



ANA: A Natural Language System with Multimodal Interaction for People Who Have Tetraplegia

Maikon Soares^{1(✉)}, Lana Mesquita^{2(✉)}, Francisco Oliveira^{2(✉)},
and Liliana Rodrigues^{3(✉)}

¹ Cear State University, Fortaleza, Cear, Brazil
maikon@dellead.com

² Federal University of Cear, Fortaleza, Cear, Brazil
{lana.beatriz,fran}@dellead.com

³ Le@d Lab, Fortaleza, Brazil
liliana.rodrigues@dellead.com

Abstract. To interact with a computer, users with tetraplegia must to use special tools/devices that, in most cases, require a great effort. In online education, these tools normally become a distraction, which might hinder learning. Solutions like tongue mouses, smart glasses and computer vision systems, although promising, still face problems of use. This paper introduces ANA, a natural language system which can hear the student and see what is being presented on the interface. With new affordance, learning objects (LO) can have their own grammar, which allows a much more natural voice interaction. LOs respond either by audio or performing the requested action. Tests performed with people with tetraplegics show that the creation of such a shared workspace brings a statistically significant reduction in effort while taking on online lessons and their respective workshops.

Keywords: People with tetraplegia · Education · Smart agent

1 Introduction and Theoretical Background

Promoting accessibility and inclusion of people with disabilities (PwD) in the various teaching and learning environments, besides being an important step in the exercise of citizenship, brings dignity to the PwD. The popularization of the Internet allows educational contents to be distributed on large scale, which makes distance education (DE) a great ally on creating opportunities of skill development for that population. Tetraplegics need special devices like orthoses and rods in the mouth, normally mounted by a caregiver and installed in an environment adapted to their reality. In addition, this type of equipment can generate discomfort to the user, as well as making the interaction cumbersome, frustrating and effortful. Putting on Norman's terms, this is a problem of execution gulf: tetraplegic user must to over great lengths to have the computer understand their intentions.

People Expect Other People to See What They See. When the locatability constraint is lifted, discourse production and comprehension effort is greatly reduced. We shall use Clark’s candle example to illustrate the concept. Imagine two people sitting at a table, conversing. There is a candle on the table. Conversant A asks conversant B: “Would you please light the candle for me?”. Since both are seeing the candle, conversant B needs no more extra detail to understand conversant A’s intent. They are both sharing the same workspace. Communication effort is at its minimum, as they are physically co-present [1].

With ANA, we try to replicate the scenario above. The creation of a shared workspace between student and assistant allows discourse interaction with objects (either learning objects or just regular widgets). In this way, different from other common wizards or chatbots, ANA is able to see what the user sees. LOs can have their own grammar, which adds considerably more context, which contributes even more for lowering communicative effort. For instance, imagine a LO which implements a quizz is being displayed. The student, looking at the screen would say: “The answer is letter B”, as a response to the object’s onset.

Another way to show the benefits of our approach is that, in many circumstances, users have to become accustomed to the interaction, and it might take sometime. In this particular case, the navigation task becomes a second to the learning. Multiple tasks are not desirable, specially for online education, as they compete for the user’s attentional resources. With time, the navigation becomes automatic and the “competition” lowers. We argue that, with ANA, because it is based on natural discourse interaction, time to reach maximum user performance, will much shorter than with other similar interactions.

This paper is organized as follows: after introduction and Theoretical background, Sect. 2 introduces the scenarios of use and target people of the study and is followed by related Work in Sect. 3. Section 4 shows the learning environment and describes ANA. Section 5 describes how this study was conducted, the methodology used, subject recruitment and results. Section 6 shows the conclusion of the study.

2 Scenarios of Use

2.1 Target People

According to Torrecilha [2], spinal cord injury is defined as any impairment in the spinal cord that causes deficits in the motor, sensory, and visceral function. According to the American Spinal Injury Association [3], the total absence of sensory and/or motor and/or autonomic functions below the level of the lesion, including at the sacral levels, characterizes a complete lesion. While the preservation of some function sensory and/or motor and/or autonomic levels below the level of the lesion, including in the sacral segments, characterizes the incomplete lesion. Injuries to the C1, C2, C3, and C4 vertebrae are characterized as “high cervical lesion” and result in paralysis of the arms, hands, trunk, and legs; lack

of control over spontaneous breathing; difficulties in speech. Lesions in vertebrae from C5 through C8 are characterized as “low cervical lesion” and result in paralysis in the hands, arms, trunk, intestine, and bladder, but the functions of cognition and speech remain.

Thus, people with injuries above C5 need the help of medical equipment, such as oxygen tubes to survive, and require the assistance of a third person for all other actions. Therefore, this research has targets students with low cervical lesion.

As previously mentioned, people with tetraplegia need special equipment to handle the computer. These devices can be corrective orthoses or even sticks that, placed in the mouth, allow the PwD to use the keyboard.



Fig. 1. Tetraplegic PwD using a computer

Figure 1 shows a quadriplegic PwD positioned in front of the computer with a stick in the mouth to compose a text in an electronic editor.

3 Related Work

Over the years, some research has been developed to improve somehow the way a person with quadriplegia use the computers. Steriadis [5] proposes an interface

adapted for tetraplegics. His approach projects the interaction using widgets to highlight elements that can be clicked by the user using a single-switch input device. The intervention still builds a virtual keyboard with word prediction to lessen the user's effort. Although its results bring a decrease of clicks to type, the user still has to reach the keys and widgets using the single-switch input device.

Alqudah [4] researches on a mouse controller that, connected to the face of the PwD by electrodes, can capture muscular movements and translate to mouse/pointer actions. For that, a hardware component constructed with Arduino processes signals obtained from electrodes positioned on the face of the quadriplegic user. Although the results show that it is possible to interact with the electrodes, it was still necessary to couple them in the face of the user, and this can be annoying and reduce freedom of movement. Figure 2 displays Alqudah's system.

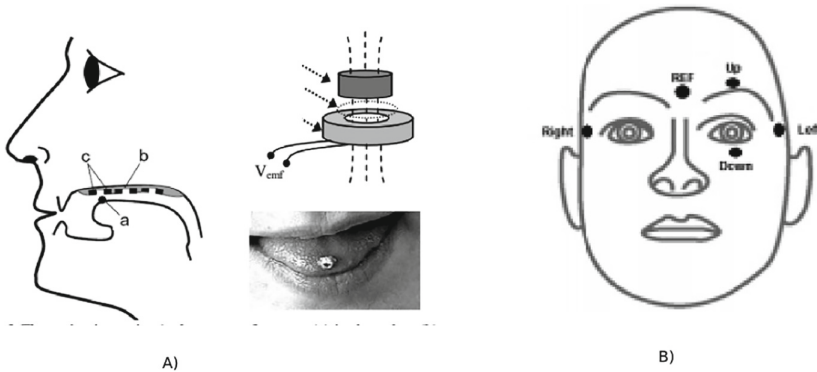


Fig. 2. (A) Tongwise mouse [5]. (B) face controlled mouse [4]

With ANA the student interacts with the computer through natural language, that is, no special equipment/gadget is required, unless, for reasons of comfort, one chooses to use a headphone.

Computer vision (CV) is a robust research topic within the scope of helping quadriplegic PwDs in the handling of the computer. Works such as Middendorp and others [6] investigate the use of visual tracking technology in online education systems for people with quadriplegia to navigate in classrooms. The study makes use of infrared cameras to capture the face image of the PwD. These images are processed and converted into commands. CV based systems either use computer's builtin camera or an external one (normally more expensive). They often need calibration, sometimes their accuracy are less than optimal and definitely constrain head movement, which is an issue for that population.

EID [7] proposes a visual tracking system for a quadriplegic PwD to send commands to a computer that controls a wheelchair. Soares and others [8], propose a gadget called TGlass. This equipment is a low budget smartglass designed to be distributed as material for online courses, respecting the anthropometry

of users and providing them comfort and freedom of movement. This equipment has input and output peripherals and uses eye tracking to provide PwD and computer interaction. The works cited above use hardware resources in order to provide accessibility. This approach often encounters barriers to development and deployment cost; besides, physical devices need maintenance and may be defective and require replacement of parts or entire equipment during a course, which can bring about high costs of logistics as well as disrupt the progress of the student. Solutions that require to be coupled in some way to the PwD user's body may need medical-orthopedic follow-up so that they will not cause you physical harm.

4 ANA the Accessible Navigation Assistant

4.1 Learning Environment Description

Oliveira [9] proposed an accessible Virtual Learning Environment (LE). His lab offers, free of charge, professional training courses in information technology, especially computer programming. However, this LE offers, to date, courses for PwDs with deafness or blindness. Therefore, to increase the number of users who can use benefit from better professional training, it was decided to face the challenge of providing accessibility to PwD with tetraplegia.

Figure 3 shows the initial screen to the introduction to programming logic course. Each lesson contains several Learning Objects (LOs), which together offer all course content interactively. Some LO's are: Webaula, Forums, Workshops, Exercises, and evaluation. Of these, WebAula contains all the theoretical information and some typing exercises and is subdivided into topics with a few pages of content on each topic, so it is the main LO and gateway to all others.

ANA assists the user in various navigation and interaction functions. With ANA, one can navigate through pages and topics within a Webaula, respond to quizzes, click buttons, interact with videos, listen to text reading and image description, and click on hyperlinks.

To navigate to a page 3 of topic 2 in a Webaula, for example, one needs to click on the topic he/she wants and then navigates using the navigation arrows at the bottom of the screen. Using ANA, the user can go to the desired page just by giving a single command, "Open page 3 of topic 2".

Virtual assistants are a reality that we live with on a daily basis. We can find them in the various operating systems: Apple Siri, Amazon Alexa, and Google assistant, as well as some video game consoles. Although very "intelligent", unlike ANA, these assistants still do not implement speech-addressable objects and therefore can not see and interact with the screen elements next to the user.

4.2 ANA's Description

This work proposes the Accessible Navigation Assistant (ANA). ANA allows a student to send commands to the computer in the form of speech. This speech



Fig. 3. Screen of programming logic web class object tool

is processed in some command known by the agent and then the action is performed on the system that returns multimodal feedback to the student. Figure 4 illustrates this interaction process. To listen to the user, the system has a voice recognition agent, which listens to the microphone connected to the user's computer. Once the voice is picked up, the actuating agent is called to display what has been recognized on the screen. Parallel to this process, the speech recognition agent synthesizes the speech into text and then passes the text to the natural language processing agent (PLN) to search for a command pattern there. The PLN agent makes use of the dialog flow service [10]. This service provides a range of features for creating a chatbot. The web service receives a text corpus called "user says" and through machine learning translates it into an intent or returns an error if it can not find an intent that matches the input text.

In the web context, HTML structures do not necessarily represent the real role of the component it generates within the system, a list of links, for example, can represent from a menu up navigation tabs to a higher level of abstraction. For the end-user of the system, the abstraction can reach even higher levels, and navigation tabs can have other meanings, such as "Class topics" or other types of tools.

In the AVA to which ANA was inserted, the screen interaction objects were cataloged according to their roles. These objects can be: "icons", "topics",

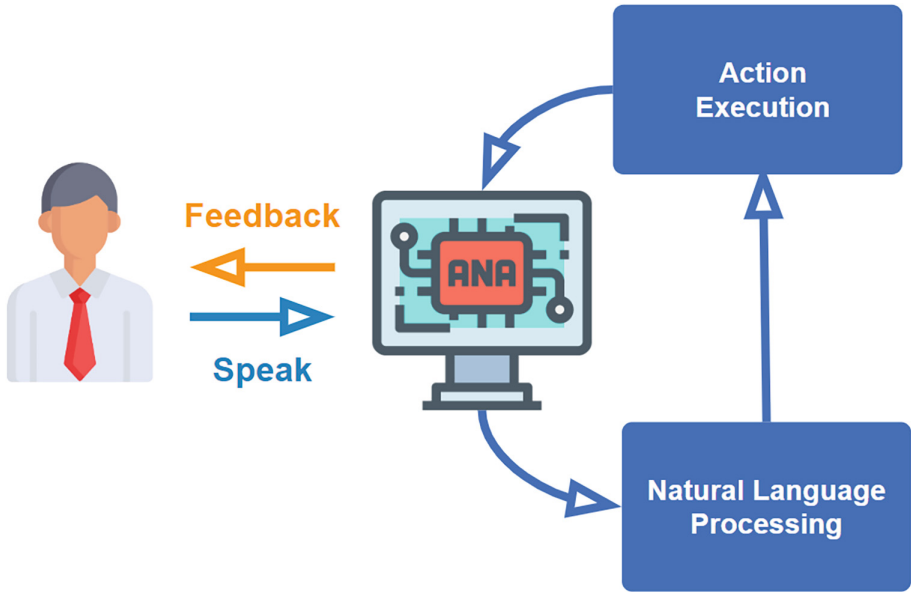


Fig. 4. ANA interaction flow

“pages”, “quiz” or others according to the need for task interaction. Assigning roles to objects allows the creation of a shared workspace between student and assistant. Once the dialog flow response has been obtained, the assistant knows exactly where each interaction object is, i.e. ANA can see what the user sees and therefore can access each object directly with only one command, making LOs addressable by the speech.

5 The Study

5.1 Methodology

The study compares differences in the execution gulf [11] of a user with tetraplegia when navigating in a Webaula of course of programming logic using their standard adaptation and using ANA. Table 1 displays subject profile, type of injury and adaptation tool. All subjects were male and volunteers. No remuneration or reward was offered to participants. Each participant performed the tasks at the location and time of their choice and using their computers. Thus, a scenario of actual use of a distance learning course could be simulated more faithfully. Participants were free to leave the experiment at any time, before or during the study procedures.

The data collection took place in the within-subjects model, that is, each participant performed the task using their standard adaptation and also using ANA. All trials were recorded on video to later be counted the number of steps for the execution of each task (execution gulf).

Table 1. Subjects division

Subject	Injury degree	Adaptation tool
1	C5	Stick in the mounth
2	C5	Stick in the mounth
3	C7	Finger pointer
4	C7	Finger pointer

The study was comprised of two phases: training and data collection. During training, subjects received a tutorial explaining how to use ANA and how the tests would be performed. Subjects were then allowed to use ANA for fifteen minutes. Subjects were trained in the customer service course.

The data collection phase was divided into 4 navigation tasks in a webaula LO of the course of programming logic. These task were as follows: (1) to navigate to the topic 1 and search in the pages for the image of the bust of Aristotle; (2) to navigate to page 5 of topic 2; (3) to find a quiz with 4 options, mark an option and observe the feedback of the chosen option; (4) to navigate to page 5 of topic 3 and click on a link and then navigate to page 8 of topic 3 and interact with a video (Play, Pause, Forward and Back).

The Fig. 5 shows a screenshot of the webaula LO used in experiment. In the figure one can see the topic navigation structure, a four option quiz and the page navigation structure.

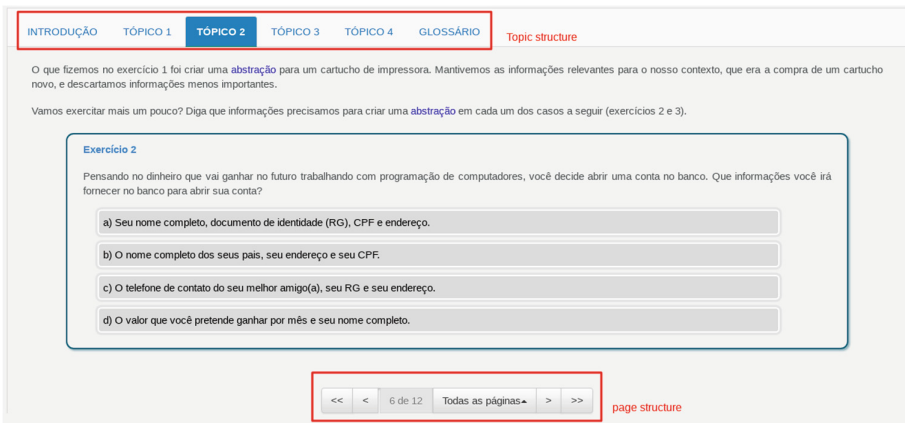


Fig. 5. Webaula LO

5.2 Results

Once data were collected, the number of steps performed by each participant took to perform the tasks using their standard means of accessibility and using ANA were counted. Tables 2 and 3 shows the number of steps of each subject in each task, without using ANA and using ANA, respectively.

Table 2. Step number of each subject to perform each task without ANA

Task	Subject 1	Subject 2	Subject 3	Subject 4
1	8	10	5	6
2	51	29	8	7
3	58	15	6	9
4	65	40	15	25

Table 3. Step number of each subject to perform each task with ANA

Task	Subject 1	Subject 2	Subject 3	Subject 4
1	4	6	2	6
2	5	5	4	5
3	4	6	3	3
4	15	10	10	7

Data do not conform to normal distribution (failed on Shapiro Wilk normality test [12]). Thus we ran Wilcoxon Signed-Ranked test at 95% confidence level. The results were: $z = 3.4078$ and $p\text{-value} = 0.00032$, quite significant.

The results show that using the ANA the student with tetraplegia can perform the proposed tasks with a smaller number of steps.

6 Conclusion

Provide accessibility in teaching people with tetraplegia is a major challenge for the academic community. Interaction must not get in the way of the user's goals. Making LOs "see" what the learner is seeing allowing a more natural and direct interaction can pave the way of online education for that population.

In this paper we introduce ANA, a conversational agent that enables the creation of LOs to whom PWD can talk while making discourse references to what is on the computer screen. We ran an initial study to compare the effort of tetraplegics while doing routine tasks on our accessible learning environment either with ANA or with their preferable method of interaction. Results show a great difference on the number of steps necessary to manifest their intentions to the computer.

This research is only in its infancy. We plan to develop a framework and specific grammars to enable content creators and curators ANA-like/google dialog interactions on their courses. Furthermore, more tests are needed to understand the impacts of ANA on learning outcomes.

References

1. Clark, H.H., Brennan, S.E.: Grounding in communication. *Perspect. Soc. Shared Cogn.* **13**(1991), 127–149 (1991)
2. Torrecilha, L.A., et al.: O perfil da sexualidade em homens com lesão medular. *Fisioterapia em Movimento*, pp. 39–48 (2014)
3. ASIA. American Spinal Injury Association. <http://asia-spinalinjury.org/>
4. Alqudah, A.M.: EOG-based mouse control for people with quadriplegia. In: Kyriacou, E., Christofides, S., Pattichis, C.S. (eds.) *XIV Mediterranean Conference on Medical and Biological Engineering and Computing 2016*. IP, vol. 57, pp. 145–150. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-32703-7_30
5. Steriadis, C.E., Constantinou, P.: Designing human-computer interfaces for quadriplegic people. *ACM Trans. Comput.-Hum. Interact. (TOCHI)* **10**(2), 87–118 (2003)
6. Van Middendorp, J.J., et al.: Eye-tracking computer systems for inpatients with tetraplegia: findings from a feasibility study. *Spinal Cord* **53**(3), 221 (2015)
7. Eid, M.A., Giakoumidis, N., El-Saddik, A.: A novel eye-gaze-controlled wheelchair system for navigating unknown environments: case study with a person with ALS. *IEEE Access* **4**, 558–573 (2016)
8. Soares, M.I.D.S., et al.: VISUAL JO2: Um Objeto de Aprendizagem para o Ensino de Programação Java a Deficientes Físicos e Auditivos através do Estímulo Visual—Um Estudo de Caso. *RENOTE* **12**(2), 1–10 (2014)
9. Oliveira, F.C.D.M.B., et al.: IT education strategies for the deaf. In: Hammoudi, S., Maciaszek, L., Missikoff, M.M., Camp, O., Cordeiro, J. (eds.) *Proceedings of the 18th International Conference on Enterprise Information Systems (ICEIS 2016)*, SCITEPRESS - Science and Technology Publications, Lda, Portugal, pp. 473–482. <https://doi.org/10.5220/0005922204730482>
10. DIALOGFLOW. DialogFlow <https://dialogflow.com/>
11. Norman, D.A.: Cognitive engineering. *User Centered System Design*, vol. 31, p. 61 (1986)
12. Shapiro, S.S., Wilk, M.B.: An analysis of variance test for normality. *Biometrika* **52**(3 and 2), 591 (1965)