READING COMPREHENSION PROCESSES: A REVIEW BASED ON THEORETICAL MODELS AND RESEARCH METHODOLOGY

ANTONIJA BLAŽI OSTOJIĆ

Center of Speech and Language Pathology and Rehabilitation Blaži, Prva Prigorska 10A, Zagreb, Croatia, contact: antonijabo@blazi.hr

Abstract: *Reading comprehension is a complex, multifaceted process in which a number of components are appropriately and often simultaneously activated. The importance of different background subprocesses, its complex nature, as well as the interconnection of various components has been demonstrated by various 'models of reading and reading comprehension'. Based on these models, assessment methods and materials for reading comprehension are developed.*

However, the models and assessment materials developed so far are mostly derived from research on languages with non-transparent orthography (e.g., English). Therefore, the question arises regarding the extent to which they can or should be applied in languages with shallow, transparent orthography (e.g., Croatian) that have clear and consistent relationships between letters and sounds, as opposed to languages with deep orthography.

The main aim of this study was to present a brief review of prominent reading comprehension models and their interconnections through specific levels of language processing (single word, sentence, discourse), as well as to discuss the methodological aspects of assessing reading comprehension processes arising from the presented theoretical models. Motivated by the fact that there is a lack of studies on models and assessment materials in languages with transparent orthography, the application of existing models and assessment methods will be discussed in the context of transparent orthography languages.

This study provides comprehensive insights, based on theory, on the key elements to consider when developing an assessment method/tool for reading comprehension, both for research or diagnostic purposes.

Keywords: *reading comprehension, theory, methodology, language processing*

INTRODUCTION

Reading is a biologically secondary skill built from language, perception, memory, and learning. It is one of the most important forms of human communication and one of the most complex domains of language use (Snow, 2021). When we read, we read from printed meaning. The main purpose of reading is to understand what is written to acquire a wide range of information through different everyday situations (social, academic, employment, and so on) (Clarke et al., 2010). Reading comprehension is a multifaceted process that involves word recognition skills – identifying and decoding words, integrating lexical features, activating contextual meaning, computing sentence meaning, inferring causal connections, and extracting inferences (Scarborough et al., 2009;

Cain & Oakhill, 2006; Perfetti & Stafura, 2014). To be a skilled at comprehension, the activation of a wide range of skills, knowledge, and all of the above-mentioned processes must be done effortlessly, automatically, and unconsciously.

This multi-level complexity has been recognised in previous studies, leading to the development of various theoretical models of reading comprehension. These models mainly focus on individual pressure points, whose role has been shown to be significant in activating the process of understanding written language. There are only a few recently published review papers on this specific topic (see McNamara & Magliano, 2009; Babashamsi et al., 2013; Davoudi & Moghadam, 2015, etc.), with most of them focusing predominantly on few target models of reading

comprehension. Hence, a review of the previous literature found that there is a lack of comprehensive reviews of models of reading comprehension concerning specifically the level of language processing (Oakhill & Cain, 2007). Also, the lack of review papers is particularly noticeable in terms of the analysis of theoretical insights through methodological aspects that must be considered when constructing an assessment methods and tasks in this specific area. Finally, it is known that orthographic transparency plays an important role in reading processes (Seymour, 2006). If these models are not analysed with languages of different orthographic depth, it can be wrong to directly apply the premises of a model that has been tested using only one language (Share, 2008).

Considering the above-mentioned knowledge, the main purpose of this review paper was to provide a comprehensive description, analysis, and comparison of theoretical models of reading comprehension, followed by an analysis of methodological aspects (challenges, methods, and measures) of the written language comprehension assessment arising from a theoretical point of view. The final aim was to provide a foundation for future research in the area of reading comprehension processes and the encouragement for the future development of a comprehensive assessment tools for this construct in languages other than English (e.g., Croatian). Therefore, the possibility of implementing theoretical knowledge from existing models commonly developed in languages with opaque orthography (e.g., English) into our understanding of transparent orthography languages (e.g., Croatian, Spanish, Finnish) will be discussed.

MODELS OF READING COMPREHENSION

Models of reading comprehension can be defined from different points of view. In the present study, they will be presented by considering the specific level of language processing that they mainly refer to and specific processing components through which the individual differences can be observed (Perfetti, 2001). Additionally, these models differ in their *simplicity* or *compre-*

hensiveness. Because research shows that comprehension is not a standalone skill, most models now tend to develop a multicomponent approach to explain reading comprehension processes, such as *situation* models (Knitsch & Rawson, 2005). Naturally, multitudes of models stand, in their essence, as the improvement of the prior ones, thus giving a modernist approach to component interrelationships. Although there is no model so far that accounts for all the different components of reading comprehension, the intention of earlier models is not to separate specific components according to the order in which they are activated, but rather to explain their mutual connection, since both *bottom-up* and *top-down* processes are required simultaneously, particularly in the first stages of development (Gough & Turner, 1986; Knitsch & Knitsch, 2005).

In the following section, the fundamental, comprehensive, and most prominent models of reading comprehension process and their common points will be analysed and discussed. Of course, there are other, valuable models that further clarify the processes of reading comprehension. The following models were chosen because they represent the foundation for implementation of research on the process of reading comprehension and the development of a comprehensive test for the assessment of this construct. However, when studied within the framework of language typology, certain premises arising from these models can or cannot be directly applicable.

THEORETICAL BACKGROUND OF READING COMPREHENSION MODELS

Top-down, bottom-up, and interactive approaches

The psycholinguistic top-down (Goodman, 1967) and bottom-up (Stanovich, 1986; Rieben & Perfetti, 2013) approaches are among the first to attempt to clarify the nature of behaviour in the background of written comprehension processes. The two models differ in their explanation of how a reader *behaves* when reading and comprehending a text. According to the top-down model (Goodman, 1967), when the reader is faced with the

text, he/she first activates higher processing skills such as background knowledge, and cognitive and metacognitive strategies to compare and verify the information obtained from the text. In this way, the reader will speed up the process of word recognition. Respectively, this process will be activated only to confirm the previous, higher-level processing. In contrast, the bottom-up view (Stanovich, 1986; Barnet, 1989) directs one's attention to a gradual build-up of comprehension. In other words, the reader begins the comprehension process by activating the orthographic, phonological, syntactic, and semantic (lower-level processing, word recognition) processing skills, merely unconsciously and independently and without much regard for background knowledge. Since these two models led to great debates between researchers, the need for a more sufficient approach was emphasised and the *Interactive model* was developed (Rumelhart, 1977; Perfetti, 1985). This model proposes a unification of bottom-up and top-down models. Thus, it assumes that the reader reads and understands the text simultaneously – recognising and decoding the words, while comparing it with relevant background knowledge. Therefore, the subsequent models were developed in this direction.

Lower–level processing models (word-level)

Simple View of Reading

The development of the *Simple View of Reading* theory by Gough and Tunmer (1986) can be considered as the beginning of deeper explanations of the complexity of the phenomenon of reading comprehension. It is one of the most widely used theories and starting points in the reading comprehension research area. In this interactive and *not so simple* framework, the proportion of variance in reading comprehension that can be explained by decoding (based on phonology) and language comprehension (based on other language knowledge) is discussed using the following equation: **R** (reading comprehension) = **D** (decoding) x **C** (language comprehension). If reading comprehension is a product of both decoding and language comprehension skills (Gough & Tunmer, 1986;

Hoover & Gough, 1990), then all these components must be successfully activated in order to comprehend what is written. Therefore, one must visually recognise the word, adequately apply the grapheme-phoneme rule, and effectively process the linguistic information received. Observing the SVR principles through the well-known Scarborough's Reading Rope (2001), this is a significantly demanding task for a reader, especially in languages with no clear grapheme-phoneme relationship. Also, this theoretical point of view can be discussed in the case of a child with a language disorder – for example, children with specific and significant problems in receptive and expressive language; poor phonological/ morphological/ syntactic/ semantic/ pragmatic skills (e.g., Developmental language disorder, Dyslexia; Bishop & Snowling, 2004). To explain further, if any one of the components is violated (e.g., if a child has lower decoding skills; D = 0), the entire reading comprehension process is violated (the child will not understand the content read). This child can have a reading comprehension problem that can be attributable to poor word recognition skill (decoding), or general language comprehension problem, or both. Hence, when making a conclusion about one's reading comprehension, one must consider both decoding and general language comprehension skills, and their interrelationships between the two.

The SVR theory has been primarily developed and tested in languages with non-transparent orthography (Nation & Snowling, 2004; Oakhill & Cain, 2012), where there is no clear and consistent relationship between graphemes and phonemes. Interesting findings from such studies include the strong influence of decoding skills on reading and reading comprehension, even after the early stages of reading development (Goff et al., 2005; Vellutino et al., 2007). However, the same residual influence of decoding was not found in languages with shallow orthography (e.g., Spanish, Finnish, and so on; Florit & Cain, 2011), where greater influence was found in oral comprehension (e.g., Italian; Tobia & Bonifacci, 2015; e.g., Croatian; Zaretsky et al., 2009; e.g., Dutch; Patel et al., 2004). This can be explained by the fact

that, when decoding is not so challenging as it is in languages with deep orthography, most of cognitive resources can be devoted to comprehension. This knowledge must be considered when assessing reading comprehension because it is precisely in language comprehension that the fundamental problem can manifest itself, not in decoding.

Nevertheless, authors of the *Simple View of Reading* theory agree that once the reader has established the grapheme-phoneme relationship, further processes are almost the same as those that occur in oral language comprehension (activation of language features). This gives rise to the need for deeper clarification – what defines the concept of comprehension above the phonological level (e.g., word recognition/ decoding) of language processing?

Lexical Quality Hypothesis (LQH) and Reading System Framework (RSF)

If the *word* (oral or written) is observed through Levelt's model of language production (1993), then it is defined by its non-linguistic conceptual features (e.g., imageability) at the conceptual level, morphological and phonological features at the lexeme level, and syntactic and semantic aspects at the lema level. Putting this model in the framework of reading comprehension processes, the reverse order of events can be observed. After the visual recognition of words and the activation of phonological and morphological word features, the process of lexicalisation begins (the retrieval of mental representations of words from the mental lexicon). This is followed by selection of the lemma to choose the meaning that best matches the target word, according to the syntactic word features (Jescheniak & Level, 1994).

The need for rapid and efficient access to all word representations for successful reading comprehension is further emphasised through the newer Perfetti's *Lexical quality hypothesis* (Perfetti & Hart, 2002; Perfetti, 2007). This theoretical view integrates two main pressure points: word recognition and the comprehension system. To understand the background processes of reading comprehension, this hypothesis additionally points to the im-

portance of defining the word by its (linguistic) components and explains that particular internal properties of the lexical unit, the (psycho)linguistic word features, determine the word activation threshold. Accordingly, this process requires access to a highly qualitative linguistic information stored in the mental lexicon, for efficient and fast word recognition and retrieval of word meanings to form so called *propositions* – the word meaning units (Perfetti, 2007). A closer look at this theoretical view reveals similarities with the Levelt's model of language production (1993). The main idea is that the quality of the lexical code (knowledge of linguistic word features: phonological, orthographical, and semantical word meanings) determines how efficiently and reliably the reader can retrieve the word features from the mental lexicon and thereby understand what is written. In other words, if the quality of the lexical code is lower and the reader must exert a high cognitive effort to access the target lexical representations, then the quality of the lexical representations will be low and consequently may lead to difficulties in reading comprehension. Therefore, the ability to effectively access and retrieve lexical representation that define the word should be included in the overall assessment of reading comprehension. Moreover, this problem is also thought to disrupt reading comprehension processes at a higher level of language processing than that of a single word (e.g., inferencing).

Given the apparent complex nature of the reading comprehension process, the *Reading System Framework* (RSF; Perfetti & Stafura, 2014) was developed as an extension of the SVR theory, emphasising and linking the subprocesses such as word identification and word-to-text integration. The RSF model relies on three main sources of knowledge: linguistic, orthographic, and general knowledge. It implies that the reader uses these sources of knowledge in both a limited and an interactive way, to comprehend the written text. By activating word meanings from the mental lexicon, after graphemes have been converted into sounds and combined into words, the process of integrating individual word meanings into sentences and text model is enabled. Finally, the read-

er can create a mental situation model of the text by considering relevant background knowledge and text structure. These processes are observed within the cognitive system, whose pathways pass through visuo-perceptive and long-term memory and limited processing resources (Perfetti & Stafura, 2014). Therefore, in the case of poor comprehension, the problem can potentially arise in a particular source of knowledge that consequently affects the processes, or there may be a problem in the processes that leads to the failure of comprehension. However, Cain et al. (2001) pointed out that it could be very demanding and difficult to isolate one process from the others to test such theories.

This framework emphasises the quality of word mental representation and the role of word identification in reading comprehension, which is beyond pure word recognition (decoding skills) (Perfetti, 2017). In languages with transparent orthography, where letters are usually uniquely mapped to sounds, word processing is faster and more efficient, resulting in better mental representations of words (Perfetti, 2001, 2007). In other words, good readers who have adequate word recognition skills and can successfully integrate word meanings have more precise and richer mental lexicons, and therefore can understand the read text faster and efficiently. The LQH suggests that different aspects of the mental lexicon, such as number of different words stored, lexical diversity, and lexical retrieval influence the reading comprehension process (Perfetti & Stafura, 2014; Perfetti, 2017). So far, the LQH was narrowly tested in languages with shallow orthography. For example, Verhoeven et al. (2019) examined early lexical quality as a predictor of reading comprehension in Dutch children and found a great impact of lexicon quality in the beginning reading stages. Therefore, the role of mental lexicons and their quality could be of great importance in studying reading comprehension processes in languages with shallow orthography, especially when the phonological component (decoding skill) is under control.

DIME model

The so-called DIME model (Direct and Inferential Mediation Model of Reading Comprehension; Cromley & Azevedo, 2007) stands as an extension of the Theory of Lexical quality (Perfetti & Hart, 2002), alongside the Situation model (van Dijk & Kintsch, 1983) described in the next paragraph. The DIME model, which was developed based on a study of academic text comprehension in high school students, hypothesises that specific pressure points - background knowledge and reading vocabulary in the foreground - along with inferencing and word reading skills, and specific strategies (e.g., summarising) have a major effect on reading comprehension success. These processing skills support each other simultaneously in the process of understanding written text and can also act separately. Findings suggest that vocabulary and background knowledge have a strong direct contribution to literal reading comprehension. On the other hand, when deeper inferencing is required, these two components are mediated by inferencing skills, and inferencing is further influenced by comprehension strategies. According to this model, one's reading comprehension will be impaired if a greater than expected effort is made to activate only one of the listed components (Cromley & Azavedo, 2007). However, further research is needed to test the hypotheses of this model using other types of text, measures (combining offline and online measures, e.g., eye tracking), and different clinical groups, while considering the multidimensionality of the components (e.g., strategy, inference). Nevertheless, this model refers to the main components that could influence the processes and product of reading comprehension, and therefore must be included in the assessment.

As already mentioned, languages with transparent orthography have a clear distinction between graphemes and phonemes. Therefore, the emphasis can be placed on inferencing, monitoring, and building a mental model of the text, which are the fundamental components of the DIME model. This can probably explain how readers process and understand the text deeper, draw conclusions, and remember what they read.

Higher-level processing models (sentence- and discourse-level)

Situation and mental models

Although some models of reading comprehension tend to focus on the sentence or discourse level, it seems logical that higher-level processes are also based on information gained through the lower level of language processing (i.e.) the single word level (Hersch & Andrews, 2012). To understand a sentence or a text, a reader, after recognising a single word and extracting its adequate meaning, must integrate the meaning of all words presented in each sentence and form meaningful and mental representations, considering their syntactic and semantic structures (Perfetti & Stafura, 2014). Therefore, the overall product of comprehension is to build a mental model of the text presented and to be able to gain and learn new information from it.

At the same time, the reader must be able to adequately access high quality lexical representations (word concepts, lemas, and lexemes), syntactic relations and specific word functions, as well as knowledge of word order and inflection (e.g., verb tense and so on). Major aspects that differentiate comprehension of a single word from discourse are interpretation of figurative meaning (Hoogan et al., 2011) and context (Perfetti, 1999). Because individual words usually have literal meanings, the reader must interpret semantic-syntactic sentence context to decide whether the information in the text has a literal or figurative meaning (Hoogan et al., 2011; Tong et al., 2013). Otherwise, the retrieval of individual word meanings leads to a misunderstanding of the entire sentence/discourse concept. The important role of context was highlighted by Stanovich in 1980 through his *Interactive compensatory model*. According to this model, which emphasises the interactive approach of *bottom- up* and *top-down* processes, the context can be used in two ways: for word recognition or the prediction and understanding the meaning of the next unknown word. Furthermore, Stanovich (1980) believed that a reader is skilled at comprehension if s/he is 'context sensitive' and has a knowledge of contextual relations in addition to the automatised word recognition process. Hence, those who are skilled at comprehension will not rely on adequate context to recognise all words, but only those that are unfamiliar and new. With that, the greater part of their cognitive resources is available for the creation of meaningful and coherent text representations.

Moreover, discourse-level processes are commonly described through the *Situation model* (van Dijk & Kintsch, 1983) and the *Coherent mental model* (Johnson – Laird, 1981). These models are important because they explain the complexity of reading comprehension even more extensively by defining comprehension as a creation of mental representations of text and by emphasising the main role of inference generation. Discourse is inherently organised in such a way that all the information is presented by linguistic structures, whose purpose is to establish and maintain the cohesion and coherence. This is usually done through anaphora as a linguistic device, whose function is to organise important information in the text and reduce unimportant ones (Halliday & Hasan, 1976). Along with the aforementioned, written discourse comprehension therefore implies linking information perceived from multiple sentences into connected mental representations in order to *infer* implicit and explicit information in the text (Bowyer-Crane & Snowling, 2005), through the establishment of bridging and elaborative inferences. These mental text representations include above-mentioned meaning units – *propositions* (coherent semantic text structures), information derived from the written discourse, and the reader's background knowledge about the topic of the text. The integration of all these components is extremely important for the overall comprehension of the text. By building a mental model that constitutes both the information presented in the text and that obtained from ones' background knowledge and experience, the skilled reader can more efficiently understand the text. This can be especially viewed in languages with shallow orthography, when word recognition, i.e., decoding, is simple and automated, and the building process is faster and more efficient. It could be possible to study how readers build their

understanding about the text - how they integrate multiple pieces of information from the text with background knowledge, make inferences, and/or how this changes the model after re-reading a text.

During the discourse reading and afterwards, readers with skilled comprehension abilities should be able to make causal connections, predict further textual content, monitor the text comprehension process, and decide about the meaningfulness of the textual content that was read. But not all readers are skilled at comprehension, and to recognise them, to detect and diagnose a reading comprehension difficulty, or to test theories about underlying processes and skills, one must be aware of a construct being measured, the variety of assessment tools, as well as their strengths and weaknesses.

ASSESSING AND MEASURING READING COMPREHENSION

The knowledge about pressure points of reading comprehension, which are identified in theoretical models, brought valuable tools (measures and materials) to scientific and clinical community, for examining the strengths and weaknesses of specific populations with respect to specific skills. Furthermore, to better understand the nature and development of comprehension processes and potential causes of comprehension failures (Nation & Snowling, 1998; Cain & Oakhill, 2007; Elwér et al. 2015), studies often focus on how individuals who have difficulties with comprehension (e.g., those with *poor comprehension abilities*) differ from typical readers (Hulme & Snowling, 2011). Also, by relying on specific theories and measures, researchers can identify the main subgroups of children with reading comprehension difficulties (*e.g., poor readers, poor comprehension abilities*). However, the methods used both in research and clinical work vary considerably. Some of this variability arises from, among other things, the theoretical basis of what is being measured with a specific test and task (Cain & Oakhill, 2006; Keenan et al., 2008, 2014). Similar to the theoretical models, assessment materials have been widely developed for non-transparent orthography (mainly English language). For many

reasons discussed further in the present study, the direct implementation of formal assessment material is not justified in languages with high consistency of grapheme-phoneme correspondence.

CHALLENGES OF READING COMPREHENSION ASSESSMENTS

Given the theoretical background, it is obvious that measuring reading comprehension is much more complex than *just* assessing word recognition skill (i.e.,) decoding. To be able to make a conclusion about someone's reading comprehension skills, the assessment must include measures of various processing components with a proven role in reading comprehension processes, as well as their mutual interrelationships, as emphasised in, for example, the Reading System Framework (Perfetti & Stafura, 2014) and other above-mentioned models. The unique role of different pressure points in reading comprehension depends on how individuals are asked to express their understanding of the read stimuli (Pearson & Hamm, 2005). Therefore, it is important to know and understand the specific theoretical construct of the specific test, regardless of the purpose that it is used for: to test a theoretical view or assess one's own reading comprehension skills for diagnostic purposes.

Another challenge of reading comprehension assessments is that some tests are designed and marked as measures of reading comprehension, but in reality, they depend heavily on another skill. For example, de Jong & van der Leij (2002) found a high correlation (0.70) between some components such as word reading (decoding) and comprehension, especially in the early years of reading and reading comprehension development. If the specific aim is to assess *reading comprehension*, then this finding points to the need for a clear separation of these two components in the assessment process (Tunmer & Chapman, 2012). In contrast, reading comprehension difficulties may go unnoticed (e.g., *poor comprehension abilities, hyperlexia, and so on;* Oakhill et al., 2003; Bowyer-Crane & Snowling 2005; Saldaña & Frith, 2007). Additionally, listening comprehension plays a dominant role in reading compre-

hension and should therefore be included as an independent measure in the assessment process (Catts et al., 2005). This is stated in accordance with the SVR theory (Gough & Tunmer, 1986). Furthermore, the contribution and interrelations of the various components of reading comprehension may vary throughout development (Cain et al., 2017). In the period of *learning to read* (early stages), reading comprehension success relies heavily on phonological skills (decoding), while with age and transitions to the period of *learning through reading,* other language variables (such as vocabulary, syntax, inference generation) play a greater role. Therefore, it is important to consider the age range for the implementation of certain measures and development of specific tests for reading comprehension. Given the complexity of the nature of reading comprehension, another potential issue is that the test assesses comprehension only at one level, e.g., sentence level. Some children may have considerable difficulties reading a discourse and answering related questions, in contrast to their results on measures of single sentence comprehension. The higher-level processing models described earlier, i.e., the situation and mental models, describe and confirm this very fact. Hence, by examining the processes at a higher level, researchers or clinicians could gather more data on the integration of multiple components and the associated cognitive effort. Most studies and assessment methods had focused on measuring single word or single sentence comprehension. Part of the reason for this was the problem of controlling for various covariates that could potentially mask the ultimate performance outcome (results). However, there is a lack of data on the effect of these at a level higher than an individual word (when lexical units enter the syntactic constructions that determine the discourse).

Additional factors that contribute significantly to the variability of assessment methods are the nature of questions being asked (literal vs. inferential) and the response format (e.g., multiple choice, open-ended questions, and so on), silent vs. reading aloud, as well as the length, structure and lexical complexity of the text, and the required cognitive load (Cain & Oakhill, 2006;

Deane et al., 2006; Basaraba et al., 2013; Keenan et al., 2014).

Recent studies also point out the importance of individually assessing the cluster of skills that have been found to have an impact on reading comprehension, such as syntactic and phonological skills, vocabulary breadth and depth, listening comprehension, general cognitive resources, and working memory (Nation, 2005; Cain & Oakhill, 2007). In reviewing the existing literature, it is possible to highlight specific skills that provide a reliable distinction between those who are skilled at comprehension and those who are not (high vs. low comprehension competence). These are mainly higher-level language processing skills that become strongly important with age, especially during the transition into the phase of learning through reading. The most prominent ones are inference and integration skills (Cain & Oakhill, 1999; Cain et al., 2001; Oakhill, 1984), comprehension monitoring (Oakhill & Yuill, 1996; Ehrlich & Remond, 1997; Oakhill et al., 2005), story structure understanding (Cain & Oakhill, 1996), and working memory (Daneman & Carpenter, 1980; Yuill et al., 1989).

Another very important item to consider is the fact that most of the tests for assessing reading comprehension are developed often exclusively for the English language. English is a language with deep orthography and does not rely on consistent relations between graphemes and phonemes, as opposed to languages with shallow and transparent orthography (e.g., Croatian, Spanish, and so on) Therefore, tests developed in languages with non-transparent orthography are not directly applicable in languages with transparent orthography because of differences in phonological and orthographical language structure (Seymour et al., 2003; Zaretsky et al., 2009). Also, the possible differences in syntactic and semantic complexity among languages must be considered. Tests developed for the English language may have some inappropriate tasks for assessing reading comprehension in languages with transparent orthography, where decoding skills are simpler and develop faster (Ziegler et al., 2010). This could lead to

misinterpretation of the results achieved on such tests.

Thus, assessing reading comprehension represents a challenging piece of the riddle that must be carefully assembled, given the complexity of the process itself. Also, some of the components that contribute to it are hard to observe and measure, and therefore, implement in tests. Nevertheless, the importance of assessing these components with the purpose of collecting data about one's reading comprehension skills remains unquestionable. Given that an assessment is usually divided into formal (standardised; based on the use of standardised tests) and informal (alternative; includes criterion- based measures) assessments, both types will be presented in the present review, in the context of reading comprehension measurement methods. The description of both types of assessments is important because their combination provides the most comprehensive representation of someone's ability to understand written language.

METHODS AND MEASURES OF READING COMPREHENSION

Informal reading comprehension measures

Specific methods for assessing reading comprehension have changed significantly in the 20th and 21st century (Pearson & Hamm, 2005) and the primary reason for this is the recognition that reading comprehension is an essential component of overall reading competence. The assessment of reading comprehension mostly refers to standardised tests. However, it has also been expanded to experimental procedures on different processing levels. Such informal (experimental, criterion- based) assessment methods are of great value both in research and clinical work, as they imply the creation and manipulation of the specific stimuli, to answer specific hypothesis questions and to make a clinical decision (Murray & Coppens, 2013).

Over time, it became clear that the principles of comprehension methods for the listening paradigm are transferable to the reading paradigm,

with some minor differences such as the phonological decoding skills required for reading. In both paradigms, children or adults have to respond to a presented stimulus (word, sentence, discourse) by selecting the correct picture, word/sentence, acting out the target stimulus, verifying the sentence, retelling, preferentially looking at the target, or assessing comprehension through functional neuroimaging methods, judgement methods (grammatically or truth-value judgements), and/or answering questions in different response formats (Cain & Oakhill, 2006; Keenan et al., 2008; Paris & Hamilton, 2014; Ambridge & Rowland, 2013). Some of these methods can be used in real time (such as online methods - eye tracking, neuroimaging), while others are implemented after reading a stimuli. It depends on the specific purpose and aims of the specific study.

If the specific research question and aim is to determine how reading comprehension processes of a target stimulus occur in real time, then online methods such as eye tracking, functional neuroimaging methods (fMRI), or electroencephalography (EEG) are the best choices. The advantage of these methods is that they provide data on what is happening in language processing while reading the presented stimulus and when, why, and how failure occurs. For example, Matić et al. (2018) conducted a study aiming to understand the effect of phonological versus lexical knowledge on lexical processing of children with dyslexia in Croatian. They found that both children with dyslexia and typical language development (TLD) do not rely mostly on lexical knowledge during lexical processing, but children with TLD use phonological knowledge to help them process. This is a very interesting finding in the context of reading comprehension and orthographic transparency, given that Croatian is a language with deep orthographic transparency. Furthermore, eye tracking presents a newer, modern, and innovative method in the reading paradigm, whose ''previous'' version in the listening paradigm is looking while listening (Rayner & Reichle, 2010; Ambridge & Rowland, 2013). By observing the participant's eye movements while reading a particular stimulus (e.g., sentences with anaphoric relations), one can mea-

sure individual reaction time, how long the participant dwells on the target word before moving to the next word, or how many regressions the participant took during reading, and so on. Additionally, Palmović et al. (2018) conducted an eye tracking study to examine and compare the anaphora resolution strategies through visual paradigms in children and adults in Croatian. Furthermore, studies implementing eye tracking methods have shown that the more difficult the text, the more likely it is that more regressive eye movements will be observed. This would indicate that a reader did not understand part of the discourse adequately, and therefore needs to go back in the text to re-read it (Rayner et al., 2006)

In general, researchers can observe and collect data on what children or adults do during the activation of reading comprehension processes and where and why some of them have difficulties. For example, Hessel et al. (2021, corrected manuscript) studied comprehension monitoring in children and found longer gaze durations (i.e., the time they spent reading inconsistent words), suggesting slower lexical access in the inconsistent condition. Furthermore, fMRI and EEG are somewhat similar when a target stimulus is presented, except that they measure brain response and local activity/changes in a particular brain area rather than eye movements. By measuring these variables, valid data can be obtained about the potential difficulties in processing certain linguistic constructions (Ozernov-Palchik et al., 2021).

On the other hand, valuable data about one's reading comprehension skills can be obtained by using offline methods, for example, by measuring the extent to which someone has understood a particular text. Unlike online methods, these offline methods of reading comprehension do not require special and expensive equipment such as hardware, devices, and statistical expertise. As already mentioned above, it is possible to measure comprehension through listening and reading paradigms. Similarly, one can measure two related constructs independently: 1) listening comprehension, which is inevitably important for reading comprehension (SVR; Gough & Tunmer, 1986), and 2) written discourse (text) comprehen-

sion. The research methods most often mentioned in studies aimed at measuring comprehension are: picture selection task, act out task, retelling, judgement methods, and so on.

Some of these tasks are less cognitively demanding (e.g., picture selection task), while others can require more cognitive effort (e.g., act out task, judgement methods). For example, the picture-choice or -selection task is often used to assess comprehension of passive constructions (actional vs. non-actional verbs) and inferences (Cain et al., 2001; Nation et al., 2007). In this task, participants need to select the correct picture (static or dynamic) that credibly represents the unambiguous stimulus, after listening or reading a target sentence. Other tasks that require participants to do more and are somewhat more demanding are the act out task and retelling. In the act out task, which is commonly used to assess the role of syntax awareness in comprehension, participants need to perform the target action (often with figures) in relation to the referents. Retellings are often defined as time-consuming and demanding, but previous research has shown their value as an assessment tool (Cao & Kim, 2021). How this method is used depends on the theoretical starting point. It could be used to calculate the percentage of semantic propositions retold by the reader (Kintsch & van Dijk, 1978), story elements (Stein & Glenn, 1979), grammatical or content words. Thus, before deciding on this type of method, one must be aware of the chosen theoretical point of view and the way points are assigned. Reed and Vaughn (2012) showed in their systematic review that retelling as a method correlates moderately with other standardised measures of reading comprehension and the correlation coefficient between retelling and phonological skills (e.g., decoding) decreases with age. Additionally, students with reading difficulties expressed more difficulty in retelling and needed additional support.

As far as the measurement of text comprehension is concerned, the tasks that stand out had an emphasis on the special control of making literal versus inferential inferences. In addition, these types of comprehension methods also differ in relation to cognitive demands (Figure 1).

Figure 1. *Type of text comprehension measurement methods, arranged from least to most cognitively demanding (Cain & Oakhill, 2006)*

Cloze type of questions (fill in the blank), in which participant completes a series of one-word sentences (up to 5 offered) that can be presented independently or embedded in a coherent text, are found to be significantly affected by decoding and syntactic skills (Francis et al., 2005). To ensure that the comprehension of the written text is being measured, the stimuli basis of this task must be a coherent text with the possibility of choosing the answer at the sentence and text level, and the accompanying measure of single word reading to control decoding skills.

Furthermore, the true/false sentence recognition task is more likely to require recall of memorised literal information than the sentence verification task (SVT), with lower processing demands. After reading the text, the participant is required to simply answer a question with yes/no. However, this does not reflect a true measure of inferencing, as children only need to recognise whether a target sentence matches with the one in the paragraph/text (Cain & Oakhill, 2006). On the other hand, the SVT implies memorising the meaning extracted from the text (but not the literal words). In this task, four sentences are presented, in which the meaning, content words, and syntax are manipulated, with one being a distractor (filler) (Royer et al., 1979). After reading, the reader needs to determine which of the four sentences have the same meaning as the one in the target paragraph. Thus, the semantic component of comprehension matters more. Moreover, it should be emphasised that this method is significantly influenced by reading skills (Carlisle, 1989; Ambridge & Rowland, 2013; Collins et al., 2018).

Multiple choice (or forced choice) questions are widely used in reading comprehension assessment studies (Magliano et al., 2007) and are recognised as being associated with higher cognitive demands in comparison to previously mentioned tasks. Even though multiple-choice questions do not require a verbal response, participants need to compare the responses presented and select the correct one, partially relying on processes of automatic retrieval or familiarity of information. The task minimally includes three sentences that represents possible answers to questions about the read text that can vary in length, availability, or type of questions (literal vs. inferential). The decision of whether the text will be long, available after reading, how rich the cues (target answers) will be, or the number of inferential questions depends on whether the researcher tends to control the effect of working memory (Daneman & Hannon, 2001; Cain et al., 2004). Obviously, the longer the text and the higher number of inferential questions versus literal, with the inability of looking back at the source text, places more emphasis on working memory. Therefore, they are more cognitively demanding to process, but research showed they are sensitive and reliable (Carlson et al., 2014).

Turning to an even more demanding task, open-ended questions often require more active and generative processing, as well as verbal responses. Therefore, the comprehension skills of children with more expressive language difficulties can be obscured if the listening comprehension component is not assessed in addition to the reading paradigm (Woolley, 2008). Even though there is no clear distinction between (cognitive) processes that are tapped by multiple choice vs. open-ended questions (Campbell, 1999), it is known that these types of questions require more cognitive effort but provide a considerable amount of information about one's inferencing skills, working memory, and the impact of additional background factors on reading comprehension (Cain & Oakhill, 2006). Questions are commonly developed to examine a participant's literal and

inferential understanding of the passages, with minor amount of the retrieval cues available to the participant, in comparison to the multiple-choice task (Magliano et al, 2007, Graesser et al., 2010).

If the aim is to assess a reader's higher-level skills, such as comprehension monitoring, a typical task that would be used is to read a text with logical (internal) or empirical (external) inconsistencies, ambiguities, or long words and then determine whether the story makes sense. However, Ruffman (2013) argues that if the purpose is to clearly capture the comprehension monitoring, it is important to consider the instructions given to the participants as well as the types of material (text) used, explaining that this could affect more of participant's use of general knowledge rather than monitoring skills.

Finally, when measuring reading comprehension, one should be guided by the findings of Keenan et al. (2008) regarding the methodologies used. According to their study, the greatest impact of decoding skills can be found in cloze or multiple-choice tasks and in short text passages, whereas open-ended questions and longer passages are more influenced by oral comprehension. This finding is important in deciding regarding which task and what to use when assessing reading comprehension in languages with transparent orthography.

Formal reading comprehension measures

The more formal measures that involve interpretation of results and comparison of normative data are standardised tests. Studies that have used standardised tests can provide information about variability in reading comprehension. There are several standardised tests on the market that measure reading comprehension. This review presents some of the widely used tests, developed exclusively for English-speaking children. Even though these tests cannot be directly applied in other languages, they provide researchers and developers a valuable structure and foundation for how the material should look like for formal reading comprehension measurement (for instance, for Croatian). A brief list of the rest of the well-known standardised reading comprehension tests is provide in Table 1.

Table 1. *Additional examples of reading comprehension assessment tests (Fletcher, 2006; Westerveld, 2009)*

Reading comprehension assessment	Procedure and format
Wechsler Individual Achievement Test (WIAT; Wechsler, 1992)	• short passages of 2 to 3 sentences • silent reading • open-ended questions, both literal and inferential • short-answer approach
Gates–MacGinitie Reading test (MacGinitie et al., 2000)	• passages up to 15 sentences • silent reading • multiple choice task
Gray Oral Reading Test - Third Edition (Wiederholt & Bryant, 1992)	• passages of about 6 to 7 sentences • reading aloud • multiple choice task (without viewing the passage)
Passage Comprehension from Woodcock-Johnson-III (Woodcock et al., 2001)	• sentences or short passages • cloze procedure
Diagnostic Assessment of Reading Comprehension - DARC (August et al., 2002)	• an experimental test • passages of three sentences • control of the required level of decoding • manipulation of components related to text memory and inferencing, as well as knowledge access and integration • measures comprehension using linguistic discourse methods
TORCH: *Test of Reading Comprehension (Mossenson et al., 1987)*	• 12 passages • silent reading • modified cloze tasks
WRMT–R: *Woodcock Reading Mastery Test – Revised*: Passage Comprehension (Woodcock, 1998)	• one/two sentences or a short passage • silent reading • cloze task

As mentioned earlier, each of these tests has its own strengths and limitations that researchers and clinical professionals need to be aware of.

For example, one of the widely known tests is the Neale Analysis of Reading Ability (NARA-II): NFER-Nelson (Neale, 1997), which is commonly used to obtain normative data of children's reading skills (ages six to twelve years), their strengths and weaknesses in reading, and one's interest in reading. It measures reading accuracy and speed, as well as reading comprehension. Comprehension is measured by answering short questions (open-ended format) about main story theme and events, using both literal and inferential questions, thus obtaining information about the participant's inferencing skills, working memory, and so on. It includes both short and extended passages. Thus, it assesses both sentence and text-level comprehension, and the overall achievement is the total number of correct answers. In addition to strengths of the test, there are some limitations related to written text comprehension such as the possibility to answer the question *accurately* just by directly referring to the text (without the need to extract and integrate meaning), as well as some limitations related to reading measures (Cain & Oakhill, 2006; Keenan et al., 2008). The same limitations are found in the York Assessment of Reading for Comprehension (YARC) Primary: GL Assessment (Snowling et al., 2012). However, the advantage of this test (ages four to eleven years) is that it assesses comprehension of words, sentences, and texts, as well as various types of inferences, using open-ended questions. The two previously mentioned tests are used for individual assessment of skills; the next two are designed for group assessments. The Suffolk Reading Scale (ages 6 to 14; mean 11 years; Hagley, 1987) and the Group Reading Test (GRT 2; age 6-14 years; Macmillan Unit, 2000) were designed to primarily assess reading comprehension in a group setting, through the silent reading of a simple sentence or a simple sentence and a short passage. Both tests include a cloze and a multiple-choice response format. The major disadvantages of these tests are that they do not assess text level strategies, global coherence, and rely heavily on decoding skills.

At last, one of the newest, more integrative instruments for assessing language and literacy skills is the Test of Integrated Language & Literacy Skills (TILLS; Nelson et al., 2015). The test includes 15 subtests and is intended for the children aged 6-18 years. The scope of this test is reflected in the extended principles of Simple View of Reading theory, which emphasises the interrelationships between oral and written language modalities and different language level processing. That is, it measures and compares reader's skills at different language levels – sound/word (lexical and sub-lexical knowledge) and sentence/discourse (processing of meaning of information and knowledge of sentence structure) – in both oral and written language, unlike other assessment instruments. Thus, it assesses behavioural performance in different modalities (listening, speaking, reading, writing) and at different language levels, with additional tasks to assess memory across these language levels.

Although the tests mentioned above were not directly developed for languages with transparent orthography, they can provide a valid basis for constructing tests in those other languages if certain components are taken into account (linguistic characteristics). The above will be discussed on the example of reading comprehension assessment in Croatian, as a language of deep orthographic transparency.

ASSESSMENT OF READING COMPREHENSION IN CROATIAN AS AN EXAMPLE OF TRANSPARENT ORTHOGRAPHY LANGUAGE

In most languages other than English, the selection of reading comprehension assessment tools is limited (e.g., in Spanish: PROLEC-R; Cuetos et al., 2007; LEE; Defior et al., 2006) or non-existent, such as in shallow orthography languages (particularly in minority languages and those from areas with smaller populations, such as Croatia). Currently, there is no standardised tool to assess the reading and reading comprehension of school-aged children in Croatian: a few tools are in the process of being developed (mainly for early grade school students). Since the tests de-

scribed above are standardised mainly on the population of English-speaking children, they cannot be used directly for Croatian-speaking children. Even though there are some general language features, Croatian and English differ in their specific language typology, including their lexicon and grammar, but mainly in phonology and orthographic features. Consequently, these differences may influence measurement outcomes. To be able to compare assessment results across languages for both clinical and research purposes, the development of assessment tools for a particular language requires more than just a simple translation of literal words from the existing one – it requires test standardisation, norming, and validation (corroboration that both versions of assessments are equivalent and are built to measure the same concepts). A fundamental step of the adaptation process is to select adequate material to be adapted while consider its potential limitations, and a construct to be measured based on the selected theoretical framework. Furthermore, it is essential to take into consideration specific linguistic and psycholinguistic variables and parameters of the target languages underlying the test items, as well as cultural differences between language communities and equivalence of the scores (Hambleton et al., 2012). So far, there is a lack of evidence of factors affecting the reading comprehension processing skills in children with typical and atypical language development in Croatian. It is important to collect such data for the purpose of adaptation. Specifically, previous research in Croatian language was primarily focused on the role of decoding skills in the reading processes, while the process of reading comprehension itself and other components with potential influence were not the main subjects of research studies. Till now, it is known that the (meta)phonological aspect plays an important role in emergent reading skills of Croatian children (e.g., Zaretsky et al., 2009; Blaži et al., 2011; Ivšac Pavliša & Lenček, 2011; Lenček, 2012; Matić et al., 2018; Kelić, 2019; Kelić et al., 2021). Hence, it is obvious that an assessment tool in Croatian for overall reading competence (including reading comprehension) should include, among others, word reading

tasks (on different processing levels) to determine phonological skills, and phonological representations (also in accordance with the SVR theory), as observed in, for example, NARA-II or TILLS. However, considering the orthographic transparency of the language and its characteristics, it is inevitable to study, for example, how lexical representations, vocabulary, grammar, and other components affect one's understanding of what is being read, while controlling for decoding at the same time.

Referring to theoretical insights (e.g., Lexical Quality Hypothesis), the adaptation of target items of reading comprehension tests also includes choosing the adequate item length (word, sentence, text), lexical and syntactical complexity of the material, text type, and the type of task that will be used to assess the success of the task performance. Hence, the above-mentioned standardised tests (such as NARA-II and TILLS, which are quite comprehensive) represent a good basis for designing a language-specific assessment tool (e.g., for Croatian).

In Croatia, there is a tendency among speech and language therapists to primarily use traditional, norm-referenced assessment measures in clinical work. Despite the recognised need and the effort invested in adapting and developing standardised assessment tools for language (comprehension) measurement, there is still a lack of suitable standardised tests in Croatia. As previously mentioned, informal assessment methods represent a valuable source of data about one's reading comprehension skills, and therefore should be used alongside with standardised tools. In fact, a combination of both formal and informal methods stands out as the best practice for obtaining a complete clinical picture of the specific disorder (ASHA, 2004). However, the non-norm-referenced materials created by clinicians that are currently in use are not always valid, reliable, or representative enough, and they do not always measure the target construct with certainty. This could led to a serious problem – the misinterpretation of the participant's target skills and clinical misdiagnosis. Therefore, when planning and creating assessment materials, it is important to

follow the main principles in developing experimental measurement methods (mentioned in the previous chapter) based on a chosen theoretical framework.

CONCLUSION

This paper provides a comprehensive review of theoretical models and the methodology of assessing reading comprehension and related component skills. Without sufficient knowledge and understanding regarding the models that underpin various reading comprehension measures, there is a possibility of measuring an incorrect theoretical construct, or the correct construct, but with the wrong approach. Additionally, the importance of language typology in the selection and application of tests should not be overlooked. Consequently, this can lead to the misinterpretation of background processes that are being activated and one's written comprehension skills. Finally, it can lead to incorrect implications with respect to instructions that should be implemented in future therapy settings. This is especially important in the context of the development of a new assessment tool in languages with transparent orthography, where such tools do not exist (e.g., Croatian).

Decoding is an important phonological skill that refers to the ability to convert graphemes into phonemes and vice versa, while comprehension refers to the ability to process and interpret the meaning of a text after reading. These two components form the essence of the SVR theory. However, in languages with transparent orthography, such as Croatian, the decoding process is relatively simple (1:1) and in the context of typical language development and cases other than, for example, dyslexia, they should not represent a problem. This means that good and efficient language processing plays a significant role in understanding what is being read. Therefore, SVR theory can represent a good starting point for studying the process of reading comprehension in languages with shallow orthography and it can provide a direction towards expanding this framework by studying the role of language processing in reading comprehension. Therefore, future research

should focus on considering the characteristics of languages with transparent orthography.

On the other hand, other theoretical models such as the Lexical Quality Hypothesis (Perfetti & Hart, 2002), DIME (Cromley & Azevedo, 2007), and the Mental Model (van Dijk & Kintsch, 1983) are potential more applicable to studying reading comprehension in languages with transparent orthography. Given that research in languages with transparent orthography indicates that decoding skills do not sufficiently explain the variance in reading comprehension, LQH represents a good framework through which the role of the lexicon and the quality of stored mental representations in reading comprehension can be studied. Furthermore, starting from the assumptions of the Mental Model, it is possible to study how readers create mental models of the text, what levels they go through during this process (within the framework of the DIME model,) and how this affects understanding.

As for the assessment itself, it can be concluded that formal reading comprehension measures (as well as informal) in languages with non-transparent orthography, such as English, are widely developed and tend to place more emphasis on decoding skills, including the correspondence between letters and sounds, which is often more complex and unpredictable. These tests include tasks that assess phonological awareness, word recognition, and phonemic decoding. Additionally, comprehension tasks may be more challenging and require more inferencing and critical thinking skills due to the increased difficulty of decoding the text.

On the other hand, there is a lack of standardised tests developed for languages with transparent orthography. However, these tests could be more straightforward than those in languages with non-transparent orthography. This is because in languages with transparent orthography, the relationship between letters and sounds is more consistent and predictable, which makes it easier for readers to decode words accurately and quickly. As a result, reading comprehension assessment measures in these languages should place emphasis on lexical and semantic knowledge, and

higher-level comprehension skills. Finally, while transparent orthographies may require less emphasis on decoding, it continues to be a necessary skill for reading comprehension and should be included in such assessments to obtain a complete understanding of reading proficiency, as well as to identify potential areas of difficulty for targeted support.

The final purpose of this review paper was to encourage researchers in languages other than English to study reading comprehension processes, both in language-specific domains and general domains, as well as to conduct cross-linguistic studies to further clarify the nature of this process to provide a set of valuable methods and tools for reading comprehension assessment.

REFERENCES

Ambridge, B., & Rowland, C. F. (2013). Experimental methods in studying child language acquisition. *Wiley Interdisciplinary Reviews: Cognitive Science*, *4*(2), 149-168.

American Speech-Language-Hearing Association. (2004). Preferred Practice Patterns for the Profession of Speech-Language Pathology [Preferred Practice Patterns]. Available from www.asha.org/policy/. doi:10.1044/policy.PP2004-00191

August, D., Francis, D. J., & Calderón, M. (2002). Diagnostic assessment of reading comprehension (DARC). *Washington, DC: Center for Applied Linguistics.*

Babashamsi, P., Bolandifar, S., & Shakib, N. (2013). Various models for reading comprehension process. *International Journal of Applied Linguistics and English Literature*, *2*(6), 150-154.

Basaraba, D., Yovanoff, P., Alonzo, J., & Tindal, G. (2013). Examining the structure of reading comprehension: Do literal, inferential, and evaluative comprehension truly exist?. *Reading and Writing*, *26*(3), 349-379.

Barnett, M. A. (1988). Teaching reading strategies: How methodology affects language course articulation. *Foreign Language Annals*, *21*(2), 109-119.

Blaži, D., Buzdum, I., & Kozarić-Ciković, M. (2011). Correlation of reading skill success with some aspects of phonological development. Croatian Review of Rehabilitation Research, 47(2), 14-25.

Bishop, D. V., & Snowling, M. J. (2004). Developmental dyslexia and specific language impairment: Same or different?. *Psychological bulletin*, *130*(6), 858.

Bowyer- Crane, C., Snowling, M.J. (2005). Assessing children's inference generation: What do tests of reading comprehenson measure?. British journal of educational psychology, 75 (2), 189-201.

Cain, K., & Oakhill, J. (1996). The nature of the relationship between comprehension skill and the ability to tell a story. *British Journal of developmental psychology*, *14*(2), 187-201.

Cain, K., Oakhill, J.V. (1999). Inference making ability and its relation to comprehension failure in young children. Reading and writing: AN interdisciplinary Journal, 11, 489-503.

Cain, K., Oakhill, J. V., Barnes, M. A., & Bryant, P. E. (2001). Comprehension skill, inference-making ability, and their relation to knowledge. *Memory & cognition*, *29*(6), 850-859.

Cain, K., Oakhill, J., & Lemmon, K. (2004). Individual differences in the inference of word meanings from context: The influence of reading comprehension, vocabulary knowledge, and memory capacity. *Journal of educational psychology*, *96*(4), 671.

Cain, K., Oakhill, J. (2006). Profiles of children with specific reading comprehension difficulties. British Journal of Educational Psychology, 76 (4), 683–69.

Cain, K., Oakhill, J. (2007). Reading comprehension difficulties: Correlates, causes and consequences. In children's comprehension problems in oral and written language, New York: Guilford Press, 41–75.

Cain, K., Compton, D. L., & Parrila, R. K. (2017). Introduction to big questions. *Theories of Reading Development*, *15*, 7.

Campbell, J. R. (1999). Cognitive processes elicited by multiple-choice and constructed-response questions on an assessment of reading comprehension. Temple University.

Cao, Y., & Kim, Y. S. G. (2021). Is retell a valid measure of reading comprehension?. *Educational Research Review*, *32*, 100375.

Carlisle, J. F. (1989). The use of the sentence verification technique in diagnostic assessment of listening and reading comprehension. *Learning Disabilities Research*.

Carlson, S. E., Seipel, B., & McMaster, K. (2014). Development of a new reading comprehension assessment: Identifying comprehension differences among readers. *Learning and Individual Differences*, *32*, 40-53.

Catts, H. W., Hogan, T. P., & Adlof, S. M. (2005). Developmental changes in reading and reading disabilities. In: H. W. Catts, A. G. Kamhi (Eds.), *The connections between language and reading disabilities* (pp. 38-51). Psychology Press.

Clarke, P. J., Henderson, L. M., & Truelove, E. (2010). The poor comprehender profile: Understanding and supporting individuals who have difficulties extracting meaning from text. *Advances in child development and behavior*, *39*, 79-129.

Collins, A. A., Lindström, E. R., & Compton, D. L. (2018). Comparing students with and without reading difficulties on reading comprehension assessments: A meta-analysis. *Journal of learning disabilities*, *51*(2), 108-123.

Cromley, J. G., & Azevedo, R. (2007). Testing and refining the direct and inferential mediation model of reading comprehension. *Journal of educational psychology*, *99*(2), 311.

Cuetos, F., Rodríguez, B., Ruano, E. & Arribas, D. (2007). PROLEC-R. Evaluación de los procesos lectores –Revisada. TEA

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of verbal learning and verbal behavior*, *19*(4), 450-466.

Daneman, M., & Hannon, B. (2001). Using working memory theory to investigate the construct validity of multiple-choice reading comprehension tests such as the SAT. *Journal of Experimental Psychology: General*, *130*(2), 208.

Davoudi, M., & Moghadam, H. R. H. (2015). Critical review of the models of reading comprehension with a focus on situation models. *International Journal of Linguistics*, *7*(5), 172-187.

Deane, P., Sheehan, K. M., Sabatini, J., Futagi, Y., & Kostin, I. (2006). Differences in text structure and its implications for assessment of struggling readers. *Scientific Studies of reading*, *10*(3), 257-275.

Defior, S., Fonseca, L., Gottheil, B., Aldrey, A., Jiménez Fernández, G., Pujals, M. et al. (2006). LEE. Test de lectura y escritura en Español. Paidós

de Jong, P. F., & van der Leij, A. (2002). Effects of phonological abilities and linguistic comprehension on the development of reading. *Scientific studies of Reading*, *6*(1), 51-77.

Elwér, Å., Gustafson, S., Byrne, B., Olson, R. K., Keenan, J. M., & Samuelsson, S. (2015). A retrospective longitudinal study of cognitive and language skills in poor reading comprehension. *Scandinavian journal of psychology*, *56*(2), 157–166.

Ehrlich, M. F., & Remond, M. (1997). Skilled and less skilled comprehenders: French children's processing of anaphoric devices in written texts. *British journal of developmental psychology*, *15*(3), 291-309.

Fletcher, J. M. (2006). Measuring reading comprehension. *Scientific studies of reading*, *10*(3), 323-330.

Florit, E., & Cain, K. (2011). The simple view of reading: Is it valid for different types of alphabetic orthographies?. *Educational Psychology Review*, *23*, 553-576.

Francis, D. J., Fletcher, J. M., Catts, H. W., & Tomblin, J. B. (2005). Dimensions affecting the assessment of reading comprehension. In: G. Paris, S. A. Stahl (Eds.), *Children's reading comprehension and assessment* (pp. 387-412). Routledge.

Goodman, K. (1967). Reading: A psycholinguistic guessing game. *Journal of the Reading Specialist, 6*(4), 126-135.

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and special education*, *7*(1), 6-10.

Graesser, A. C., Ozuru, Y., & Sullins, J. (2010). What is a good question? In M. McKeown & L. Kucan (Eds.), *Bringing reading research to life* (pp. 112–141). New York, NY: Guilford Press.

Hagley, J. F. (1987). The Suffolk Reading Scale. Windsor: NFER-Nelson.

Halliday, M. A. K., & Hasan, R. (2014). *Cohesion in english*. Routledge.

Hambleton, R. K., Merenda, P. F., & Spielberger, C. D. (Eds.). (2004). *Adapting educational and psychological tests for cross-cultural assessment*. Psychology Press.

Hersch, J., & Andrews, S. (2012). Lexical quality and reading skill: Bottom-up and top-down contributions to sentence processing. *Scientific Studies of Reading*, *16*(3), 240-262.

Hessel, A. K., Nation, K., & Murphy, V. A. (2021). Comprehension monitoring during reading: An eye-tracking study with children learning English as an additional language. *Scientific Studies of Reading*, *25*(2), 159-178.

Hoogan, T., Sittner Bridges, M., Justice, L.M., Cain, K. (2011). Increasing higher level language skills to improve reading comprehension. *Focus on exceptional children, 44*, 1-20.

Hoover, Wesley A. & Philip B. Gough. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal, 2*(2), 127-160.

Hulme, C., & Snowling, M. J. (2011). Children's reading comprehension difficulties: Nature, causes, and treatments. *Current Directions in Psychological Science*, *20*(3), 139-142.

Jescheniak, J.D., Levelt, J.M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of experimental psychology: Learning, memory and cognition, 20,* 824-843.

Johnson-Laird, P. N. (1981). Comprehension as the construction of mental models. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, *295*(1077), 353-374.

Keenan, J., Betjemann, R., & Olson, R. (2008). Reading comprehension tests vary in the skills they assess: Differential dependence on decoding and oral comprehension. *Scientific Studies of Reading*, *12*, 281–300.

Keenan, J. M., & Meenan, C. E. (2014). Test differences in diagnosing reading comprehension deficits. *Journal of learning disabilities*, *47*(2), 125-135.

Keenan, J. M., Hua, A. N., Meenan, C. E., Olson, R. K., Pennington, B. F., Willcutt, E. G. (2014). Issues in identifying poor comprehenders. L'année Psychologique/ *Topics in Cognitive Psychology, 114*(4), 753–777.

Kelić, M. (2019). From phonemic segmentation and blending to spoonerisms – deficits in phonological awareness in children with dyslexia in Croatian. *Govor*, *36*(2), 139-166.

Kelić, M., Zeba, M. Z., & Kraljević, J. K. (2021). Reading Predictors in Croatian: Contribution of (Meta) Phonological Variables. *Psychological Topics*, *30*(2), 161-184.

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85*(5), 363-394.

Kintsch, W., & Kintsch, E. (2005). Comprehension. In *Children's reading comprehension and assessment* (pp. 89-110). Routledge.

Kintsch, W. & Rawson. KA (2005). Comprehension. *The science of reading*, 209-226.

Lenček, M. (2012). Assessment of Dyslexia in Croatian: Some Characteristics of Reading and Writing in Students with Dyslexia. *Croatian Review of Rehabilitation Research*, *48*(1), 11-26.

Levelt, W. J. (1993). Lexical selection or how to bridge te major rift in language processing.

Levelt, W. (1999). Producing spoken language. *The neurocognition of language*, 83-122.

Matić, A., Coumel, M., & Palmović, M. (2018). Lexical processing of children with dyslexia: An eye-tracking adaptation of the Reicher-Wheeler task. *Govor*, *35*(1), 27-52.

MacGinitie, W. H., MacGinitie, R. K., Maria, K., & Dreyer, L. G. (2000). *Gates-MacGinitie reading tests (4th ed.):* The Riverside Publishing Company.

Macmillan Unit. (2000). *The Group Reading Test II*. Windsor: NFER Nelson.

Magliano, J. P., Millis, K., Ozuru, Y., & McNamara, D. S. (2007). A multidimensional framework to evaluate reading assessment tools. *Reading comprehension strategies: Theories, interventions, and technologies*, 107-136.

Martin, C. (2011). York Assessment of Reading for Comprehension (YARC).

McNamara, D. S., & Magliano, J. (2009). Toward a comprehensive model of comprehension. *Psychology of learning and motivation*, *51*, 297-384.

Mossenson, L., Hill, P., & Masters, G. (1987). *Test of reading comprehension TORCH.* Australian Council for Educational Research.

Murray, L., & Coppens, P. (2013). Formal and informal assessment of aphasia. *Aphasia and related neurogenic communication disorders*, 66-91.

Nation, K., & Snowling, M. J. (1998). Semantic processing and the development of word-recognition skills: Evidence from children with reading comprehension difficulties. *Journal of Memory and Language*, *39*(1), 85-101.

Nation, K., & Snowling, M. J. (2004). Beyond phonological skills: Broader language skills contribute to the development of reading. *Journal of research in reading*, *27*(4), 342-356.

Nation, K. (2005). Children's Reading Comprehension Difficulties.

Nation, K., Snowling, M. J., & Clarke, P. (2007). Dissecting the relationship between language skills and learning to read: Semantic and phonological contributions to new vocabulary learning in children with poor reading comprehension. *Advances in Speech Language Pathology*, *9*(2), 131-139.

Neale, M. D. (1997). *Neale analysis of reading ability-revised (NARA-II)*. Windsor, UK: NFER.

Nelson, N. W., Plante, E., Helm-Estabrooks, N., & Hotz, G. (2015). *Test of Integrated Language and Literacy SkillsTM (TILLS™)*. Brookes Publishing Co., Inc.

Oakhill, J. (1984). Inferential and memory skills in children's comprehension of stories. *British journal of educational psychology*, *54*(1), 31-39.

Oakhill, J. & Yuill, N. (1996). Reading comprehension difficulties. *Children's comprehension problems in oral and written*, 41.

Oakhill, J. V., Cain, K., & Bryant, P. E. (2003). The dissociation of word reading and text comprehension: Evidence from component skills. *Language and Cognitive Processes, 18*(4), 443–468.

Oakhill, J., Hartt, J., & Samols, D. (2005). Levels of comprehension monitoring and working memory in good and poor comprehenders. *Reading and writing*, *18*(7), 657-686.

Oakhill, J. V., & Cain, K. (2012). The precursors of reading ability in young readers: Evidence from a four-year longitudinal study. *Scientific studies of reading*, *16*(2), 91-121.

Ozernov-Palchik, O., Centanni, T. M., Beach, S. D., May, S., Hogan, T., & Gabrieli, J. D. (2021). Distinct neural substrates of individual differences in components of reading comprehension in adults with or without dyslexia. *NeuroImage*, *226*, 117570.

Palmović, M., Matić, A., & Kovačević, M. (2018). Resolution of anaphoric expressions in children and adults: evidence from eye movements. *Suvremena lingvistika*, *44*(85).

Patel, T. K., Snowling, M. J., & de Jong, P. F. (2004). A cross-linguistic comparison of children learning to read in English and Dutch. *Journal of Educational Psychology, 96*, 785–797.

Pavliša, J. I., & Lenček, M. (2011). Phonological skills and phonological memory as basic literacy predictors- some differences between children with typical language development, children with perinatal brain lesions and children with specific language impairmant. *Croatian Review of Rehabilitation Research*, *47*(1), 1-17.

Paris, S. G., & Hamilton, E. E. (2014). The development of children's reading comprehension. In *Handbook of research on reading comprehension* (pp. 56-77). Routledge.

Pearson, P. D., & Hamm, D. N. (2005). The assessment of reading comprehension: A review of practices—Past, present, and future. *Children's reading comprehension and assessment*, 31-88.

Perfetti, C. A. (1985). *Reading ability.* Oxford University Press.

Perfetti, C.A. (1999). Comprehending written language: A blueprint of the reader. The neurocognition of language, 167-208.

Perfetti, C. A. (2001). Reading Skills. *Psychiatry*, *50*, 1125-29.

Perfetti, C.A., Hart, L. (2002). The lexical quality hypothesis. Precursors of functional literacy, *11*, 67-86.

Perfetti, C.A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of reading*, *11*, 357-383.

Perfetti, C., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific studies of Reading*, *18*(1), 22-37.

Perfetti, C. A. (2017). Lexical quality revisited. *Developmental perspectives in written language and literacy: In honor of Ludo Verhoeven*, 51-67.

Rayner, K., Chace, K. H., Slattery, T. J., & Ashby, J. (2006). Eye movements as reflections of comprehension processes in reading. *Scientific Studies of Reading*, *10*, 241–255.

Rayner, K., & Reichle, E. D. (2010). Models of the reading process. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(6), 787-799.

Rayner, K., Reichle, E. D., & Pollatsek, A. (2010). Cognitive processes in reading: The EZ Reader model of eye-movement control. In *From Inkmarks to Ideas* (pp. 152-178). Psychology Press.

Reed, D. K., & Vaughn, S. (2012). Retell as an indicator of reading comprehension. *Scientific studies of reading*, *16*(3), 187-217.

Rieben, L., & Perfetti, C. A. (2013). *Learning to read: Basic research and its implications*. Routledge.Royer, J., Hastings, C., Hook, C. (1979). A Sentence Verification Technique for Measuring Reading Comprehension. *Journal of Reading Behavior, 11.* 10.1080/10862967909547341.

Ruffman, T. (2013). Reassessing children's comprehension-monitoring skills. In *Reading comprehension difficulties* (pp. 33-67). Routledge.

Rumelhart, D. E. (1977). Toward an integrative model of reading. In S. Dornic (Ed.), *Attention and performance* (pp. 573-603.). New York: Academy Press.

Saldaña, D., & Frith, U. (2007). Do readers with autism make bridging inferences from world knowledge? *Journal of Experimental Child Psychology, 96*(4), 310 319.

Scarborough, H. S., Neuman, S., & Dickinson, D. (2009). Connecting early language and literacy to later reading (dis) abilities: Evidence, theory, and practice. *Approaching difficulties in literacy development: Assessment, pedagogy and programmes*, *10*, 23-38.

Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology, 94*, 143–174.

Seymour, P. H. (2006). Theoretical framework for beginning reading in different orthographies. *Handbook of orthography and literacy*, 441-462.

Snow, P. C. (2021). SOLAR: The Science of Language and Reading. *Child Language Teaching and Therapy, 37*(3), 222–233.

Snowling, M. J., Stothard, S. E., Clarke, P., Bowyer-Crane, C., Harrington, A., Truelove, E. & Hulme, C. (2009). *YARC York Assessment of Reading for Comprehension.* London: GL Assessment

Stanovich, K.E. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly, 16*, 32-71.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequence of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*(4), 360-407.

Stein, N. L., & Glenn, C. G. (1975). *An Analysis of Story Comprehension in Elementary School Children: A Test of a Schema.*

Tobia, V., & Bonifacci, P. (2015). The simple view of reading in a transparent orthography: the stronger role of oral comprehension. *Reading and Writing*, *28*(7), 939–957.

Tong, X., Deacon, S. H., & Cain, K. (2014). Morphological and syntactic awareness in poor comprehenders: Another piece of the puzzle. *Journal of learning disabilities*, *47*(1), 22-33.

Tunmer, W. E., & Chapman, J. W. (2012). The simple view of reading redux: Vocabulary knowledge and the independent components hypothesis. *Journal of Learning Disabilities, 45*(5), 453–466.

Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension.* NY Academic Press.

Verhoeven, L., Voeten, M., & Vermeer, A. (2019). Beyond the simple view of early first and second language reading: The impact of lexical quality. *Journal of Neurolinguistics*, *50*, 28–36.

Wechsler, D. (1992). *Wechsler Individual Achievement Test.*

Westerveld, Marleen. (2009). Measuring reading comprehension ability in children: Factors influencing test performance. Acquiring Knowledge in Speech, Language and Hearing. 11, 81-84.

Wiederholt, J. L., & Bryant, B. R. (1992). *Gray oral reading tests: GORT-3*. Austin, TX: Pro-ed.

Woodcock, R. W. (1998). *Woodcock Reading Mastery Tests-Revised: Forms G and H.. Examiner's Manual*. American Guidance Service.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). Woodcock-Johnson III NU Complete. *Rolling Meadows, IL: Riverside Publishing*.

Woolley, G. (2008). The assessment of reading comprehension difficulties for reading intervention. *Australian Journal of Learning Difficulties*, *13*(1), 51-62.

Yuill, N., Oakhill, J., & Parkin, A. (1989). Working memory, comprehension ability and the resolution of text anomaly. *British journal of psychology*, *80*(3), 351-361.

Zaretsky, E., Kuvac Kraljevic, J., Core, C., & Lencek, M. (2009). Literacy predictors and early reading and spelling skills as a factor of orthography: Cross-linguistic evidence. Written Language and Literacy, 12, 52–81.

Ziegler, J. C., Bertrand, D., Tóth, D., Csépe, V., Reis, A., Faísca, L., Saine, N., Lyytinen, H., Vaessen, A., & Blomert, L. (2010). Orthographic Depth and Its Impact on Universal Predictors of Reading: A Cross-Language Investigation. *Psychological Science*, *21*(4), 551–559.