ARTIS: a digital interface to promote the rehabiliatation of text comprehension difficulties through Artificial Intelligence

Martina Galletti[1,2,*,†], Eleonora Pasqua[3,2,†], Manuela Calanca[3,2], Caterina Marchesi[3], Donatella Tomaiuoli[3,2,5] and Daniele Nardi[2,4]

[1]Sony Computer Science Laboratories-Paris (Sony CSL-Paris) - France
[2]Sapienza University of Rome - Italy
[3]Centro Ricerca e Cura di Roma - Italy
[5]Tor Vergata University of Rome, Italy
[4]CINI-AIIS - Italy

Abstract

This study introduces ARTIS, an AI-powered interface aimed at enhancing text comprehension in children with language and learning disorders and deficits. Leveraging neuro-psycholinguistic models, ARTIS offers personalized practice at varying difficulty levels, targeting specific linguistic components of text processing. Through AI algorithms, ARTIS autonomously extracts keywords, associates them with pictograms, identifies complex words, generates semantic networks, and proposes exercises on grammatical components. By automating cognitive processes and providing tailored interventions, ARTIS represents a significant advancement in promoting inclusive hybrid speech and language therapy practices, and improving text comprehension skills in children with special educational needs.

Keywords

Natural Language Processing, Text comprehension, Language and learning disorders, Speech and Language Therapy

1. Introduction

Understanding a text allows us to acquire information, access new content, enrich the knowledge we already possess. Conversely, having difficulty in understanding what one reads is a great limitation to personal growth, and restrict opportunities, not only at school, but also in life. Text comprehension is thus a fascinating but also very complex ability. It requires motivation, attention, memory, but also specific language skills. These can be problematic for children with language and learning disorders. In fact, subjects with poor text comprehension show difficulties related to the processing of syntactic and semantic sentence components [1], the analysis of lexical components of words [2] and deficits in the syntactic representation of words and oral comprehension skills [3]. Moreover, [4] stated that the same subjects report significant deficits in receptive vocabulary and semantic processing. Finally, [5] and [6] addressed the issue of grammar, claiming that children and adolescents with problems in text comprehension show difficulties in understanding the role of pronouns within sentences, especially if these are in clitic form.

Artificial intelligence-powered clinical and education tools have the potential to revolutionise learning and speech and language therapy for children with special educational needs and disabilities. It can personalize interventions and adaptive content that meet individual needs, thus promoting a more inclusive and equitable educational experience. Within this context we developed ARTIS — an interface designed to facilitate the rehabilitation and instruction of text comprehension skills through artificial intelligence. Starting from neuropsycholinguistic models of reading comprehension and focusing on the linguistic components of text processing, ARTIS enables personalized practice on texts at different levels. Thanks to AI algorithms, the interface is able to automatically extract pictograms from keywords, identify more complex words, generate semantic networks, and to propose exercises on certain grammatical components. ARTIS is aimed at primary and secondary school children with difficulties in understanding text, but can also be used as support for English as an L2. The output of our contribution is twofold. First, we designed and deployed the interface. Second, we tested whether AI can be a valid support tool for text comprehension from

Ital-IA 2024: 4th National Conference on Artificial Intelligence, organized by CINI, May 29-30, 2024, Naples, Italy

*Corresponding author.
†These authors contributed equally.
martina.galletti@sony.com (M. Galletti); e.pasqua@crc-balbuzie.it (E. Pasqua); m.calanca@crc-balbuzie.it (M. Calanca); c.marchesi@crc-balbuzie.it (C. Marchesi); d.tomaiuoli@crc-balbuzie.it (D. Tomaiuoli); nard@diag.uniroma1.it (D. Nardi)
0009-0002-2079-8999 (M. Galletti); 0000-0002-7153-6094 (E. Pasqua); 0000-0001-6606-200X (D. Nardi)

a clinical perspective.

2. Pedagogical and Psycho-linguistics Foundations

Over the years researchers in psycho-linguistics had been trying to formalise how people understand connected text. Text comprehension is defined as a complex cognitive task which involves an active process of meaning construction, dependent not only on the information in the text but also on the information possessed by the reader [7]. Researchers in psycho-linguistics have strived to formalize the mechanisms underlying this process. As a result, four main theories emerged as the main ones in the psycho-linguistics landscape:

1. **The Simple View of Reading** by Gough and Tunmer [8] where the decoding process and language skills are considered central to text comprehension. The key idea of this model is that comprehension, **C**, is defined as the product of decoding factors, **D**, defined as fluent word reading and language comprehension, **L** [8].
2. **Structure Building Framework** by Gernsbacher [9, 10, 11] where text comprehension is defined with the metaphor of the construction of a building where one's must start from the foundations i.e. from the first element contained in the text, for later one integrating the new information and enriching the departing structure in a cohesive one.
3. **Memory-Focus Model** of Sanford and Garrod [12, 13] in which is central the process of co-reference resolution. In fact, the key idea of this theory is that when comprehending a text, a reader must first establish a coherent interpretation of the text comprehended so far and must therefore establish whether a given element had been discussed previously or not.
4. **Construction-Integration Model** of Kintsch and van Dijk [14, 15, 16] focuses on how information from the text, at the granularity level of propositions, is connected to and completed by information stored in the long-term memory of the reader. Besides literal understanding of the words and syntax in the text, it's essential to be able to gather propositions from the text, their concepts and connections, and organise them into interconnected hierarchical structures. It is only with this representation of information, integrated with the reader's preexisting knowledge makes it possible to gain deep access to the content of the text[7, 16, 15]. To reflect this view, Kintsch and van Dijk identified two main levels of processing: local and global. Local processing involves the understanding of individual sentences and phrases, while global processing focuses on integrating these local meanings into a coherent representation of the entire text. Central to their model is the concept of "coherence," where readers strive to create a cohesive mental representation by linking information within the text and connecting it to their existing knowledge.

As much as the models described may emphasize different components as central to the comprehension process, they all converge on the idea that language skills, along with semantic and inferential skills, participate in linking prior knowledge to new information, creating a coherent and more complex representation of new meanings. In designing our interface, ARTIS, we drew inspiration from these foundational theories in psycho-linguistics.

3. Existing Work

In the literature, there is one Italian systems for tele-rehabilitation targeting specific learning disorders and special educational needs, i.e. RIDInet [17]. RIDInet focuses on text comprehension skills, particularly semantic and syntactic inferential processes. Neberthless, despite the clinical context, RIDInet lacks the integration of Natural Language Processing (NLP) techniques. In English, two primary systems are 3D Readers [18] and CACSR [19]. 3D Readers offer users options for verbal or visual strategies to enhance comprehension, with immediate feedback provided. CACSR provides personalized instruction using various techniques like visual images and summarizing strategies, offering real-time feedback. Although these systems enhance reading comprehension, they are primarily used for educational contexts and lack focus on clinical categories. Moreover, Open Book [20] is the only system employing NLP, mainly for children with Autism Spectrum Disorder, focusing on text simplification and customization functions [21]. Its key features include text simplification through NLP techniques and rich customization functions, enabling users to quickly adapt the document presentation (font, font size, line spacing, colors) to their preferences. It also provides assistive tools such as dictionaries and images. However, Open Books was designed for a specific clinical population, and it focused more on improving their decoding skills than reading comprehension per se. Finally, Systems targeting tele-rehabilitation specifically for Dyslexia exist [22, 23, 24, 25], but few focus on reading comprehension skills.

4. Target Population of ARTIS

ARTIS targets two distinct target populations: students with specific reading comprehension disorders, and those identified in literature as "poor comprehenders".

Reading Comprehension Disorder, identified in the DSM-V[26], delineates difficulty in grasping the meaning of text despite proficient decoding skills. This encompasses understanding word sequences, implicit information (inferences), and the deeper meaning of text content. The distinction between decoding and comprehension abilities is highlighted in the literature [27], which underlines differences in cognitive processes, predictive factors, disorder characteristics, and treatment approaches. The functions involved in text comprehension include lexical skills (vocabulary), inferential skills, working memory, attention, and meta-cognitive control, which are distinct from decoding skills. Even if this disorder therefore often intersects with other conditions such as specific language disorder, decoding disorder (dyslexia), intellectual disability, or memory difficulties, it can be said that specific reading comprehension disorder is somehow independent of specific decoding disorder. Despite complexities in classification, understanding this disorder is crucial, as it often gets conflated with other clinical profiles. This is precisely the reason behind the beginning of this work and the reason why we involved since the beginning of its conception speech and language therapists.

The literature defines "poor readers" or "poor comprehenders" students who experience specific problems in comprehension in the face of decoding skills that are instead within the normal range [28],[29], [30]. Distinguishing poor comprehension from a reading or writing learning disability necessitates diagnostic test results. For example, dyslexia manifests as reading effort, characterized by decoding and reading difficulties, error prevalence, and fluency deficits. In contrast, poor readers and writers exhibit milder characteristics, including adequate reading speed with occasional fluency issues. Poor comprehenders' difficulties often stem from factors such as inadequate reading and writing process automation, environmental stimulation deficiencies, cognitive, memory, or attention issues, rather than neurological causes.

5. Architecture of ARTIS

ARTIS is an online tool designed to enhance reading comprehension skills. It comprises four distinct levels aimed at fostering a foundational understanding of language. The tool assists users in comprehending the lexical and syntactic aspects of the text, beginning with understanding sentences and progressing to words. Furthermore, the interface aids users in constructing coherent sequences and hierarchical structures within the text through exercises that focus on co-reference structures. Finally, in a preliminary stage, ARTIS endeavors to develop users' ability to construct broader mental models of language by integrating textual information with their existing knowledge [31]. Subsequent sections will delve into the functionalities of each level, providing detailed insights into the underlying algorithms. The initial prototype of the interface was extensively described in [32].

5.1. Understanding Sentences

The interface initially presents the text at a sentence level using the Spacy Sentencizer tool [1]., allowing users to focus on individual sentences. To aid comprehension, users can listen to the text through speech synthesis powered by the Google text-to-speech API, improving accessibility for those with reading difficulties. Keywords from each sentence are extracted and displayed using Picture Communication Symbols (PCS) on the left side of the screen, facilitating comprehension for individuals with literacy challenges. The extraction process involves using a customized version of Keybert [33] and manual verification by speech and language therapists. These keywords are linked to appropriate pictograms through the Arasaac API [2], replacing real images to ensure suitability for vulnerable audiences.

Figure 1: Section on understanding sentences.

5.2. Understanding Words

In a following step, the interface offers a closer examination of the text at the word level. It displays the ten rarest words for each selected text, identified using WordFreq [34], a tool providing frequency estimates across languages. This feature aims to assist individuals with reading comprehension difficulties by addressing their hesitancy towards unfamiliar vocabulary. Upon user selection, the word's definition, along with a PCS representing its meaning and the original sentence context, is provided. Definitions, sourced from the Oxford Dictionary API [3], are presented in both written and spoken forms through speech synthesis, enhancing accessibility.

5.3. Understanding Paragraphs

Based on insights from [6], a co-reference resolution algorithm was incorporated to aid users in grasping the connections between different entities and pronouns within

[1]https://spacy.io/api/sentencizer
[2]https://arasaac.org/
[3]https://developer.oxforddictionaries.com/

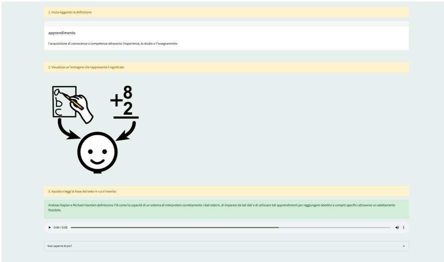

Figure 2: Section on understanding words.

the text. Pronouns were prioritized due to the challenges associated with their recognition. In this level of the interface, sentences are segmented, and grammatical elements such as common nouns, proper nouns, and pronouns are highlighted with distinct colors based on their Part of Speech Tags. This color-coded approach assists in immediate differentiation between elements and suggests the function of each lexical morpheme, guiding the understanding of pronouns by identifying their correct referents. Users are presented with propositions and asked to select the correct referent from four alternatives. One is accurately identified by the co-reference resolution algorithm, while two are intentionally misleading, and the fourth closely resembles the correct answer. Positive feedback is provided upon selecting the correct answer. Spacy part-of-speech tagger [4] was used for parsing, and a fine-tuned version of Coreferee [5] served as the coreference resolution algorithm.

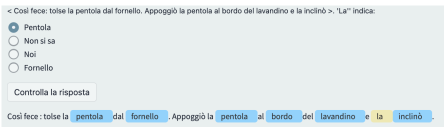

Figure 3: Section on understanding paragraphs.

5.4. A first attempt for bridging Textual Understanding with Prior Knowledge

As emphasized by [31], text comprehension extends beyond its surface-level representation. To foster deeper understanding, Synset Networks were introduced. Each selected word is presented alongside its associated Synsets from the Merriam Webster thesaurus, facilitating vocabulary expansion and recall. This approach enhances comprehension by elucidating the various meanings a term can encompass within a text and linking it with familiar words, thereby aiding in the integration of new information with existing knowledge. Through the Synset Network, users explore the interconnections of words, fostering a deeper understanding of their multifaceted meanings.

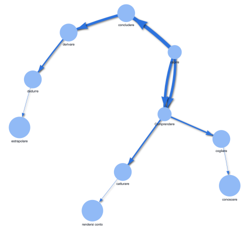

Figure 4: Sectiong on bridging existing and newer knowledge.

6. Empirical Evaluation

A pilot study was conducted at the CRC in Rome, a developmental rehabilitation center, from January to June 2023. The study involved 24 Italian-speaking children from primary 2 to 5, who were assessed using the platform under the guidance of speech therapists from the CRC. Each child had a clinical diagnosis and an updated functional profile, including IQ, language, and learning profiles. The study included children with impairments in text comprehension and language skills, particularly in receptive and expressive vocabulary not aligned with their age reference. The participants comprised children diagnosed with Specific Learning Disorder (7 subjects), Primary Language Disorder (8 subjects), Borderline Intellectual Functioning (7 subjects), and High Functioning Autism (3 subjects).

The children who participated in the feasibility study underwent standardised pre- and post-treatment evaluations, with particular attention in the post-evaluation to data on reading comprehension and language skills. The speech therapists who followed the children in the trial were also asked to fill in a questionnaire (google form) to investigate usability, functionality and the level of perceived effectiveness of ARTIS in its different components, in relation to the different pathologies, with qualitative questions structured according to a five-point likert scale (Very Little, Slightly, Sufficient, Very Much). The data collected focused on the platform usage, work area preferences, exercise types, and their correlation with different clinical profiles. The therapists' feedback regarding exercise functionality, including the need for additional support or information beyond AI-generated content, was also evaluated.

Results indicated that all participants (24 children) utilized the platform to visualize sentences and aid comprehension of selected texts. This feature included displaying relevant images corresponding to individual sentences and the option for speech synthesis reading. No

[4] https://spacy.io/usage/linguistic-features
[5] https://github.com/richardpaulhudson/coreferee

supplementary material or therapist intervention was required for this aspect. Children with more significant language deficits (DPL, FIL, AUTISM) tended to use the vocabulary area (focused on words) more frequently. The inclusion of definitions accompanied by pictures proved highly beneficial. Occasionally, additional information sourced from Google searches was needed to clarify or enhance the meaning of certain words. A small subset of the sample (4 subjects diagnosed with DSA) engaged with the semantic network. Utilizing the semantic network required meta-cognitive and generalization skills, with therapist support in tasks such as image searches, reflection on different word uses, and identification of texts or phrases highlighting semantic distinctions.

The results of the pre and post-treatment evaluations show a significant increase in text comprehension and some language skills with particular reference to receptive and expressive vocabulary.

7. Risks and Mitigations

To ensure the development of a trustworthy AI system, extensive involvement of key stakeholders, particularly speech and language pathologists, was maintained throughout the work. Their participation ranged from the initial decision-making phases to the final evaluation of the interface, with continuous exchange of advice and feedback. Moreover, the administration of exercises always occurs under the supervision of practitioners. Before conducting clinical trials involving children, precautionary measures were taken to mitigate risks. Texts were pre-selected and categorized by school age to cater to users' individual levels. Additionally, manual clinical control was implemented to identify and rectify any potentially unsafe pictograms or misleading keywords extracted by the AI models. A stop button was integrated into the interface to prevent exposure to misleading material. Furthermore, certain functionalities, such as the semantic network, are selectively displayed based on individual needs and clinical judgment, ensuring tailored support for each user.

8. Conclusions & Future Work

In this paper, we described the development of ARTIS, an AI-powered interface designed to enhance text comprehension in children with language and learning disorders. By leveraging neuro-psycholinguistic models and thanks to the integration of AI algorithms, ARTIS autonomously extracts keywords, associates them with pictograms, identifies complex words, generates semantic networks, and proposes exercises on grammatical components. These features aim to improving text comprehension skills in children with special educational needs by promoting inclusive hybrid speech and language therapy practices.

Looking ahead, future work could focus on several aspects. Firstly, conducting an extensive clinical evaluation is essential to assess the effectiveness of the interface in clinical settings. Additionally, we aim to integrate concept-map-based document summaries, as highlighted in [35], to enhance the bridging between knowledge present in the text read and prior knowledge.

References

[1] J. Oakhill, Instantiation in skilled and less skilled comprehenders, The Quarterly Journal of Experimental Psychology Section A 35 (1983) 441–450.

[2] K. Nation, M. J. Snowling, Factors influencing syntactic awareness skills in normal readers and poor comprehenders, Applied psycholinguistics 21 (2000) 229–241.

[3] R. Padovani, La comprensione del testo scritto in età scolare. una rassegna sullo sviluppo normale e atipico, Psicologia clinica dello sviluppo 10 (2006) 369–398.

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

[5] F. Arosio, C. Branchini, L. Barbieri, M. T. Guasti, Failure to produce direct object clitic pronouns as a clinical marker of sli in school-aged italian speaking children, Clinical linguistics & phonetics 28 (2014) 639–663.

[6] A. Zachou, E. Partesana, E. Tenca, M. T. Guasti, et al., Production and comprehension of direct object clitics and definite articles by italian children with developmental dyslexia, Advances in language acquisition (2013) 464–471.

[7] W. Kintsch, On comprehending stories, Cognitive processes in comprehension (1977) 33–62.

[8] P. B. Gough, W. E. Tunmer, Decoding, reading, and reading disability, Remedial and special education 7 (1986) 6–10.

[9] M. A. Gernsbacher, Language comprehension as structure building, Psychology Press, 2013.

[10] M. A. Gernsbacher, Cognitive processes and mechanisms in language comprehension: The structure building framework, in: Psychology of learning and motivation, volume 27, Elsevier, 1991, pp. 217–263.

[11] M. A. Gernsbacher, Two decades of structure building, Discourse processes 23 (1997) 265–304.

[12] A. J. Sanford, S. C. Garrod, S. Garrod, Understanding written language: Explorations of comprehension beyond the sentence, John Wiley & Sons, 1981.

[13] A. J. Sanford, S. C. Garrod, Selective processing in text understanding. (1994).

[14] T. A. Van Dijk, W. Kintsch, et al., Strategies of discourse comprehension (1983).

[15] W. Kintsch, The role of knowledge in discourse comprehension: A construction-integration model, in: G. Stelmach, P. Vroon (Eds.), Text and Text Processing, volume 79 of *Advances in Psychology*, North-Holland, 1991, pp. 107–153. URL: https://www.sciencedirect.com/science/article/pii/S0166411508615514. doi:https://doi.org/10.1016/S0166-4115(08)61551-4.

[16] W. Kintsch, How readers construct situation models for stories. The role of syntactic cues and causal inferences, Typological Studies in Language 31, John Benjamins., Amsterdam; Philadelphia, 1995. URL: https://www.degruyter.com/database/COGBIB/entry/cogbib.6551/html, 2010.

[17] RIDInet, https://www.anastasis.it/ridinet/), ????

[18] 3D Readers Software, https://www.sbir.gov/node/334161), ????

[19] A.-H. Kim, S. Vaughn, J. K. Klingner, A. L. Woodruff, C. Klein Reutebuch, K. Kouzekanani, Improving the reading comprehension of middle school students with disabilities through computer-assisted collaborative strategic reading, Remedial and special education 27 (2006) 235–249.

[20] E. Barbu, M. T. Martín-Valdivia, L. A. U. Lopez, Open book: a tool for helping asd users' semantic comprehension, in: Proceedings of the Workshop on Natural Language Processing for Improving Textual Accessibility, 2013, pp. 11–19.

[21] N. Pavlov, User interface for people with autism spectrum disorders, Journal of Software Engineering and Applications 2014 (2014).

[22] M. Wang, B. Muthu, C. Sivaparthipan, Smart assistance to dyslexia students using artificial intelligence based augmentative alternative communication, International Journal of Speech Technology (2021) 1–11.

[23] A. Zingoni, J. Taborri, V. Panetti, S. Bonechi, P. Aparicio-Martínez, S. Pinzi, G. Calabrò, Investigating issues and needs of dyslexic students at university: Proof of concept of an artificial intelligence and virtual reality-based supporting platform and preliminary results, Applied Sciences 11 (2021) 4624.

[24] P. D. Barua, J. Vicnesh, R. Gururajan, S. L. Oh, E. Palmer, M. M. Azizan, N. A. Kadri, U. R. Acharya, Artificial intelligence enabled personalised assistive tools to enhance education of children with neurodevelopmental disorders—a review, International Journal of Environmental Research and Public Health 19 (2022) 1192.

[25] S. Rajapakse, D. Polwattage, U. Guruge, I. Jayathilaka, T. Edirisinghe, S. Thelijjagoda, Alexza: A mobile application for dyslexics utilizing artificial intelligence and machine learning concepts, in: 2018 3rd International Conference on Information Technology Research (ICITR), IEEE, 2018, pp. 1–6.

[26] A. P. Association, et al., Diagnostic and statistical manual of mental disorders, Text revision (2000).

[27] A. P. Association, et al., Manuale diagnostico e statistico dei disturbi mentali, Quinta edizione. Raffaello Cortina Editore (2014).

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

[29] J. Oakhill, K. Cain, Issues of causality in children's reading comprehension, Reading comprehension strategies: Theories, interventions, and technologies (2007) 47–71.

[30] B. Carretti, R. De Beni, C. Cornoldi, et al., Disturbi della comprensione del testo, in: Difficoltà e disturbi dell'apprendimento, Il Mulino, 2007, pp. 143–162.

[31] W. Kintsch, T. A. Van Dijk, Toward a model of text comprehension and production., Psychological review 85 (1978) 363.

[32] M. Galletti, E. Pasqua, F. Bianchi, M. Calanca, F. Padovani, D. Nardi, D. Tomaiuoli, A reading comprehension interface for students with learning disorders, in: Companion Publication of the 25th International Conference on Multimodal Interaction, 2023, pp. 282–287.

[33] M. Grootendorst, Keybert: Minimal keyword extraction with bert., 2020. URL: https://doi.org/10.5281/zenodo.4461265. doi:10.5281/zenodo.4461265.

[34] R. Speer, rspeer/wordfreq: v3.0, 2022. URL: https://doi.org/10.5281/zenodo.7199437. doi:10.5281/zenodo.7199437.

[35] M. Galletti, M. Anslow, F. Bianchi, M. Calanca, D. Tomaiuoli, R. Van Trijp, E. Pasqua, Interactive concept-map based summaries for send children, 2022, pp. 5236–5243.