A Reading Comprehension Interface for Students with Learning Disorders

Martina Galletti
martina.gallett@sony.com
Sony Computer Science Laboratories-Paris, Sapienza Università di Roma
Italy

Eleonora Pasqua
Centro Ricerca e Cura di Roma, Sapienza Università di Roma
Italy

Francesca Bianchi
Centro Ricerca e Cura di Roma
Italy

Manuela Calanca
Centro Ricerca e Cura di Roma,Sapienza Università di Roma
Italy

Francesca Padovani
Sony Computer Science Laboratories-Paris, Università di Trento
Italy

Daniele Nardi
Sapienza Università di Roma, CINI-AIIS
Italy

Donatella Tomaiuoli
Centro Ricerca e Cura di Roma, Sapienza Università di Roma,Università Tor Vergata Roma
Italy, Italy

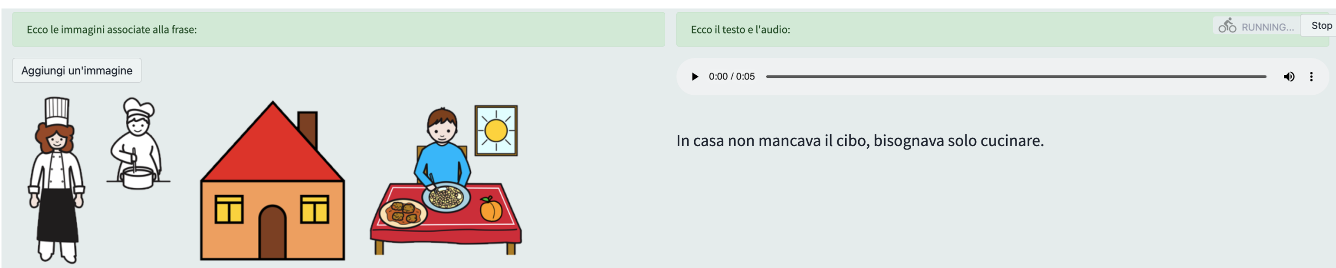

Figure 1: The first module of the NLP-powered reading comprehension interface

ABSTRACT

A strong reading comprehension ability is not only one of the cornerstones of a successful education but also a crucial skill for various aspects of our daily lives, both in the workplace and personal life. Nonetheless, text comprehension is a complex cognitive task that can prove challenging and frustrating for many children with learning disabilities. In this paper, we introduce an Natural Language Processing (NLP)-powered reading comprehension interface to explore the potential of modern NLP techniques in hybrid therapeutic practices, aiming to make the rehabilitation of text comprehension difficulties more accessible, automatic, multi-modal, and truly inclusive for students with special educational needs and disabilities in presence or at a distance. Our contributions are twofold. Firstly, drawing inspiration from psycho-linguistic theories, we have designed, together with speech and language therapists, an NLP-powered reading comprehension interface as a proof of concept for the treatment of reading comprehension. Secondly, we have deployed and evaluated our proof of concept in a general public setting and initiated the evaluation of our preliminary results in a clinical setting.

CCS CONCEPTS

• **Human-centered computing → Accessibility systems and tools**; *User studies*; • **Computing methodologies → Information extraction**.

KEYWORDS

Text Comprehension, Natural Language Processing, Learning Disorders

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

ICMI '23 Companion, October 09–13, 2023, Paris, France

© 2023 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 979-8-4007-0321-8/23/10…$15.00
https://doi.org/10.1145/3610661.3616176

ACM Reference Format:
Martina Galletti, Eleonora Pasqua, Francesca Bianchi, Manuela Calanca, Francesca Padovani, Daniele Nardi, and Donatella Tomaiuoli. 2023. A Reading Comprehension Interface for Students with Learning Disorders. In *INTERNATIONAL CONFERENCE ON MULTIMODAL INTERACTION (ICMI '23 Companion), October 09–13, 2023, Paris, France.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3610661.3616176

1 BACKGROUND

Text comprehension is a complex process that requires the construction of coherent and well-formed semantic representations [7, 11, 12] that integrates syntactic, semantic and narrative aspects of language [13]. This includes challenges in grasping the sequence, relationships, implicit information (inferences), and deeper meaning of the text content" [3]. The whole process is then controlled by one's meta-cognitive skills [6]. Despite this complexity, the existence of a diagnostic category for text comprehension disorder is still a matter of debate. In the literature, children with text comprehension problems have been referred to in various ways: bad readers [14], readers with learning problems [19], and finally poor readers [6]. Text comprehension difficulty is associated with different clinical profiles related to language difficulties or disorders, dyslexia, and cognitive delays.

Despite the absence of clear diagnostic categories, reading comprehension is a fundamental skill in our age. It not only facilitates the acquisition of knowledge and cultural enrichment but also unlocks one's full potential in education and professional life. This is precisely why we believe there should be a concerted effort to develop systems specifically designed to rehabilitate in presence or at a distance reading comprehension skills. This paper presents a reading comprehension interface that utilizes advanced Natural Language Processing techniques. The goal is to investigate how these modern technologies can enhance hybrid therapeutic approaches for rehabilitating text comprehension challenges among students with special educational needs and disabilities. In the literature, we are aware of a few systems available in Italian for the integrated tele-rehabilitation of specific learning disorders and other special educational needs, including RIDInet [2]. RIDInet includes a series of activities, referred to as "applications," that address text comprehension skills, specifically working on semantic and syntactic inferential processes related to text comprehension. However, in the English language, two primary systems are available: 3D Readers [1] and CACSR [10]. 3D Readers allow users to read texts and choose between four options, adopting either a verbal strategy (creating a question) or a visual strategy (interactively moving images on the screen) to enhance text comprehension. The generated questions and images are immediately evaluated, providing learners with feedback to improve subsequent performance [1]. On the other hand, CACSR provides personalized instruction to enhance comprehension strategies [10]. It employs visual images, graphic organizers, mnemonics, self-questioning, and summarizing techniques [18], offering real-time feedback for continuous assessment.

Although these projects contribute to improved reading comprehension, it is important to note that 3D Readers, CACSR were primarily developed for educational contexts rather than therapeutic ones. Moreover, they are mostly used in schools and lack a specific focus on distinct clinical categories. Moreover, even if RIDInet was developed for clinical contexts, it does not incorporate Natural Language Processing (NLP) techniques. To the best of our knowledge, there is only one system that employs NLP algorithms: Open Book [4]. However, Open Book primarily serves as a reading tool for children with Autism Spectrum Disorder (ASD) [15]. Its key features include text simplification through NLP techniques and rich customisation functions, enabling users to quickly adapt the document presentation (font, font size, line spacing, colours) to their preferences. It also provides assistive tools such as dictionaries and images. However, Open Books was designed for a specific clinical population, namely children with ASD and focusing more on improving their decoding skills than reading comprehension per se.

Finally, it's important mentioning all the systems targeting telerehabilitation [5, 16, 20, 23] specifically for Dyslexia. Even if some of them include some exercises to rehabilitate the decoding of single words and sentences and engage with written content, we are not aware of systems specifically targeting reading comprehension skills.

For all these reasons, we believe that exploring the feasibility and usability of an NLP-powered interface, designed to aid in the rehabilitation of reading comprehension for children with special educational needs and disabilities, could address a significant gap in existing literature.

2 SYSTEM DESCRIPTION

When designing our proof of concept, we had three primary goals in mind. Firstly, our aim was to develop a system specifically tailored to children from the elementary school up to 11 years old students who have been diagnosed with reading comprehension difficulties. Secondly, we wanted to create an interface that could be used during therapeutic sessions, always accompanied by a professional when working with younger age groups. Lastly, we sought to involve potential end-users and stakeholders right from the initial design phase, this is why our proof of concept was designed together with speech and language therapists. To achieve this, we collaborated with highly skilled speech and language therapists who not only conducted the clinical evaluation but also actively participated in the development of our proof of concept. Their valuable feedback and comments on preliminary exercises and versions greatly contributed to its advancement.

In line with Kitsch and Van Dijk's model [11], we specifically integrated functionalities to enhance the creation of a superficial representation of the language, i.e. the first level of Kitsch and Van Dijk's model [11]. This involved enabling users to develop a strong understanding of the lexical and morphosyntactic aspects of the words present in the text. In our interface, this is accomplished in what we will indicate as "First Module" through the paper. The "Second Module" would like to help users training, with what Kitsch and Van Dijk's call the sequence representations and hierarchical organisation of information [11]; this effort is underway despite the fact that the exercises in this direction are currently in a very preliminary form; further work will be conducted in the future. Finally, in a very preliminary form, the "Third Module", aims to train the capacity of creating a broader mental model of language, i.e.integrating the information found in the text with the knowledge already possessed by the users [11]. This need is addressed by the Synset map, which will be further explained later.

In the next subsection, we will describe each module and the algorithms behind each functionality in more details[1].

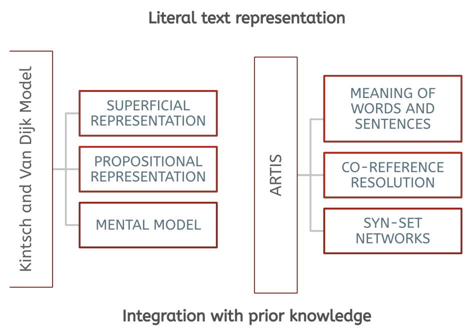

Figure 2: The different levels of [11] and the corresponding implementation

2.1 Superficial Representation: understanding words and sentences meaning

The interface starts by presenting the original text through an intermediate level of granularity: the sentence. This first sub-module is visible in Figure 1. As we can see, in this sub-module, the text, appears on the right divided into sentences by the Spacy Sentencizer [2]. For each sentence, users have the possibility of listening to the text through speech synthesis, integrated by calling the Google text-to-speech API. We decided to introduce a speech synthesis module, since we wanted our proof of concept to be accessible to those who have a technical difficulty in reading, together with a reading comprehension fragility. In fact, speech synthesis is extremely useful in cases where slowness and difficulty in decoding are present, as it allows users to bypass the reading process, thus enabling them to focus on the content rather than on form of the passage.

On the left part of the screen, for each sentence, the three main keywords are extracted and their meaning is represented in the form of PCS (Picture Communication Symbols), the most widely used form of pictographs for our use case[3]. As a first step this visualisation module, the ability to visualize what is written can act as a support for text comprehension. Moreover, it allows for the activation of bottom-up strategies, in which the perception of the external stimulus guides the process of information processing. For extracting relevant keywords, Keybert [9] was used.

For evaluating whether the right keywords were extracted and where not misleading, speech and language therapists manually checked whether for the sentences chosen for the users studies the correct keywords were extracted. Once the keywords were extracted from the sentences, after lemmatisation, we used the Arasaac API [4] to link them to pictograms. In a previous version of this proof of concept, we fine-tuned Clip [5] on real images, using the Unsplash dataset [6]. Even if the algorithm was performing well computationally, the images returned, even if correct, were sometimes unsuitable to be seen by a fragile public. This is why, we thought that using pictograms, instead of real images, could make the whole process safer. Unfortunately, as far as we know, big dataset of PCS are lacking so data creation or augmentation techniques would be needed for creating robust text-to-pcs algorithms.

Following [13], in a second submodule, the interface presents the text through a lower level of granularity: the word. This second sub-module is visible in the supplementary material, see Figure 3. In this second step, the 10 words that are ranked as the rarest in the vocabulary of the Italian language for each selected text are presented to the user. For extracting those we used WordFreq [17], which provides access to estimates of the frequency of a word used in 40 languages, among which Italian. We decided to automatically extract the rarest words present in each text because, after long discussions with speech therapists, we found that, in many cases, children with reading comprehension fragilities are afraid of admitting that they do not know a particular word. If these words are selected by the user, the definition together with a Pcs representing the meaning of these definitions and the sentence in which the word was originally placed it's displayed. The definition, taken from the Oxford Dictionary API [7] is presented in both written form and spoken form through the use of speech synthesis.

2.2 Propositional Representation: hierarchical organisation

By taking into account the theories of Zachou et al [22] and the hierachical organisation of Kitsch and Van Dijk [11], we opted to embark on the initial phase of implementing exercises in this particular direction. As a very preliminary step, we implemented a co-reference resolution algorithm. This algorithm enables users to practice the linking relationship between different entities and pronouns within the text. While other exercises, such as cause-and-effect relationship extraction could be considered, we chose to start by pronouns precisely because difficulties in the recognition of pronouns seems to be particularly problematic. In this module, the text is divided into sentences and certain grammatical elements are highlighted using different colors depending on their Part of Speech Tags: common names with the light blue color, proper names with orange, pronouns with yellow. This second module is visible in the supplementary material, see Figure 4. This visual distinction helps users differentiate between various elements within the sentence and suggests the function of the lexical morpheme, aiding in the comprehension of the pronoun by identifying the correct referent. For each proposition, users are presented with four alternatives and asked to select the correct one. Among the four options, one is correct (determined by the co-reference resolution algorithm), two are misleading, and the last one closely resembles the correct answer. When the user provides the correct answer, the system provides positive feedback. To implement this functionality, we used the speech tagger of Spacy [8] for the parser and Coreferee [9].

[1]A recorded demonstration of our proof of concept can be found here : https://drive.google.com/drive/folders/16QuSPxHtBvOHKEkKhcbXSLVjnWzcpA1d
[2]https://spacy.io/api/sentencizer
[3]Pictographs are images created specifically for people who are illiterate due to age or disability. They offer the advantage of allowing from a very basic level of communication - adaptable to people with low cognitive levels or in very early stages - to a very rich and advanced level, although never as complete and flexible as that which can be achieved with the use of written language.
[4]https://arasaac.org/
[5]li2022cwelip
[6]https://unsplash.com/data
[7]https://developer.oxforddictionaries.com/
[8]https://spacy.io/usage/linguistic-features
[9]https://github.com/richardpaulhudson/coreferee

2.3 A first step towards the Mental Model: integrating a word with further knowledge

According to Kintsch and Van Dijk [11], comprehending a text involves more than just understanding its surface-level representation. A crucial aspect is the ability to connect newly acquired information from the text with existing knowledge. To initiate the rehabilitation of this deeper level of comprehension, we have implemented what we refer to as a Synset Network. This third module is visible in the supplementary material, see Figure 5. In the Synset Network, we focused on the smallest unit of granularity in a text, which is the word, and developed a method to establish connections between the words encountered in the text and prior knowledge. Each selected word is presented as linked to its synsets. Synsets can facilitate vocabulary expansion and retrieval, activating the user's pre-existing knowledge. Understanding the various meanings associated with a term not only ensures a deeper comprehension of its intended meaning, but also serves as an initial step towards constructing a mental model of the word. To determine the synsets of a word, we utilized the Merriam-Webster thesaurus [10].

3 EVALUATION

The first prototype of this proof of concept was unveiled at the Rome-European Edition Maker Faire [11].During the three-day event, visitors had the opportunity to explore the different sections of the platform, ask questions, and provide valuable insights and feedback. Parents, students from all walks of life, clinical specialists, teachers, software engineers, and the merely curious took turns. At the end of the experience, visitors were invited to complete an anonymous usability questionnaire which evaluated different aspects of the platform, such as sentence visualization, single word level visualization, semantic network, and pronoun exercises.A total of 110 individuals completed the questionnaire. The barplot, see Figure 6, demonstrates the platform's remarkable success among users, as the average score across all sections exceeded 8. Specifically, the section focusing on pronouns received the highest score, followed by the sentence display section, word display, and the general interface presentation along with the semantic network, which received nearly equal ratings. The differences in scores among the sections were minimal, indicating that the interface was well-received as a whole. In addition to rating each section on a scale from 0 to 10, users were also encouraged to provide their comments on the platform. One commonly shared comment pertained to the potential use of the interface for second language (L2) learning [12]. Moreover, we are testing the platform in a clinical setting with a small set of children (around twenty-five) with Neuro-developmental Disorders, selected from those attending the CRC in Rome [13], based on age, schooling and functional profiles. The aim of the clinical evaluation is to assess the effectiveness of the use of the platform in the rehabilitation of subjects with Neuro-developmental Disorders and to analyse the usability and satisfaction of both the children and the operators following them during the training.

[10]https://www.merriam-webster.com/thesaurus
[11]https://makerfairerome.eu/it/
[12]The results of the questionnaire can be accessed here: https://drive.google.com/drive/folders/16QuSPxHtBvOHKEkKhcbXSLVjnWzcpA1d
[13]https://www.crc-balbuzie.it/

4 DISCUSSION AND FUTURE CLINICAL EVALUATION

This paper introduces a reading comprehension interface specifically designed for children with language disorders. The interface was developed in collaboration with speech and language therapists and was inspired by psycho-linguistic theories. Our research has yielded two key outcomes, with one additional outcome currently in progress. Firstly, we have successfully implemented an initial proof of concept utilizing state-of-the-art Natural Language Processing tools. Secondly, we have deployed our system and gathered feedback from the general public to evaluate the interface. Additionally, we are conducting testing in a clinical setting with an initial group of approximately thirty children diagnosed with Neuro-developmental Disorders, selected from attendees of the CRC in Rome [14] based on age, schooling, and functional profiles.

The primary aim of the clinical evaluation is to assess the effectiveness of the use of the platform in the rehabilitation in presence or at a distance of subjects with Neuro-developmental Disorders,specifically targeting impaired text comprehension and fallouts on lexical semantic components and morphosyntactic comprehension (i.e. pronouns). This evaluation also seeks to analyze the usability and satisfaction levels of both the children using the platform and the professionals guiding their training. The testing subjects consists of children in primary school with diagnoses such as Specific Learning Disorder, Mild Intellectual Disability, and other Neuro-developmental Disorders associated with text comprehension difficulties. The experimentation period, currently just started, will span three months with two therapy sessions per week. The experimentation methods are taking the form of pre-testing to collect the functional profile, followed by reinforcement during therapy sessions with speech therapists and the platform (totalling 24 sessions), a post-test and satisfaction questionnaires. All subjects entering the project already have a functioning profile measured through the WISC-IV [21].

In the future, we plan to design more specific and precise exercises to rehabilitate the propositional representation studied in [11] and proceed further with the design of training exercises concerning the mental model [11]. Moreover, we will leverage the clinical results obtained from user studies to enhance the overall interface. Through this collaborative research effort, we also aim to raise awareness about language disorders, particularly highlighting the lack of formalism and computational implementations supporting reading comprehension. Moreover, we would like to use SUS, system usability scale [8] to compare the score of our interface with similary systems, given the extensive use in the literature. If the interface proves effective based on the user studies, its potential impact would be significant and wide-ranging. Firstly, success in improving text comprehension skills would take us a step closer to equitable and quality education, one of the main objectives that was added to the United Nations' 17 Sustainable Development Goals to be reached by 2030. The interface has the potential to empower individuals with special educational needs, who are often marginalized and undervalued in social and educational settings compared to their peers. As improved education correlates with better employment opportunities, the formation of informed and

[14]https://www.crc-balbuzie.it/

aware citizens, and enhanced relationships and social conditions, our interface could trigger a chain of positive effects on the society we all belong to.

REFERENCES

[1] [n. d.]. 3D Readers Software. https://www.sbir.gov/node/334161).

[2] [n. d.]. RIDInet. https://www.anastasis.it/ridinet/).

[3] American Psychiatric Association. 2013. Diagnostic and statistical manual of mental disorders. 5, 5 (2013).

[4] Eduard Barbu, M Teresa Martín-Valdivia, and L Alfonso Urena Lopez. 2013. Open book: a tool for helping asd users' semantic comprehension. In *Proceedings of the Workshop on Natural Language Processing for Improving Textual Accessibility*. 11–19.

[5] Prabal Datta Barua, Jahmunah Vicnesh, Raj Gururajan, Shu Lih Oh, Elizabeth Palmer, Muhammad Mokhzaini Azizan, Nahrizul Adib Kadri, and U Rajendra Acharya. 2022. 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, 3 (2022), 1192.

[6] Barbara Carretti, Cesare Cornoldi, Rossana De Beni, and Marta Romanò. 2005. Updating in working memory: A comparison of good and poor comprehenders. *Journal of experimental child psychology* 91, 1 (2005), 45–66.

[7] Morton A Gernsbacher, Kathleen R Varner, and Mark E Faust. 1990. Investigating differences in general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 16, 3 (1990), 430.

[8] Rebecca A Grier, Aaron Bangor, Philip Kortum, and S Camille Peres. 2013. The system usability scale: Beyond standard usability testing. In *Proceedings of the human factors and ergonomics society annual meeting*, Vol. 57. SAGE Publications Sage CA: Los Angeles, CA, 187–191.

[9] Maarten Grootendorst. 2020. KeyBERT: Minimal keyword extraction with BERT. https://doi.org/10.5281/zenodo.4461265

[10] Ae-Hwa Kim, Sharon Vaughn, Janette K Klingner, Althea L Woodruff, Colleen Klein Reutebuch, and Kamiar Kouzekanani. 2006. Improving the reading comprehension of middle school students with disabilities through computer-assisted collaborative strategic reading. *Remedial and special education* 27, 4 (2006), 235–249.

[11] Walter Kintsch and Teun A Van Dijk. 1978. Toward a model of text comprehension and production. *Psychological review* 85, 5 (1978), 363.

[12] Maria Chiara Levorato. 1988. *Racconti, storie e narrazioni: i processi di comprensione dei testi*. Vol. 246. Il mulino.

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

[14] Jane V Oakhill, Kate Cain, and Peter E Bryant. 2003. The dissociation of word reading and text comprehension: Evidence from component skills. *Language and cognitive processes* 18, 4 (2003), 443–468.

[15] Nikolay Pavlov. 2014. User interface for people with autism spectrum disorders. *Journal of Software Engineering and Applications* 2014 (2014).

[16] Sampath Rajapakse, Dasuni Polwattage, Umesha Guruge, Isuru Jayathilaka, Tharindu Edirisinghe, and Samantha Thelijjagoda. 2018. ALEXZA: A mobile application for dyslexics utilizing artificial intelligence and machine learning concepts. In *2018 3rd International Conference on Information Technology Research (ICITR)*. IEEE, 1–6.

[17] Robyn Speer. 2022. *rspeer/wordfreq: v3.0*. https://doi.org/10.5281/zenodo.7199437

[18] Maria Earman Stetter and Marie Tejero Hughes. 2011. Computer Assisted Instruction to Promote Comprehension in Students with Learning Disabilities. *International Journal of Special Education* 26, 1 (2011), 88–100.

[19] Diane L Swanson. 1999. Toward an integrative theory of business and society: A research strategy for corporate social performance. *Academy of management review* 24, 3 (1999), 506–521.

[20] Min Wang, BalaAnand Muthu, and CB Sivaparthipan. 2021. Smart assistance to dyslexia students using artificial intelligence based augmentative alternative communication. *International Journal of Speech Technology* (2021), 1–11.

[21] David Wechsler. 2003. Intelligence Scale for Children. (2003).

[22] Angeliki Zachou, Enrica Partesana, Emanuela Tenca, and Maria Teresa Guasti. 2013. Production and comprehension of direct object clitics and definite articles by Italian children with developmental dyslexia. *Advances in language acquisition* (2013), 464–471.

[23] Andrea Zingoni, Juri Taborri, Valentina Panetti, Simone Bonechi, Pilar Aparicio-Martínez, Sara Pinzi, and Giuseppe Calabrò. 2021. 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, 10 (2021), 4624.

A SUPPLEMENTARY MATERIALS

A.1 First Module : Screenshot

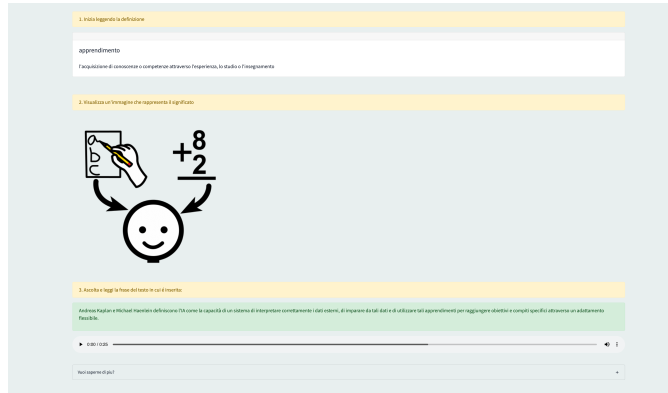

Figure 3: The second sub-module of the First Module.

A.2 Second Module : Screenshot

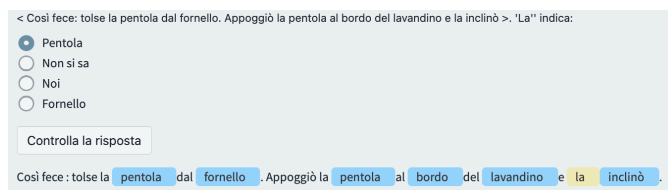

Figure 4: The Second Module of the interface: an exercise about pronouns.

A.3 Third Module : Screenshot

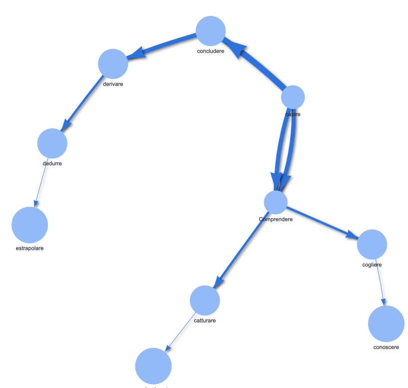

Figure 5: Example of a Synset map for the verb "to understand"

A.4 Evaluation : Barplot

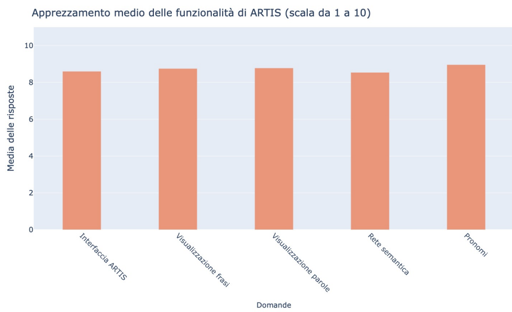

Figure 6: The results of the Google Form at Maker Faire in Rome in October 2022