MEDINFO 2015: eHealth-enabled Health
I.N. Sarkar et al. (Eds.)

doi:10.3233/978-1-61499-564-7-589

CA²JU: an Assistive Tool for Children with Cerebral Palsy

Flávio Arthur O. Santos[a], Carlos Augusto E. M. Júnior[a], Hendrik Teixeira Macedo[a], Marco T. Chella[a], Rosana C. do Nascimento Givigi[b], Luciano Barbosa[c]

[a]*Computer Science Department, Federal University of Sergipe, Brazil*
[b]*Speech and Hearing Department, Federal University of Sergipe, Brazil*
[c]*IBM Research - Brazil*

Abstract

This paper presents CA²JU, a hardware/software tool that aims to help individuals with severe speech or language problems in their communication in order to promote their social and digital inclusion. CA²JU is composed of two applications: CA²JU Accelerated, which makes typing faster by suggesting potential words to the user; and CA²JU Illustrated, which automatically converts a sentence of words into a sequence of pictographic symbols, allowing a user familiar with the symbols to verify whether the written sentence is correct. We have implemented, evaluated in a controlled scenario, and deployed CA²JU in a real environment with children with cerebral palsy. In the controlled settings, the results confirm CA²JU Accelerated speed up typing by reducing the number of clicks made by users, and CA²JU Illustrated obtained high accuracy by suggesting the correct pictograms from sentences. In the real scenario, the two use cases show that the children improved their communication and linguistic abilities.

Keywords:

Natural Language Processing; AAC; Web Interface.

Introduction

Communication is a key attribute for a human being. It is fundamental for expressing necessities and desires, enabling access to information, and socialization. Because of that, when a person lacks the power of communication, it is frustrating for her and for people around her. As a consequence, many assistive technologies have been proposed to deal with these limitations such as, for instance, Augmentative and Alternative Communication (AAC) [1],[2],[3]. According to the American Speech and Hearing Association [4]: AAC is a research and educational area of clinical practice for Speech Therapists who aim to "compensate and facilitate, temporarily or permanently for the impairment and disability patterns of individuals with severe expressive and/or language comprehension disorders". AAC allows communicative interactions of individuals with speech impairment, helping them in the process of language building and school interaction, which are fundamental for learning.

Natural language processing can have an important role in AAC [5]. For instance, it can reduce the input necessary to produce a text, increasing the communication speed; it can add semantics to written text to improve its comprehension etc. For instance, [1] described an English language word prediction tool that has a virtual keyboard. Mihalcea and Leong [6] proposed a system for the construction of pictorial representations by sentences. Finally, Nakamura and Zeng [7] described a text-pictogram conversion system to assist the creation of illustrated education for patients. Unfortunately, these and other initiatives focused primarily on the English language, lack proper integration between the software and the needed hardware, and do not provide an updated and intuitive Web interface which actually aids both the patient and therapist to use the software.

This paper presents the assistive tool called CA²JU, a hardware/software solution that aims to improve the communication in Brazilian Portuguese of children with cerebral palsy. The hardware is the interface of the system with the users. It allows the children with disabilities to provide the input to the software. The software is composed of two applications: CA²JU Accelerated and CA²JU Illustrated. CA²JU Accelerated predicts the next words to be typed by the user. It uses a NLP technique known as language modeling to perform this task. We implemented two different approaches for this task: N-gram and HMM. CA²JU Illustrated allows users to select pictograms to create sentences by converting sequences of pictograms into written sentences. To perform this task, it uses different NLP techniques such as Stemming and Name Entity Recognition to clean and add semantics to the sentences.

We evaluated CA²JU in a controlled setting and the results show that: (1) CA²JU Accelerated considerably reduced the number of clicks needed to be performed by the users to write sentences, where HMM showed a superior performance over N-gram; and (2) CA²JU Illustrated demonstrated a high accuracy of suggesting the correct sequence of pictograms from known sentences. We also report two use cases showing the experience of children with cerebral palsy using CA²JU. The child who used CA²JU Accelerated presented a great improvement in her communication and linguistic abilities. Initially, the child produced meaningless texts but, after many sessions using the application, she was able to write sentences with some semantic structure. The other child, who used CA²JU Illustrated, could communicate more effectively and improved his linguistic knowledge after starting to use the application.

The remainder of this paper is organized as follows. In Section 3, we explain CA²JU in detail, showing its input devices, architecture and the techniques used to perform its tasks. In Sections 3 and 4, we present the experimental evaluation of the deployment of CA²JU Accelerated and CA²JU Illustrated, respectively. We conclude Section 5 with our final remarks.

The Tool CA²JU

As we mentioned before, CA²JU is composed of two applications: CA²JU Accelerated and CA²JU Illustrated. In this section, we describe them in more detail.

CA²JU Accelerated

CA²JU Accelerated aims to speed up text input in Brazilian Portuguese of children with cerebral palsy. It does so by using word prediction to suggest to users the next word to input [8]. Figure 1 presents the user interface of CA²JU Accelerated. It contains a virtual keyboard used to input text, and a list of suggested words for the sentence currently being typed.

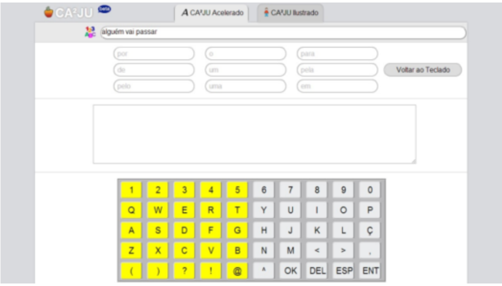

Figure 1 - Interface of CA²JU Accelerated

Language models are probabilistic models commonly used for word prediction. In this paper, we implemented two types of language models: N-gram and Hidden Markov Models (HMM).

Word prediction using N-gram

The joint probability of a given sentence *S* composed by $W_1 \ldots W_n$ words can be calculated using the chain rule:

$$P(W1, W2 \ldots Wn) = P(W1)P(W2 \mid W1) \ldots P(Wn \mid W1 \ldots Wn\text{-}1) \quad (1)$$

Since it is not computationally feasible to keep the whole previous context for all words, we limit the previous context of a word *W* to a fixed number of N-1 words. As a result:

$$P(W_n \mid W_1 \ldots W_{n-1}) \approx P(W_n \mid W_{n-N+1} \ldots W_{n-1}) \quad (2)$$

This is known as the N-gram model [9]. The assumption is that knowing the fixed previous context of *W* is enough to predict the probability of *W* (Markov assumption). For bigrams (N=2), for instance, the context is the immediately preceding word. Using bigrams, the joint probability presented in equation 1 becomes:

$$P(X_1, X_2 \ldots X_n) \approx P(X_1)P(X_2 \mid X_1) \ldots P(X_n \mid X_{n-1}) \quad (3)$$

To calculate the N-gram conditional probability, we use Maximum Likelihood Estimation (MLE). For bigrams, we have:

$$P(W_n \mid W_{n-1}) = C(W_{n-1}, W_n) / C(W_{n-1}) \quad (4)$$

where $C(W_{n-1}, W_n)$ is the number of times the words W_{n-1} and W_n appear together in the corpus, and $C(W_{n-1})$ the frequency of W_{n-1} in the corpus.

Word prediction using HMM

Another way to perform word prediction is using Hidden Markov Models [10]. In this work we have implemented the first order HMM, in which the current state depends only on the previous state and not on all previous states, so we can compare with the word prediction algorithm using N-GRAM, and then decide which algorithm to use. Hidden Markov Models is a temporal probabilistic model, where the state of the process is characterized by a single discrete random variable. In our case, the possible values of the variable representing the state are the parts of speech of the Portuguese language. The model is composed of a set of state variables:

$$X_t = \{X_1, X_2, \ldots, X_N\} \quad (5)$$

A set of evidence variables:

$$E_t = \{E_1, E_2, \ldots, E_K\} \quad (6)$$

A T_{ij} matrix of state transition probability i to state j:

$$T = [P(X_j \mid X_i) * P(X_i)] = \begin{bmatrix} T_{11} & T_{12} & \ldots & T_{1j} \\ T_{21} & T_{22} & \ldots & T_{2j} \\ \vdots & \ldots & \ldots & \ldots \\ T_{i1} & \ldots & \ldots & Ti_j \end{bmatrix} \quad (7)$$

A S_{ik} matrix of state transition probability X_i for E_k evidence variable:

$$S = [P(E_k \mid X_i) * P(X_i)] = \begin{bmatrix} S_{11} & S_{12} & \ldots & S_{1k} \\ S_{21} & S_{22} & \ldots & S_{2k} \\ \vdots & \ldots & \ldots & \ldots \\ S_{i1} & \ldots & \ldots & S_{ik} \end{bmatrix} \quad (8)$$

An array of evidence A_{1K} output variable probability belonging to the state X_i, X_i for each state:

$$A_{1k} = [P(E_1) \;\; P(E_2) \;\; \ldots \;\; P(E_k)] \quad (9)$$

To perform word prediction, we classify the input word according to its grammatical class, Part of Speech Tagging [9], then consult the matrix presented in equation (8) regarding the part of speech resulting, and then get the possible variables of evidence that we have.

CA²JU Illustrated

CA²JU Illustrated allows users to convert a sentence, written in Brazilian Portuguese, into an ordered pictogram representing the meaning of the sentence. Its goal is to help children with cerebral palsy in reading and writing learning, since it allows the child familiar with pictograms to verify whether the written sentence is correct by looking at its corresponding pictogram sequence. In addition, the application provides information about the words such as their part-of-speech tag, synonyms and the entity they belong to (e.g. if it is a place, an organization or a person), which can help them learn these concepts. Figure 2 presents the user interface of CA²JU Illustrated.

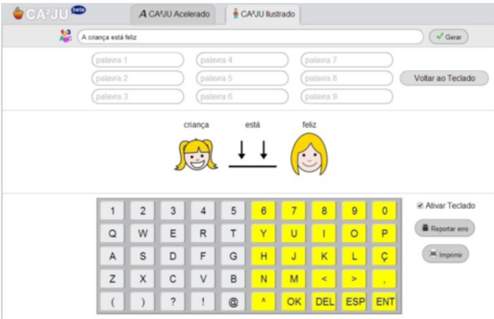

Figure 2 - CA²JU Illustrated's interface

Given a written sentence *S*, CA²JU Illustrated, first applies the RSLP Stemmer algorithm [11] in the words of *S*, so that suffixes of each word are removed. Next, it performs stopword removal, whose goal is to keep only the meaningful words in *S*. Since our original stopword removal, using a list of words, removed meaningful words, we used a Naïve Bayes classifier to identify important parts of the sentences: subject, verb and object (SVO), as well as greeting expressions, which we want to keep in the sentences.

The resulting sentences, with SVO information, are used to map to pictograms. The algorithm searches for the maximum sequence of words in each SVO portion in the sentence that maps to a pictogram in a given mapping dataset. If no pictogram is found, it does a similar process looking at the root words, identified by the RSLP Stemmer algorithm, instead of the actual words. If still no pictogram is found, a not-found pictogram is presented.

Last, CA²JU Illustrated runs a Name Entity Recognizer (NER) [9] to identify entities such as place, organization and person. From the tagged entities, CA²JU Illustrated is now able to add entity information to their respective pictograms. For the NER detection, we used a logistic regression model, also known as maxent [12]. Maxent performs a linear combination of predictor variables (features) to perform the classification. More specifically, it calculates the probability of that an instance x is member of a class c in the following way:

$$p(c \mid x) = \frac{\exp(W_c x)}{\sum_{c' \in C} \exp(W_c x)} \qquad (10)$$

Where W_c is the weight matrix of the features associated to the class c, and C is the set of classes. The instance x is classified in the class with the highest probability.

During training, the algorithm finds the model, represented by W, with the largest entropy from all models that fits the training data. To train the classifier for our task, we used the corpus of the Golden Collection of HAREM [13], which contains 225,000 words.

Buccal Input Device and Virtual Keyboard

The input module is composed of a hardware system for detecting the opening and closing of the mouth, signal conditioning, and the filtering and communication with the computer. The central element of the detector is the module composed of a magnet and a transducer that detects the magnetic field generated by the magnet into electrical signals.

The virtual keyword works in the following way. Initially, its cursor keeps switching between the left half of the virtual keyboard and the right half (see Figure 1). After receiving an input from the buccal device, the cursor vertically keeps changing from line to line on the chosen half. After a second input in the buccal device, the cursor keeps changing horizontally, going over each letter on the chosen line. Finally, the user can select a desired letter by moving the buccal device when the cursor is on the letter.

Figure 3 – Buccal Input Device

Experimental Evaluation of CA²JU Accelerated

It is difficult to assess word predictor softwares because many factors that have influence in the writing time are related to not only the software but also the user (user familiarity and level of disability) and the user environment.

Data Description

We used a generic corpus and personalized corpora for each user in Portuguese for this evaluation. The generic corpus used was Bosque Sintático [14], which has 14,844 words taken from generic news articles. The personalized corpora were obtained from social media conversations, e-mail messages, scientific works, blog posts of 8 users without disabilities. Table 1 presents information about these data.

Table 1 – Information about the personalized corpora

User	Number of words
1	15476
2	2795
3	16793
4	6084
5	33353
6	17426
7	13514
8	113211

Metrics and Test Set

To evaluate the different word prediction strategies, we use the writing speed of the users. For that, we use a metric called Keystroke (KSR) which is defined as follows:

$$\text{Ksr} = (1\text{-KP/KA})*100 \qquad (11)$$

KP is the number of inputs on the sensor needed to write the sentence using word prediction, and *KA* is the number of inputs without word prediction. We also used the *HR* metric, which calculates how many times the user selected a word from the suggested list. These metrics were used to calculate how the approaches' perform over one sentence per user (see Table 2).

GRAM and HMM Results

Table 2 presents the results for the N-Gram model and Table 3 the results for the HMM model.

Table 2 - Generic=G, Personalized=P

User	Sentence Size	Ksr in G	HR in G	Ksr in P	HR in P
U1	92	28,26%	3	39,13%	6
U2	66	18,56%	4	48,49%	7
U3	90	29,17%	7	68,06%	15
U4	89	14,61%	3	32,58%	12
U5	104	32,77%	11	33,04%	12
U6	129	14,73%	8	48,84%	14
U7	99	43%	7	51,60%	11
U8	114	15,79%	5	71,28%	18

Table 3 - Generic=G, Personalized =P

Name	Sentence Size	Ksr in G	HR in G	Ksr in P	HR in P
U1	92	34,17%	4	52,28%	10
U2	66	23,29%	4	66,66%	7
U3	90	33,43%	7	70,62%	15
U4	89	19,91%	4	46,34%	15
U5	104	37,58%	12	64,56%	16
U6	129	16,14%	8	54,91%	13
U7	99	48,38%	9	71,75%	14
U8	114	27,81%	8	71,92%	18

The results show that both word prediction models, N-gram and HMM, reduce the number of clicks for all users. In addition, the personalized corpora using HMM outperformed

N-gram in all cases. Another interesting result is that the size of the corpus not necessarily implies the reduction of clicks. For instance, user 5 has a worse result than user 7 but the corpus of user 5 (33353 words) is much bigger than user 7 (13514 words).

Use case of a teenager with cerebral palsy

We now show a successful use case of CA²JU Illustrated with a user with severe quadriplegic mixed cerebral palsy who we in this section call user M. After 74 sessions performed by the Speech and Hearing Department at Federal University of Sergipe, user M presented a great improvement in her writing abilities as we describe next.

In each session, the speech and language therapist worked with user M using CA²JU Illustrated along with linguistic activities. Before using CA²JU Illustrated, user M could not write due to her motor conditions. The first words produced by user M were meaningless, as one can read from the therapist report: "[...] we tested the virtual keyboard and user M wrote disconnected letters and then we asked whether she was writing something or just playing with the keyboard. User M smiled, looking excited and delighted with the possibility of handling the computer by herself "(12 ° Excerpt care; 13.08.12).

After understanding the operation of the system, user M wrote a single word when she wanted to convey some information and as she got more familiar with the system, she started to be able to tell stories using disconnected words, which were understandable for those who were mediating the story creation. Here is what the therapist reported from a story written by user M that she could understand: "[...] The story was that on holiday somebody had traveled to Los Angeles to play football. He had gone with his girlfriend who was a cheerleader. His team whose name was "ESFLAISSDL" had won "ERREST". It was interesting to note that she deleted letters a few times while writing these strange names to convince us that they actually existed "(Excerpt from 32 ° session; 19.11.12).

After a few more sessions, M. felt the need to better structure her writing in order to allow a more effective communication, not only with people close to her, but also to any other person. This is the therapist's repost regarding that: "[…] we realized a progress, a change in the writing of user M, since in previous sessions, user M had never worried about placing the pronoun + verb + noun to build a complete sentence" (Excerpt the 48th session; 11.04.13).

Experimental Evaluation of CA²JU Illustrated

Testing Scenarios

We evaluated CA²JU Illustrated with 10 users without disabilities as follows. First, the users were trained with respect to the pictograms by showing them some pictograms in sentences and their respective meanings. We then converted 10 sentences into sequences of pictograms, hiding the word associated with each pictogram. These sequences of pictograms were presented to the users, asking them to write associated sentences. Finally, we measured the similarity between the correct sentences and the written sentences by the users using the metric presented in Section 4.3.

Dataset and Metrics

The pictogram dataset used in the experiments contains 13,013 colored pictograms from CATEDU [15].

The metric used was the Levenshtein distance [16], which calculates the similarity between two sentences based on the number of character operations (i.e. insertions, deletions or substitutions) needed to change one word into the other. The Levenshtein distance is calculated by the following formula:

$$lev_{a,b}(|a|,|b|)=\begin{cases}\max(|a|,|b|) & ,\min(|a|,|b|)=0\\ \min\begin{cases}lev_{a,b}(i-1,j)+1\\ lev_{a,b}(i,j-1)+1\\ lev_{a,b}(i-1,j-1)+1_{(a_i\neq b_j)}\end{cases} & ,otherwise\end{cases} \tag{12}$$

Results

Tables 4 and 5 show the results for the 10 users and the 10 sentences.

Table 4 –Results from sentence 1 to 5

Users	Sent.1	Sent.2	Sent.3	Sent.4	Sent.5
User1	84,31	66,67%	100,0%	75,00%	62,22%
User2	56,00	100,0%	95,65%	83,33%	100,0%
User3	67,92	83,33%	35,00%	71,21%	100,0%
User4	30,00	100,0%	20,00%	20,41%	83,78%
User5	50,00	45,83%	73,91%	46,29%	54,34%
User6	40,00	62,50%	95,65%	65,71%	59,45%
User7	60,00	100,0%	61,76%	86,53%	89,18%
User8	72,00	100,0%	95,65%	42,26%	100,0%
User9	16,00	100,0%	41,66%	56,34%	100,0%
User10	34,00	100,0%	50,00%	25,49%	100,0%
Avg.	42,68	85,83	66,93	57,26	84,90

Table 5 - Results from sentence 6 to 10

Users	Sent.6	Sent.7	Sent.8	Sent.9	Sent.10
User1	81,48	58,82%	76,47%	62,90%	75,00%
User2	95,45	61,76%	82,97%	62,29%	75,00%
User3	98,18	58,82%	75,00%	64,70%	75,00%
User4	89,74	17,65%	72,09%	61,01%	100,0%
User5	41,66	23,52%	64,28%	57,62%	20,00%
User6	80,55	17,64%	54,76%	57,62%	40,00%
User7	94,44	52,94%	100,0%	59,32%	75,00%
User8	98,18	61,76%	95,23%	34,92%	75,00%
User9	62,04	61,76%	88,09%	66,10%	35,48%
User10	75,00	20,58%	100,0%	50,00%	100,0%
Avg.	81,67	43,53	80,89	57,65	67,05

The numbers show that the similarity between the sentences created by the users was fairly similar to the correct ones. For instance, only 2 out 10 users had an average similarity lower than 0.5. These results confirm that CA²JU Illustrated is really able to present the meaning of sentences in sequences of pictograms. Another observation is that for certain sentences the similarity among users varies a lot. This can indicate the users might have different interpretations of the pictograms, and they might need to be disambiguated in some way.

Use case of a teenager with cerebral palsy

User G is a 12-year-old male teenager, diagnosed with severe quadriplegic mixed cerebral palsy. He does not speak, neither studies at a special school. He has weekly speech and language therapy sessions where one of the goals is to teach him how to use CA²JU Illustrated. During the sessions where we used the software, user G understood well its functionalities. He interacted with CA²JU Illustrated by kicking a sensor device (see Figure 4). Using this input device, user G is able to select pictograms to form sentences.

Figure 4 – User G using the device.

The use of CA²JU Illustrated has positively contributed to his communication, improving his linguistic structure. We plan to allow him to use CA²JU Illustrated in other environments such as his home and school.

Conclusion

We presented in this paper CA²JU, our AAC tool composed of CA²JU Accelerated, which helps users with disabilities to type text by suggesting the next word, and CA²JU Illustrated, which helps users to compose sentences by using pictograms.

We evaluated both applications in controlled scenarios, as well as in real ones, with children with cerebral palsy. For CA²JU Accelerated, the results showed word prediction reduced the number of clicks needed to create a text, and its use by a teenager with disabilities helped improve his linguistic skills and communication. The results of CA²JU Illustrated confirm that it creates meaningful sequences of pictograms and also helped a teenager with cerebral palsy to improve his communication.

References

[1] Trnka, K., McCaw, J., Yarrington, D., McCoy, K. F., and Pennington, C. User interaction with word prediction: The effects of prediction quality. p. 34.

[2] BEUKELMAN, D., and MIRENDA, P. Augmentative and alternative communication.

[3] LIGHT, J. Toward a definition of communicative competence for individuals using augmentative and alternative communication systems. augmentative and alternative communication. p. 137–144.

[4] ASSOCIATION, A. S. H. Augmentative and alternative communication. p. 9–12.

[5] Newel, A., Langer, S., and Hickey, M. The role of natural language processing in alternative and augmentative communication. p. 1–16.

[6] Mihalcea, R., and Leong, C. Toward communicating simple sentences using pictorial representations. machine translation. p. 153–173.

[7] Nakamura, C., and Zeng, Q. The pictogram builder: Development and testing of a system to help clinicians illustrate patient education materials. p. 324–330.

[8] ANSON, D. The effect of word prediction on typing speed.

[9] Jurafsky, D., and James, H. Speech and language processing an introduction to natural language processing, computational linguistics, and speech. 2000.

[10] JORDAN, M., PEREIA, Paulo C. BRITTO JR, A. d. S., STELLE, L., and NOHAMA, P. Predição de palavras: Desenvolvimento de uma técnica baseada em Markov.

[11] Orengo, V. M., Buriol, L. S., and Coelho, A. R. A study on the use of stemming for monolingual ad-hoc portuguese information retrieval. Springer. p. 91–98.

[12] BERGER, Adam L.; PIETRA, Vincent J. Della; PIETRA, Stephen A. Della. A maximum entropy approach to natural language processing. Computational linguistics, v. 22, n. 1, p. 39-71, 1996.

[13] Coleção dourada do HAREM, http://www.linguateca.pt/primeiroHAREM/harem.html, 2014.

[14] Corpus bosque sintático, http://www.linguateca.pt/floresta/corpus.html, 2014.

[15] Palao, S. Catálogo de pictogramas coloridos. http://www.catedu.es, Mar. 2014.

[16] Levenshtein, V. I. Binary codes capable of correcting deletions, insertions and reversals. in: Soviet physics doklady. p. 707–710