The conception of comprehension in two eye movement tracking models: Just and Carpenter (1980) and Rayner and Pollatsek (1989)

Abstract

Previous research on language comprehension has generally focused on the eye tracking method, as it allows us to account for reading processes and lexical access in real time. This method aims to observe the comprehension processes of the entire text contents. The contribution of these studies is undeniable, though visual processing of words is not a guarantee of its comprehension. For this reason, in this work I compare how comprehension is conceived in two models used in eye-tracking studies on language processing during reading: Just and Carpenter’s model (1980) and Rayner and Pollatsek’s model (1989), as they offer an interesting insight on text comprehension. This article mainly focuses on similarities and differences between both models and the extent to which the term processing could be used alternatively to refer to comprehension.

Keywords: comprehension; eye tracking; on-line processing; fixations; perceptual span

Resumen

Las investigaciones previas sobre la comprensión del texto escrito se han centrado generalmente en el método de seguimiento de movimientos oculares, puesto que permite dar cuenta de los procesos de lectura y el acceso léxico en tiempo real. Este método tiene como objetivo observar los procesos de comprensión de todo el contenido del texto. La contribución de estos estudios es innegable, a pesar de que el procesamiento visual de las palabras no es una garantía de su comprensión. Por esta razón, en este trabajo se compara cómo se concibe la comprensión en dos modelos de seguimiento de movimiento ocular del procesamiento del lenguaje durante la lectura: el modelo de Just y Carpenter (1980) y el modelo de Rayner y Pollatsek (1989). Ambos modelos ofrecen una interesante visión sobre el término comprensión de texto. Este artículo se centra en las similitudes y diferencias entre modelos y, principalmente, hasta qué punto el término procesamiento podría utilizarse de manera alternativa para referirse a la comprensión.

Palabras clave: comprensión; seguimiento de movimientos oculares; procesamiento en línea; fijaciones; span perceptual
1. Introduction

Comprehension is a complex phenomenon, the study of which is complicated due to the great variability and processes involved. An integrative model of comprehension should account for the various cognitive, social, affective, linguistic, and cultural aspects acquired from the environment, as well as inherited biological aspects (Peronard & Gómez, 1985). At the same time, it should go beyond the verbal features of the written text, as texts are multisemiotic (Parodi, 2011).

Also, the multidimensionality of comprehension should take into account the variability stemming from the strategies that readers adopt to optimally use their cognitive resources to construct a coherent mental representation of the text based on their world knowledge (Dijk & Kintsch, 1983).

At the same time, there are other variables that influence the comprehension process including the reader’s expertise, as explained by Rayner and Pollatsek:

beginner readers do not have as much world knowledge as more skilled readers do […] do not have efficient comprehension monitoring strategies […] younger children do not know as many words as older children and this lack of vocabulary knowledge hinders comprehension (1989: 391).

Thus, difficulties may arise during the reading process due to a lack of world knowledge, as can be seen in children’s inefficient word recognition and reduced vocabulary. This difference can affect the amount of information processed by the subject: 11 characters in beginning readers and 15 in experts (Veldre & Andrews, 2014).

In addition to reading expertise, metacognitive strategies play an important role in determining comprehension success. The reader of a text has to be aware and have control of the strategies that let him know that he is understanding, and thus regulate and monitor
the strategies according to his reading goal (Flavell, 1976, Baker, 2008). In this case, the reader constructs a representation of the meaning of the text based on these objectives (Graesser, Singer, & Trabasso, 1994).

The social and cultural context also influences comprehension (Carrell, 1981; Wyer, 2004). Prior knowledge of beliefs and assumptions forms a schema that helps understanding the text and, according to Steffensen, “once the appropriate schema is accessed, it drives comprehension of the text, and, among other things, inhibits the inappropriate reading of ambiguous items” (1986: 83).

In the words of Rumelhart (1981: 5), the schema is conceived as a “data structure for representing the generic concepts stored in memory”. The prototypicality of these schemes allows setting a common ground that is assumed to be known by the comprehender and other readers. This way, the contents of the schema does not have to be repeated, as it is already known information that does not have to be shared with others (Schank & Abelson, 1975).

As a result, the schemes provide a cognitive economy, since they help the reader to infer non-explicit contents in the text. Although reading is a cognitive process, learning to read is a social act in which the subject learns ideological and cultural schemes that predetermine its comprehension (Rumelhart, 1981). Given the emergence of studies such as linguistic relativism (Sapir, 1921; Whorf, 1956) and contrasting rhetoric (Kaplan, 1966) that raise questions about comprehension at the intercultural level, it is necessary to elucidate the extent to which such schemes are culturally shared.

In addition, knowledge of gender (Zwaan, 1994; Fountas & Pinnell, 2001; Best, Floyd, & Mcnamara, 2008), emotional and motivational factors (Ellis, Ottaway, Varner, Becker, & Moore, 1997; Law, 2009) and the incorporation of multimodal elements have proved to be vitally important for text comprehension (Sarıço- ban & Yürük, 2016; Wolfe, 2018).

Eye movement tracking studies have shown that processing during reading is subject to great variability among readers which,
in turn, depends on the complexity of the text they are confronted with: the greater the complexity of the text, the longer the fixation time and the more regressions the readers make. While the contribution of these studies is undeniable, how much a model can account for the sophistication of the comprehension phenomenon varies from model to model.

This paper aims to examine how comprehension is approached in two models used eye-tracking studies: the model proposed by Just and Carpenter (1980) and the one proposed by Rayner and Pollatsek (1989). Both models are reviewed in the first and second sections of the article, and, in the third section, they are compared and their differences discussed. The conclusions and future research are proposed in the final section.

2. The model of Just and Carpenter (1980)

This model studies the mechanisms and phases of processing words, sentences, and text segments during moment-to-moment reading. It aims to interpret comprehension in terms of the number of fixations that college-level students make while reading scientific texts. The authors consider that the reading pace matches the comprehension processes and that the reader’s attention is an indicator of its comprehension. As highlighted by Just and Carpenter:

A reader can take in information at a pace that matches the internal comprehension processes. By examining where a reader pauses, it is possible to learn about the comprehension processes themselves (1980: 329).

The burden of this processing is greater when the word is infrequent and involves making inferences about its meaning, especially if it is located at the end of a sentence. Thus, the number of fixations on the target word is an indicator of its degree of difficulty and the processes demanded to read it.
The model covers the reading stages based on two complementary principles: the principle of immediacy and the principle of eye-mind. The former holds that the word is interpreted once read in the text, while the latter states that the word is being processed at the same time it is fixated. Just and Carpenter (1980) place this interactive model of parallel processing midway between top-down and bottom-up models. The authors consider their model to be top-down since activation is supported by both the context and the frequency of the word itself. Also, it is bottom-up because it encompasses five phases of processing: word coding, lexical access, case assignment, inter-clauses integration, and sentence closure. Each of these phases is detailed below.

2.1. Word coding

Word representation is activated by its visual features. Several candidates are activated in the short-term memory and after this visual information is retrieved from the long-term memory, those not corresponding to the word are inhibited. If the same word is processed again, it is accessed more quickly, since it acquires greater connection weight because of its frequency (Morton, 1969). Hence, the more frequent the word is, the shorter the fixation time it requires. According to the model proposed by Just and Carpenter (1980), when a word has more than one meaning the most frequent one becomes activated thanks to its occurrence context, and the unselected meanings decay.

2.2. Lexical access

Studies focusing on disambiguation through oral context investigate how readers and listeners recover the most contextually relevant meaning of lexical items with multiple meanings. In this sense, there are three approaches: selective lexical access, exhaustive lexical access, and ordered lexical access.
The first approach states that the context biases the interpretation of an ambiguous word, so that only the desired meaning is accessed. In essence, the context provides sufficient information so that only the meaning most compatible with the context in which the ambiguous word is inserted is activated (Just & Carpenter, 1980; Tabossi, 1988; Simpson & Krueger, 1991; Simpson, 1994).

On the other hand, in non-selective or exhaustive lexical access, the word meaning is accessed independently of its context. Therefore, access is non-selective when the context does not influence the selection of the meaning to be activated, as all possible meanings of the word are activated (Conrad, 1974; Schvaneve, Meyer, & Becker, 1976; Swinney, 1979; Seidenberg, Tanenhaus, Leiman, & Bienkowski, 1982; Rayner & Pollatsek, 1989).

Finally, the third approach assumes that lexical access is ordered by frequency, so that the most dominant meaning is activated without any interference from the context (Hogaboam & Perfetti, 1975; Neill, Hilliard, & Cooper, 1988). This approach suggests that the word frequency determines its selection as the first meaning to be processed when the word has multiple meanings.

### 2.3. Assigning semantic roles

Successful case assignment is key in word processing. This assignment is not easy in a language such as English where the word *hammer* can be interpreted as a tool or can be used as a verb. In the same way, the animacy of the noun *woman* in example (1a) allows it to be the subject of the sentence even though it can also be a direct object. Therefore, some latency might occur when reading the first sentence, as it is hard to determine who saw whom exactly.

(1) a. *The woman* who saw the dentist.
   b. *The test* the dentist saw.

By contrast, the inanimate noun *test* observed in example (1b) produces less latency, as it is less likely for it to be misread as the
subject of the sentence. The immediacy principle supposes that adult readers would assign the case to the animate subject even if there is another inanimate subject in the sentence, and would make regressions if this assignment is found to be erroneous during sentence integration.

2.4. Interclausule integration

New information presented in the text must be integrated with information stored in the long-term memory. Integrating information requires a strategic reader who knows how to connect the information presented in the text with its previous world knowledge in order to make sense of what he or she is reading (Van Dijk & Kintsch, 1983). In this context, a self-conscious reader can construct a coherent mental representation that links the new information with that previously known by making inferences (Peronard & Gómez, 1985). In this sense, some readers are better skilled than others in establishing memory cues and recovering structures based on their expertise (Ericsson & Kintsch, 1995).

At the same time, a text must have certain features that facilitate its integration, such as causal, temporal, and referential relationships between its parts. Conversely, a text that lacks coherence or has a greater syntactic complexity requires a longer time to read (Just & Carpenter, 1980). The reader can detect some inconsistencies in the text when the various possible interpretations of a word create a garden path effect (Luo, 2017). That is, ambiguity occurs when a dominant but inappropriate meaning of the word is activated. Once the reader realizes that he or she has misinterpreted this word, a backward fixation is needed to process the content again (Carpenter & Daneman, 1981).

2.5. Sentence wrap-up

The sentence wrap-up involves searching for references that coherently link two parts of the sentence or the current sentence with
the previous or subsequent one. According to Just and Carpenter (1980: 346), integration takes place at the end of the sentence since it clarifies within-sentence ambiguity: “ends of sentences are unambiguous, they have the same role across sentences, and they may be processed more uniformly than the cues to within-sentence clause boundaries”.

However, in a sentence such as example (2) it is not possible to determine whether it was the president who had the accident or his daughter; the referent has to be found out in a previous or subsequent sentence.

(2) The journalist interviewed the daughter of the president who had the accident.

Finally, Just and Carpenter (1980: 350) recognize reading variations between individuals due to causes such as the “function of who is reading, what they are reading, and why they are reading it”. This means that such variation can take place depending on the reading goals and the reader’s working memory capacity to integrate new information with existing information.


According to this model, comprehension is studied at the level of the individual’s reading processing. Rayner and Pollatsek (1989: 2) focus on word and sentence processing, since “the process of comprehending text is much more complex than that”. The authors state that their work focuses on the mental cognitive processes involved in reading and argue that a longer fixation time requires a high cognitive load which, in turn, denotes a complexity in processing.

The authors distinguish between componential processes and higher-order processes. Componential processes are mostly automatic and better developed in expert readers thanks to their experience in recognizing familiar words. For Rayner and Pollatsek (1989: 69) higher-order processes are “those processes that put
the meaning of the words together...”. They are integration processes aimed to revise the text and construct a coherent interpretation based on the reader’s world knowledge, with the purpose of “constructing the correct syntactic structure, relating word meanings, and fitting the text into what the reader understands about the world” (Rayner & Pollatsek, 1989: 62). According to Reichle, Rayner, and Pollatsek (2003: 450), higher-order processes intervene in the reading “only when something is wrong and either send a signal to stop moving forward or a signal to execute a regression”.

In this regard, the authors identify three higher-order reading processes that require a longer fixation time on the target word: 1) The sentence wrap-up means that more fixations are needed when the target word is the last one in the sentence. The long processing time proves that the information is integrated at the end of the sentence or clause (Rayner, Kambe, & Duffy, 2000). 2) The search for antecedents makes use of the fact that the text components are interlinked in the sense that the pronoun alludes to some antecedent in the text that coincides with it in gender and number. The referent has to replace some element explicitly mentioned in the text and linked to the topic addressed in it, as Rayner and Pollatsek indicate (1989: 42): “if a reader encounters the bird, and the current topic sentence is not about a bird, a search for the previous mention of an appropriate bird may be initiated”. 3) Elaborative inferences are inferences that the reader makes on-line to fill gaps of non-explicit information in the text. This type of inferences is not necessary for comprehending the text, but give global coherence to the processed text (León, 2001).

The model architecture relies mainly on saccadic movements and fixations that operate through various processing activities directed by the working and the long-term memory as seen in Figure 1. This long-term memory involves three components or modules: the lexicon, world knowledge, and representations of previously read texts.
The conception of comprehension in two eye movement tracking models

The working memory contains a parser that constructs the syntactic representation, a thematic processor that provides a semantic representation, and an internal processor (called inner speech) that keeps the information active until the on-going process concludes, facilitating a literal memory of what has been read.

Rules are the third component of the model and allow automatic lexical access to familiar words. The fourth module is the most important one. It is specialized in word coding and is developed by the authors in other works. Once the processing cycle is completed, attention is shifted to processing and coding the next word. Finally, the authors recognize the lack of detail in their model, and explain the following:

we are not particularly tied to this model as the truth about reading. Instead, we see the model as a (temporary) working model […] As we continue to learn more about reading process, we will refine the model (as will proponents of other models) until we (and they) hone in on one that accurately captures all the facts about reading. When that happens, we will have understood reading. However, given the complexity of the reading process, it would probably take an entire book to describe such a model (Rayner & Pollatsek, 1989: 471).

As noted above, this model focuses on the visual coding of the word, following the moving window paradigm (McConkie & Rayner, 1975) in which, according to Henderson, Singer, and Ferreira (2013: 73–74), “the amount of text presented to the reader during any given fixation is directly manipulated by changing the display as a function of eye position”. Visual coding occurs in these windows by processing the foveal information in the word letters and the parafoveal information that extends to the left or right of the fixation point.

Perceptual span plays an extremely important role in the visual coding of the word. Rayner and Pollatsek (1989: 24) define
The conception of comprehension in two eye movement tracking models

perceptual span as “the region around a reader’s fixation point from which useful information can be obtained”. Studying perceptual span is difficult because its size varies depending on the reader’s age (Risse & Kliegl, 2011) and expertise (Rayner, 1986), whether he suffers from any disease or injury (Neale, 1971; Crossland & Rubin, 2006), and the font type and size in the text (Paterson & Tinker, 1947; Yan, Zhou, Shu & Kliegl, 2015). It may even be questioned whether the reading speed and the size of the perceptual span are acquired or learned and improved by practice (Sutherland, 1946).

The relevance of size in perceptual span is major, since the ability to capture several words in a fixation not only accelerates reading but also improves the comprehension of the material (Patberg & Yonas, 1978). Perceptual span is also an indicator of performance during the reading process that allows predicting the reader’s speed and level of experience (Sperlich, Meixner & Laubrock, 2016). In addition, it provides information on the reader’s strategies during processing and aspects of the language being processed. Strategic readers are experts who manage to move forward with reading and this explains their larger perceptual span relative to inexpert readers (Bai, Gao, Gao, & Wang, 2017). At the same time, a small span in an expert reader can be due to a right-to-left reading direction as in Hebrew (Pollatsek, Bolozky, Well, & Rayner, 1981) or top-to-bottom as in Chinese (Tsai & McConkie, 1995).

However, although the authors recognize the importance of emotional, motivational, pathological, or cultural aspects, they leave them aside to construct a model that characterizes the reading process itself. In this context, reading is a default process that is investigated in the model, regardless of the difficulty of the text components or higher-order processes. According to Reichle, Rayner, and Pollatsek, the reason behind this is:

Because the model was not intended to be a deep explanation of language processing, it does not account for the many ef-
fects of higher-level linguistic processing on eye movements [...] The model can therefore be viewed as the “default” reading process. That is, we view the process of identifying words to be the forward “driving engine” in reading, as the process of knitting the words into larger units of syntax or meaning would be too slow (whether successful or not) to be a signal to decide how and when to move the eyes forward for skilled readers (2003: 450).

Actually, the model characterizes higher-order processes but does not take them into account. For this purpose, Rayner published a series of papers in which he summarizes higher-order processes and separates lexical disambiguation (Rayner & Duffy, 1986; Duffy, Morris & Rayner, 1988; Rayner & Frazier, 1989; Dopkins, Morris, & Rayner, 1992; Reichle, Pollatsek, & Rayner, 2007). Hence, the purpose of this model is to study the default reading process by an ideal expert reader and compare it with the performance of a beginner reader (Rayner & Hagelberg, 1975), readers suffering from pathologies (Rayner, 1983), and readers of a wide age range (Juhasz & Rayner, 2006).

Thus, Rayner and Pollatsek (1989) provide a bottom-up modular model of expert readers that assumes that access to the word is carried out regardless of the context in which it occurs. Hence, all possible meanings of the word are accessed and then the non-corresponding ones are inhibited. Consequently, it is a non-incremental model that proposes immediate access of processed words (Garrod & Sanford, 1999). It should be noted that most eye-tracking reading models make the same methodological decision as Rayner and Pollatsek (1989): To isolate the study of reading processes from any other individual factor (Morrison, 1984; Engbert, Longtin, & Kliegl, 2002; McDonald, Carpenter, Shepart, & Shillcock, 2005; Feng, 2006; Yang, 2006). Although these studies have contributed to explain on-line reading processes, they have not been able to consider the external or internal variations between readers.
4. Discussion

The two models presented above intend to account for the reading process in real time as detailed in Table 1. Both models study comprehension in terms of the processing that takes place at the word and sentence levels. In addition to word- and sentence-level processing, Just and Carpenter (1980) also include textual segments or text units in their model. In this model, the authors confuse comprehension with processing by stating that the number of fixations is an indicator of comprehension. As per their model the number of fixations may be a processing indicator, though it does not guarantee a deep comprehension. For this reason, Rayner and Pollatsek (1989) state that their model is limited to studying processing during reading as an aspect of comprehension due to the complexity of factors that understanding a text involves.

| Table 1. Differences and similarities between the two models |
|-----------------|-----------------|-----------------|
| Model type       | Interactive     | Modular         |
| Unit of analysis | Word, sentence, text fragments | Words and sentences |
| Measurement units | Number of fixations | Size of perceptual span |
| Variability among readers | Consider variability and text complexity | Discard any variability or text complexity |
| Processing phases | Five phases | Eight modules |
| Lexical access   | Selective       | Non selective   |
| Reader           | Common, not necessarily expert | Ideal, expert |

Therefore, Rayner and Pollatsek (1989) present a modular, bottom-up model of expert readers in which all word meanings are automatically accessed, implying a longer processing time. The model leaves out all the factors that influence the reading process.
in order to characterize the process itself in a greater depth. It focuses on the visual coding of the word and the number of letters processed in the perceptual span, and it does not separate processing phases but recognizes non-recursive serial processing modules. The model encompasses componential and higher-order processes and argues that the number of fixations is influenced by the word frequency and focuses on the perceptual span as a key indicator of the processing.

In contrast, the model proposed by Just and Carpenter (1980) considers the variation and reading objectives of the subjects. In this model lexical access is selective, as the context prompts activation of the appropriate word. Although the authors do not distinguish expert from non-expert readers, they take into account the differences between individuals in working memory capacity. Consequently, they distinguish five phases of reading processing.

5. Conclusions

The model proposed by Just and Carpenter (1980) takes a step towards realizing the complexity of the comprehension phenomenon. They present a strategic reader who knows how to make optimal use of its working memory to process text during reading, which varies from one individual to another depending on their reading goals and the material they read. By contrast, Rayner and Pollatsek (1989) offer a simpler version that aims to explain in depth a specific aspect of this same phenomenon reduced to processing. For this reason, they opt for an ideal reader not influenced by the difficulty of the text or any other external aspect.

Further work is required to shed light on the brain regions involved in processing. It is also necessary to replicate the model proposed by Just and Carpenter (1980) and studying to what extent the immediacy and mind-eye principles are complementary to each other. Based on our review, we conclude that comprehension is a complex phenomenon and that further models are needed to elucidate its various aspects and dimensions.
6. References


