant 2, participant 4 (B1) made use of 3 non-conclusive labels and only 1 vocabulary-related label at the moment of analysing fragment complexity. In spite of the fact that this is contradictory with their accuracy rate (50%), a possible explanation for this may have to do with phenomena from the language acquisition process, such as the plateau effect and u-shaped trajectory of learning (Ellis, 2015) which account for students with high levels of competence facing a drop in terms of their capacity to manage language as part of the internalisation process which is necessary for learning. Thus, the participant was able to obtain a high score in reading comprehension, which is proper of linguistic competence, but performed regularly in an analysis task, such as complexity checking per fragment.

Participant 3 (B2), different from the previous two cases, was able to make varied use of labels for the identification of complexity levels. This contributed to their 50% accuracy plus the range difference identification that was mentioned in the previous sub-section. In the case of this participant, the use of reference, as well as vocabulary and syntax show a display of an advanced knowledge about the language, which is consistent with the observations made for answering RQ

2.1, and which shed light also on RQ 2.3, as this participant is the only one in using this label. Similar to this case, is participant 1 (A2), who was able to score a 50% of accuracy for complexity identification, based on the use of syntax and vocabulary-related labels, which in turn correlate with a higher level of language knowledge than participant 2, together with a higher capacity for reading comprehension.

As a means to answer to RQs 2.2 and 2.3, Table 7 shows a distribution of text complexity components and their use by participants at different CEFR competence levels.

Table 7.

Components referred to by participants per CEFR level.

	Non Conclusive	Vocabulary	Syntax	Reference
B2	✓	✓	✓	✓
B1	✓	✓	X	X
A2	X	✓	✓	X
A1	✓	X	X	X

Note. ✓ implicates the use of the label at least once, and X means total absence.

Given the twofold purpose of the task carried out by this group of participants, components highlighted for reading comprehension and complexity fulfil the same purpose. Taking into account this information, it is possible to observe that the participant with the highest level of competence is able to use all the labels identified by the group in a mixed manner which allowed them to obtain the highest results for comprehension and for complexity identification accuracy. The use of non-conclusive label for B1 and B2 students may be related to more elaborate concepts, such as reference, information structure, or other, which are probably not entirely internalized by them so as to make effective use of them for analysis purposes. Participants in levels A1 and A2, on the other hand, make somehow effective use of vocabulary and syntactic knowledge, with the only difference that A2 is aware of the concepts and is able to use them for analysis, whereas A1 is in a state where vocabulary and syntax as metalinguistic knowledge are not completely internalized yet.

As a final remark for this subsection in the results and analysis section, RQ 2 is answered partially with a positive correlation between students' scores and the complexity framework proposed, as students at different levels make use of the framework's components to aid their comprehension and identifying how complex a text is. Both quantitative and qualitative analyses sustain the use of TCi for framing texts of different complexities for pedagogical purposes, addressing the needs of students at different linguistic competence levels. Nevertheless, more comprehensive testing is necessary to establish more precise parameters for this measurement, allowing more accurate measurements and ranges.

6. Conclusion

This research study aimed to re-think the concept of text complexity as a means to provide teachers and researchers with a holistic approach for text analysis. This novel approach re-opens the discussion of which aspects must be considered for the adaptation of texts, which has to have more than two hundred different models and formulas led in the last thirty years. By building a complexity model based on text grammar, large-scale aspects of the text such as intentionality, cohesion and information activation and de-activation are considered. These offer a comprehensive and contextualised view of the text as a whole, and in turn, allow to better meet the needs of students.

The results obtained from this experimental design allowed us to partially confirm the relevance and validity of this new measurement for text complexity, not from a theoretical, but from an applied perspective, considering tasks usually addressed by teachers, such as selection of reading materials due to certain features for their classes, and also by students of EFL, who are constantly facing reading comprehension tasks, interacting with texts that are challenging, yet readable for their level of competence. Positive results from the quantitative and qualitative perspectives allow us to confirm vocabulary, syntax, reference and information structure as the ideal components for calculating the complexity of texts, merging the perspectives of various researchers on the topic in the last decade. Moreover, formulae selected for this framework showed to have an impact on teachers' judgement of complexity as well as on students' performance in reading comprehension tasks, since both measurements for maximum and minimum complexity were recognized by teachers and had an impact on students' scores across levels of competence.

This work affects the selection, adaptation and creation of materials for the EFL classroom directly, as it invites teachers to broaden their construct of complexity and invites them to re-focus on text construction, reverting the tendency to focus on contextual and learners' factors only. Based on the results obtained, a text that is crafted for a certain level of competence has a real impact on student comprehension, and in turn, affects their perception of complexity, which impacts directly on other learner factors such as motivation and fatigue. The use of this tool by teachers in all contexts of EFL teaching and learning also serves as a message of empowerment, giving teachers back the power to select their own materials based on a theoretically updated, accurate and context-sensitive perspective. This decision-making process will allow teachers to

make use of the texts they find interesting for their students, and adapt them based on an empirically proven tool rather than imprecise expert knowledge.

The reach of this work also extends beyond the barriers of reading comprehension materials for the EFL classroom; if adapted, this tool can also be applied to the selection of listening materials for both teaching and learning processes, and also for classroom and high-standards evaluation. Moreover, this tool can also be adjusted to be used in the literacy process of young learners, allowing teachers to emphasize those features of language to be acquired by children in their first years of education. Inside and outside school, this framework may allow to carry out adaptation processes for texts for people with special educational needs, as a means to optimize their comprehension of materials, but also in case means of communication want to become inclusive. Regarding research purposes, this tool also allows the measurement of features in different languages, allowing linguists and translators to identify traits and complexities which belong exclusively to certain languages.

In terms of the limitations of this study, it is relevant to mention that the range of TCIs selected for the study is limited, as the material was intended to be brief, with the intention of minimizing fatigue in the participants. Moreover, the group of students who participated in this study was not mathematically significant so as to validate the measurement entirely in contrast with CEFR groups. The original number of participants for this study was 25 students, distributed in one group per CEFR level (A1 to C2). However, context limitations reduced that number to four participants. The number and type of comprehension questions can also be addressed as limitations since at some language competence levels, general and specific information questions are not enough so as to measure comprehension. Finally, the type of text can also be addressed as a limitation, since its layout is a list of data, which has low levels of cohesion among paragraphs, and can be framed in a particular genre; thus, the complexity framework should be tested in narrative and informative texts which have a regular structure.

Finally, regarding further research, it is highly recommended to carry out a contrastive analysis of formulae for lexical and syntactic diversities, as well as a revision of the formulas proposed for reference and information structure, with the aim of identifying the most accurate constructs for each of the components. Moreover, it is recommended to integrate other aspects that were not taken into account for this approach, such as lexical tightness, frequency, mutual

information, and other aspects that have been analysed by previous research, as a means to make the approach become more comprehensive and accurate. Another interesting aspect to address in future research is the connection between different levels of complexity and anxiety, fatigue and motivation. Finally, carrying out a longitudinal study with a group of students along their L2 learning process and testing texts with different levels of complexity would be an excellent exercise to achieve higher levels of validity and establish accurate TCis for each stage of L2 learning.

7. References

- Akbari, R., & Yazdanmehr, E. (2014). A Critical Analysis of the Selection Criteria of Expert Teachers in ELT. *Theory & Practice in Language Studies*, 4(8).
- Alderson, J. C. (1984). Reading in a foreign language: A reading problem or a language problem. *Reading in a foreign language*, 1, 24.
- Barbaresi, L. M. (2011). A "natural" approach to text complexity. *Poznań Studies in Contemporary Linguistics*, 47(2), 203.
- Barbaresi, L. M. (2003). Towards a theory of text complexity. *Complexity in language and text*, 23.
- Bulté, B., & Housen, A. (2012). Defining and operationalising L2 complexity. *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp 23-46). John Benjamin's Publishing.
- Cook, V. (2016). *Second language learning and language teaching*. 5th Ed. Abingdon: Routledge.
- Corvalán, P. (2007). Practical English Text Grammar. *Cuadernos de la Facultad*. Santiago: Colección Monografías UMCE.
- Corvalán, P., & Zenteno, C. (2009). La gramática textual: un paradigma en la enseñanza de la gramática del inglés como lengua extranjera. *Lenguas modernas*, 34, 97-116. <http://repositorio.uchile.cl/handle/2250/123145>
- Crossley, S.A., Dufty, D.F., McCarthy, P.M., & McNamara, D.S. (2007). Toward a new readability: A mixed model approach. In D. S. McNamara and G. Trafton (Eds.), *Proceedings of the 29th Annual Conference of the Cognitive Science Society*. Nashville, TN: Cognitive Science Society.
- Dasclu, M. (2016). *Analyzing Discourse and Text Complexity for Learning and Collaborating: A Cognitive Approach Based on Natural Language Processing*.
- De Beaugrande, R. A., & Dressler, W. U. (1981). *Introduction to text linguistics* (Vol. 1). London: Longman.
- Donato, R. (2000). Sociocultural considerations to understanding the foreign and second language classroom. *Sociocultural theory and second language learning*, 78(4), 27.

- DuBay, W. H. (2007). *Smart Language: Readers, Readability, and the Grading of Text*. Surge.
- Ellis, R. (2015). *Understanding second language acquisition 2nd Edition-Oxford applied linguistics*. Oxford university press.
- Evans, J. & Sahgal, N. (June 29, 2021). Pew Research Center. *Key findings about religion in India*.
<https://www.pewresearch.org/fact-tank/2021/06/29/key-findings-about-religion-in-india/>
- Flesch, R. (1948). A new readability yardstick. *Journal of applied psychology*, 32(3), 221.
- Flor, M., Klebanov, B. B., & Sheehan, K. M. (2013, June). Lexical tightness and text complexity. In *Proceedings of the Workshop on Natural Language Processing for Improving Textual Accessibility* (pp. 29-38).
- Frantz, R. S., Starr, L. E., & Bailey, A. L. (2015). Syntactic complexity as an aspect of text complexity. *Educational Researcher*, 44(7), 387-393.
- Gasperin, C., Maziero, E., Specia, L., Pardo, T., & Aluisio, S. M. (2009). Natural language processing for social inclusion: a text simplification architecture for different literacy levels. *The proceedings of SEMISH-XXXVI seminário integrado de software e hardware*, 387-401.
- Goodwyn, A. (2010). *The expert teacher of English*. 1st Ed. Abingdon: Routledge.
- Goldman, S. R., & Lee, C. D. (2014). Text complexity: State of the art and the conundrums it raises. *the elementary school journal*, 115(2), 290-300.
- Grabbe, W., & Stoller, F. (2011). *Teaching and Researching: Reading*. 2nd Ed. Abdingdon: Routledge.
- Graesser, A. C., & McNamara, D. S. (2011). Computational analyses of multilevel discourse comprehension. *Topics in Cognitive Science*, 3 (2), 371–398.
- Halliday, M. A. K. and R. Hasan. 1976. *Cohesion in English*. Harlow, Essex: Longman Group Ltd.
- Hiebert, E. H. (2011). Beyond single readability measures: Using multiple sources of information in establishing text complexity. *Journal of Education*, 191(2), 33-42.

- Hiebert, E., & Pearson, D. (2014). Understanding text complexity: Introduction to the special issue. *the elementary school journal*, 115(2), 153-160. <https://doi.org/10.1086/678446>
- Housen, A., Kuiken, F., & Vedder, I. (Eds.). (2012). *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (Vol. 32). John Benjamins Publishing.
- Janan, D. (2011). *Towards a new model of readability* (Doctoral dissertation, University of Warwick).
- Justicia, J. M., & Padilla, M. A. S. (2011). Análisis cualitativo asistido por ordenador con ATLAS.ti. *Investigar en psicología de la educación. Nuevas perspectivas conceptuales y metodológicas*, 299-363.
- Kuiken, F., & Vedder, I. (2007). Task complexity and measures of linguistic performance in L2 writing. *International Review of Applied Linguistics in Language Teaching*, 45(3), 261-284. <https://doi.org/10.1515/iral.2007.012>
- Korzen, I. (2021). Are some languages more complex than others? On text complexity and how to measure it. *Globe: A Journal of Language, Culture and Communication*, 12, 18-31.
- Kumaravadivelu, B. (2006). *Understanding language teaching: From method to postmethod*. Routledge.
- Mc Laughlin, G. H. (1969). SMOG grading-a new readability formula. *Journal of reading*, 12(8), 639-646.
- Mesmer, H. A., Cunningham, J. W., & Hiebert, E. H. (2012). Toward a theoretical model of text complexity for the early grades: Learning from the past, anticipating the future. *Reading research quarterly*, 47(3), 235-258.
- Miestamo, M. (2004, May). On the feasibility of complexity metrics. In *FinEst linguistics, proceedings of the annual Finnish and Estonian conference of linguistics, Tallinn* (pp. 11-26).

- Mosquera, A., Lloret, E., & Moreda, P. (2012). Towards facilitating the accessibility of web 2.0 texts through text normalisation. In *Proceedings of the LREC workshop: Natural Language Processing for Improving Textual Accessibility (NLP4ITA)* (pp. 9-14).
- Nation, I. S., & Webb, S. A. (2011). *Researching and analyzing vocabulary*. Boston, MA: Heinle, Cengage Learning.
- Nelson, J., Perfetti, C., Liben, D., & Liben, M. (2012). Measures of text difficulty: Testing their predictive value for grade levels and student performance. *Council of Chief State School Officers, Washington, DC*.
- O'Donnell, M. E. (2009). Finding middle ground in second language reading: Pedagogic modifications that increase comprehensibility and vocabulary acquisition while preserving authentic text features. *The Modern Language Journal*, 93(4), 512-533.
- Oh, S. (2001). Two types of input modification and EFL reading comprehension: Simplification versus elaboration. *TESOL quarterly*, 35(1), 69-96. <https://doi.org/10.2307/3587860>
- Ortega, L. (2008). *Understanding Second Language Acquisition*. 1st Ed. Abdington: Routledge.
- Palmer, D., Stough, L., Thomas K. Burdenski, J., & Gonzales, M. (2005). Identifying Teacher Expertise: An Examination of Researchers' Decision Making. *Educational Psychologist*, 40, 13 - 25.
- Perfetti, C. A., Landi, N., & Oakhill, J. (2005). The Acquisition of Reading Comprehension Skill. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 227–247). Blackwell Publishing. <https://doi.org/10.1002/9780470757642.ch13>
- Pearson, P. D., & Hiebert, E. H. (2014). The state of the field: Qualitative analyses of text complexity. *The Elementary School Journal*, 115(2), 161-183.
- Pienemann, M. (2005). An introduction to processability theory. *Cross-linguistic aspects of processability theory*, 30, 179-199.
- Pienemann, M. (2010). A cognitive view of language acquisition: Processability theory and beyond. In *Conceptualising 'learning' in applied linguistics* (pp. 69-88). Palgrave Macmillan, London.

- Pienemann, M. (2015). An outline of processability theory and its relationship to other approaches to SLA. *Language Learning*, 65(1), 123-151.
- Pollatsek, A., & Treiman, R. (2015). *The Oxford handbook of reading*. 1st Ed. New York: Oxford University Press.
- Pollitt, A. (2018). *The Meaning of Oxford Placement Test Scores*. Oxford: Oxford University Press.
- Proctor, C. P., August, D., Snow, C., & Barr, C. D. (2010). The interdependence continuum: A perspective on the nature of Spanish–English bilingual reading comprehension. *Bilingual Research Journal*, 33(1), 5-20.
- Purpura, J. (2007). *The Oxford Placement Test: What does it measure and how?*. Oxford: Oxford University Press.
- Ravid, D., Dromi, E., & Kotler, P. (2010). Linguistic complexity in school-age text production: Expository versus mathematical discourse. In M. A. Nippold & C. M. Scott (Eds.), *Expository discourse in children, adolescents, and adults: Development and disorders* (pp. 123–150). New York, NY: Taylor & Francis.
- Reed, D. K., & Kershaw-Herrera, S. (2016). An examination of text complexity as characterized by readability and cohesion. *The Journal of Experimental Education*, 84(1), 75-97.
- Riddle, E. M. (2008). Complexity in isolating languages: Lexical elaboration versus grammatical economy. *Language complexity: Typology, contact, change*, 131-151.
- Saville-Troike, M. (2012). *Introducing Second Language Acquisition*. 2nd Ed. Cambridge: Cambridge University Press.
- Sheehan, K. M., Flor, M., & Napolitano, D. (2013, June). A two-stage approach for generating unbiased estimates of text complexity. In *Proceedings of the Workshop on Natural Language Processing for Improving Textual Accessibility* (pp. 49-58).
- Sheehan, K. M., Kostin, I., & Futagi, Y. (2008). When do standard approaches for measuring vocabulary difficulty, syntactic complexity and referential cohesion yield biased estimates

of text difficulty. In *Proceedings of the 30th Annual Conference of the Cognitive Science Society, Washington DC*.

Smith, D. R. et al (1989). The Lexile Scale in Theory and Practice. Final Report for NIH Grant IID - 19448.

Štajner, S., Evans, R., Orasan, C., & Mitkov, R. (2012). What can readability measures really tell us about text complexity. In *Proceedings of workshop on natural language processing for improving textual accessibility* (pp. 14-22).

Swartz, C. W., Burdick, D. S., Hanlon, S. T., Stenner, A. J., Kyngdon, A., Burdick, H., & Smith, M. (2014). Toward a theory relating text complexity, reader ability, and reading comprehension. *Journal of applied measurement, 15*(4), 359-371.

Temnikova, I. (2012). Text complexity and text simplification in the crisis management domain (Publication No. 297482) [Doctoral dissertation, University of Wolverhampton]. <https://wlv.openrepository.com/handle/2436/297482>

Vajjala, S., & Meurers, D. (2012, June). On improving the accuracy of readability classification using insights from second language acquisition. In *Proceedings of the seventh workshop on building educational applications using NLP* (pp. 163-173).

Valencia, S. W., Wixson, K. K., & Pearson, P. D. (2014). Putting text complexity in context: Refocusing on comprehension of complex text. *The Elementary School Journal, 115*(2), 270-289.

Van Dijk. T. A. 1977. Text and context: explorations in the semantics and pragmatics of discourse. Harlow: Longman Group UK Ltd.

VanPatten, B. (2002). Processing the content of input—processing and processing instruction research: a response to Dekeyser, Salaberry, Robinson, and Harrington. *Language Learning, 52*(4), 825-831.

VanPatten, B. (2004). Input processing in second language acquisition. *Processing instruction: Theory, research, and commentary, 5-31*.

- VanPatten, B. (2013). Input processing. In *The Routledge handbook of second language acquisition* (pp. 286-299). Routledge.
- VanPatten, B., Williams, J. (Eds.). (2015). *Theories in second language acquisition: An introduction*. Routledge.
- Williamson, G. L., Fitzgerald, J., & Stenner, A. J. (2013). The Common Core State Standards' quantitative text complexity trajectory: Figuring out how much complexity is enough. *Educational Researcher*, 42(2), 59-69.
- Wray, D., & Janan, D. (2013). Readability revisited? The implications of text complexity. *The Curriculum Journal*, 24(4), 553-562.
- Yamaguchi, Y., & Usami, H. (2017). Plural marking in spoken and written narratives: A corpus-based study of Japanese learners of English. *International Journal of Applied Linguistics and English Literature*, 6(5), 224-231.

Appendix 1. Multi-level text with TCI

Key findings about religion in India

A new Pew Research Center report, based on a face-to-face survey of 29,999 Indian adults fielded between late 2019 and early 2020 takes a closer look at religious identity, nationalism and tolerance in Indian society. The survey was conducted by local interviewers in 17 languages and covered nearly all of India's states and union territories. Here are key findings from the report.

Indians value religious tolerance, though they also live religiously segregated lives. Across the country, most people (84%) say that to be "truly Indian," it is very important to respect all religions. Indians also are united in the view that respecting other religions is a very important part of what it means to be a member of their own religious community (80%). People in all six major religious groups overwhelmingly say they are very free to practice their faiths, and most say people of other faiths also are very free to practice their own religion. But Indians' commitment to tolerance is accompanied by a strong preference for keeping religious communities segregated.

5 Sentences, 12 Clauses = 0.41

111 Tokens, 63 Types = TTR 0.57

RE 5, RT 2 / Ref 12 = 0.83

SN 5 / ST 5 *1/4 = 0.25

Total = 2.06

For many Hindus, national identity, religion and language are closely connected. Nearly two-thirds of Hindus (64%) say it is very important to be Hindu to be truly Indian. Among Hindus who say it is very important to be Hindu to be truly Indian, 80% also say it is very important to speak Hindi to be truly Indian. Hindus who strongly link Hindu and Indian identities express a keen desire for religious segregation. Together, the Northern and Central regions cover the country's "Hindi belt," where Hindi, one of dozens of languages spoken in India, is most prevalent. The vast majority of Hindus in these regions strongly link Indian identity with being able to speak Hindi.

6 Sentences, 12 Clauses = 0.5

115 Tokens, 63 Types = TTR 0.55

RE 4, RT 3/ Ref 14 = 0.85

SN 6 / ST 5 *1/4 = 0.6

Total = 2.5

Religious conversion is rare in **India**; to the extent that *it* is occurring, **Hindus** gain as many **people** as they lose. **Conversion of people** belonging to lower castes away from **Hinduism** to **other religions**, especially **Christianity**, has been contentious in **India**, and **some states** have laws against **proselytism**. This **survey** finds that **religious switching** has a minimal impact on the size of **religious groups**. Across **India**, **98% of survey respondents** give the same answer when asked to identify *their* current religion and, separately, *their* childhood religion. **An overall pattern** of stability in the share of **religious groups** is accompanied by little net change from movement into, or out of, most religious groups.

5 Sentences / 10 Clauses = 0.5

112 Tokens, 79 Types = TTR 0.71

RE 3, RT 2/ Ref 17 = 0.35

SN 5 / ST 5 *1/4 = 0.25

Total = 1.81

Most Indians believe in God and say religion is very important in their lives. **Nearly all Indians** say they believe in **God** (97%), and roughly 80% of **people in most religious groups** say they are *absolutely certain* that **God** exists. The main exception is **Buddhists**, *one-third of whom* say *they* do not believe in **God**. **Indians** do not always agree about **the nature of God**; Most **Hindus** say there is one **God** with many manifestations, while **Muslims** and **Christians** are more likely to say, simply, "there is only one **God**." But across **all major faiths**, the vast majority of **Indians** say *that religion is very important in their lives*, and significant portions of each **religious group** also pray daily and observe a range of **other religious rituals**.

5 Sentences / 14 Clauses = 0.35

128 Tokens, 71 types = TTR 0.55

RE 3, RT 3/ Ref 23 = 0.39

SN 5 / ST 3 *1/4 = 0.41

Total = 1.7

Retrieved from: <https://www.pewresearch.org/fact-tank/2021/06/29/key-findings-about-religion-in-india/>

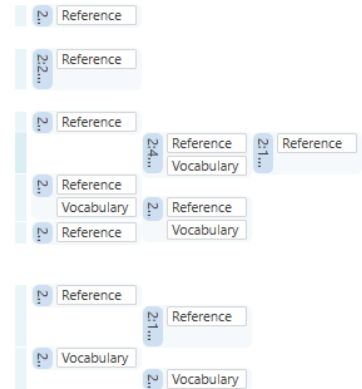
Appendix 2. Answers and labels for Fragment 1. (Experts' instrument)

1	El hecho de repetir palabras con estructura similar como communities, community, commitment... puede generar fatiga en el lector.
2	This specific fragment is not really complex to read and/or understand, but there is some repetition of words (e.g. faiths, religions, etc.) that makes its reading rather unpleasant.
3	1.- Some ideas are too long and make the reader lose the main point of the sentence. 2.- The use of the word "religious" was used 4 times within the text. A synonym could have been used in order not to make it so redundant.
4	The percentages might be problematic for some people who are not good at understanding them. However, depending on the levels of the target readers, the text seems not to be difficult to be understood by an average person regardless his or her level.
5	what makes it complex could be the vocabulary and main ideas.
6	Last paragraph should be after the first one to be more coherent, although it is really not complex to read
7	The complexity in this fragment is that the percentages are not in the same position. The first one is right after 'people', helping to identify what they mean by 'most people'; whereas in the other case the percentage is at the end of the sentence. So it makes it difficult to remember where that percentage comes from. The other complex factor I found is that there are adverbs placed in uncommon positions, so that also makes it difficult to predict what is coming in the text.
8	Regarding this fragment, I think that what makes it complex to understand is the fact that ideas are a bit contradictory and some statements seem to be disconnected from the rest of the text. For example, it mentions the fact that 84% of the people interviewed consider themselves to be "truly Indian", but then there is another statement that, in my opinion, has anything to add to the previous idea. Also, I consider that sentences are too long to read, especially if English learners at a low level read it. This may confuse them.
9	Several "long" words placed together (segregated, communities, religious, respecting, overwhelmingly, etc)
10	

S:1...	Vocabulary
S:2...	Reference
S:2...	Vocabulary
S:4...	Syntax
S:4...	Reference
S:5	Vocabulary
S:5	Vocabulary
S:5	IS
S:8 T:1...	IS
S:9...	IS
S:1...	Finding?
S:1...	Syntax
S:1...	Vocabulary

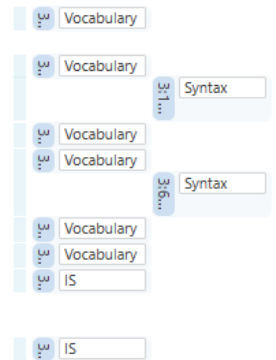
Appendix 3. Answers and labels for Fragment 2. (Experts' instrument)

1	¡Acá no sabría que decir
2	Once more there is some repetition in this fragment. For example: "it is very important to be Hindu to be truly Indian". However, it is still not complex to understand if we read it slowly.
3	1.- It made me think way too long whether the 64% represented the two-thirds of Hindus. 2.- The repetition of the sentence "it is very important to be Hindu to be truly Indian", could have been replaced in order to avoid repetition.
4	The repetition of the expression 'to be truly...' in the same paragraph might affect the message and some readers could be negatively affected by that. In the same token, the repetition of the words 'Hindus' and 'Indian' might be problematic for some readers.
5	It could be the repetition of words.
6	too many similar and repetitive words and sentences
7	In this fragment I think all the references to Indian, Hindu and Hindi are hard to follow. Although it is better because the author didn't replace it with pronouns or synonyms. I think that that would have been even more complex. But still, all the information using 'Indians', 'Hindu' and 'Hindi' are hard.
8	In this paragraph, I think that the problem is redundancy. The word "important" repeats many times along the text and it seems to me that the idea could be summarized and put in a more simple format / style.
9	Percentages used in places where quantity is expressed in a different way (Nearly two-thirds of Hindus (64%)), words that look or sound similar make it confusing (Hindu-Hindi-Indian)
10	



Appendix 4. Answers and labels for Fragment 3. (Experts' instrument)

1	El hecho de no conocer las definiciones de palabras clave como Proselytism, la verdad ni yo se lo que es y tuve que googlearlo.
2	This fragment has some complex vocabulary for some readers. For example, "proselytism" might not be a well known word for most people. Syntactically speaking, until now this was the most pleasant fragment to read.
3	1.- Vocabulary made it complex to understand, by using words like proselytism.
4	This paragraph appears to be complex mainly because of the lexis that it occupies, since some concepts like: 'proselytism' and 'conversion' may not be familiar by the average reader. Structure is not a problem but its knowledge demand might affect some readers' understanding.
5	It could be the more complex vocabulary.
6	Although is not complex, it has a "latin based" vocabulary that made me double-check the info
7	The complexity in this fragment, I think, is related to the difficulty to understand some phrases. To me, it was hard to get "to the extent that it is occurring", "Conversion of people belonging to lower castes away from Hinduism" and "An overall pattern of stability in the share of religious groups"
8	In my opinion, I would summarize the idea of this paragraph and also, re-arrange it, moving the last part (data) at the beginning and using the first lines as a sort of conclusion or closure, to round up the idea.
9	No problems with this one.
10	



Appendix 5. Answers and labels for Fragment 4. (Experts' instrument)

1	La verdad acá n encuentre nada que lo dificulte.
2	This fragment is also ok in terms of structure, but in terms of format it has a problem. Sometimes, the percentages were written between brackets, sometimes without brackets, sometimes in full and that format inconsistency makes it a little harder to understand.
3	1.- The way statistics were presented made it complex. I felt that I could have understood better if the text contained either only percentages or fractions written in words.
4	Repetition of: 'God' could affect some readers. However, the paragraph is clear and not complex, mainly because it does not need some cultural or specific knowledge to get the message it tries to spread.
5	it includes complex vocabulary, develop of supporting ideas
6	Too many quantities and items enumerated
7	I think that this is the simplest. This was not hard to understand.
8	Personally, I think that one of the problems with the text in general (including all fragments) is that it doesn't specify which religious / ethnic groups are being addressed throughout the survey and I think that this information is vital especially when starting the text. If compared to previous fragments, this is clearer in terms of organization and style, but still, there are some redundant ideas (such as the fact that Indians think that religious / God is very important)
9	Long phrases with different ideas separated by commas (But across all major faiths, the vast majority of Indians say that religion is very important in their lives, and significant portions of each religious group also pray daily and observe a range of other religious rituals)
10	

- 4 Finding?
- 4 Vocabulary
- 4 Vocabulary
- Reference
- 4 IS
- 4 Syntax

Appendix 6. Answers and labels for Fragment 1. (Students' instrument)

1		
2	There are some phrases which is weird. I don't know if they need verbs or just are extremely formal. For example; Indians value religious tolerance, though they also...	<input type="text" value="Syntax"/>
3	the paragraph is very clear.	<input type="text" value="Non conclusive"/>
4	This paragraph has some vocabulary and sentences that are difficult to understand.	<input type="text" value="Vocabulary"/> <input type="text" value="Syntax"/>
5	The paragraph is very easy to read.	<input type="text" value="Non conclusive"/>
6		

Appendix 7. Answers and labels for Fragment 2. (Students' instrument)

A veces siento que si hubieran ciertos artículos, verbos y preposiciones, podría ser más fluida mi lectura.	16...	Syntax
the whole paragraph is like a tongue twister	1..	Non conclusive
This paragraph is very confusing, as it mentions too many times the same names.	1.. 1.. 1..	Reference
This paragraph has many vocabulary problems.	1.. 1..	Vocabulary





Appendix 8. Answers and labels for Fragment 3. (Students' instrument)

The formal vocabulary
it is quite clear
The vocabulary makes it difficult to understand.
Some sentences are difficult to follow.

- Vocabulary
- Non conclusive
- Vocabulary
- Non conclusive

Appendix 9. Answers and labels for Fragment 4. (Students' instrument)

Some words that I didn't know it.
it was not too difficult
It was the most simple in the text.
It was easy but not so easy.

-  Vocabulary
-  Non conclusive
-  Non conclusive
-  Non conclusive