A Loggable Aid to Speech: A Research Proposal

JÉRÉMY BARBAY*, CAMILA LABARCA, and BRANDON PEÑA-HAIPAS, Departamento de Ciencias de la Computación (DCC), Universidad de Chile, Chile

Validating that non human animals can communicate with humans using Augmentative and Alternative Communication (AAC) requires extensive logging, and traditional techniques are costly in resources and time. We propose to implement 1) a configurable “communication board” application aimed at small non human animals able to use touch interfaces, which not only emits human words associated to each button, but also logs such interactions; 2) a hardware keyboard to extend the use of such an application to larger non human animals unable to use a touch screen, but able to use large keys and 3) a centralized back-end gathering the logs from various devices, facilitating their study by researchers. We propose to validate the usability of such prototype solutions with two Quaker parrots for the application, a dog and two cats for the keyboard (and application), and a researcher in comparative psychology for the website of the back-end.

CCS Concepts: • Applied computing → Computer-assisted instruction; Interactive learning environments; Computer games; • Human-centered computing → User interface design.

Additional Key Words and Phrases: Animal Computer Interaction, Augmentative Interspecies Communication, Comparative Psychology, Digital Life Enrichment.

*Contact author.
1 INTRODUCTION

Various projects use Augmentative and Alternative Communication (AAC) techniques designed for humans with non human animals (referred to as 'Other Animals Than Humans' or OATHs from now on) to improve the communication between humans and OATHs. For instance, Cristina Hunger [6, 7] taught the dog Stella to use Augmentative and Alternative Communication (AAC) to better communicate. The How They Can Talk research group [5] is compiling the best ways to do so, in collaboration with the company Fluent Pet [14] selling specialized buttons to help people replicating their method. Any scientific validation of such communication abilities will require extensive logging of the interactions, which is usually performed by constantly recording the subjects on video, having the guardians marking each communication interactions, and researchers analyzing the context before and after the interaction. Such extensive logging either requires taking some manual notes at each press of a button, or recording on video all the interactions, and spending time analyzing such interactions afterwards. Such process is cumbersome and time consuming (see Figure 4 for an example of set-up to take manual note at each interaction, and Figure 5 for an example of the resulting manual log). Being tied by such a time consuming process is likely to restrict the amount of such usage data, which can be gathered.

Could a dedicated application and dedicated devices give a similar experience to that procured by voice-recording buttons and digital applications? Could it have the added advantage of the automatic recording of an extensive log of usages? Could researchers and educators study such log in order 1) to design AAC techniques more appropriate for each species (e.g. in terms of the amount of pressure required to activate the buttons, of the sound frequencies to record and play back), and 2) to list, evaluate and compare methodologies to teach the use of AAC techniques to OATH subjects and to their human guardians (e.g. the order and pace at which to introduce new words and the use of combination of words)?

We propose to design, implement and validate 1) a configurable “communication board” application, which not only emits human words associated to each button, but also logs such interactions, and allows a guardian to annotate the recent interactions for future analysis; 2) a series of hard-ware keyboards to be used by medium-sized OATH animals such as dogs and cats to interact with such application; and 3) a centralized back-end, gathering the annotated logs from various devices. We propose to validate our approach in two steps: first, we hope to validate the usability of the software and devices with OATH subjects, simply showing that they are able to press buttons to trigger the playing of recorded sound, in a way equivalent to what they do with existing voice-recording buttons; and second, we hope to validate the usability of the website gathering the usage information from the software (and, indirectly, devices) for researchers who wish to work on validating or invalidating the various concepts related to advanced communication using AAC techniques. In particular, we do not propose to validate the concept of advanced communication using AAC techniques: such an objective will take years (or decades) and is beyond the
scope of this research proposal. We just aim to advance the computational tools required to perform such research.

After more formally describing some previous results concerning the use of digital live enrichment applications in comparative psychology, and the use of AAC techniques by OATH animals in Section 2, we describe the system that we propose to develop (Section 3), the experiments we are planning to perform to validate its usability and assess its impact (Section 4), and the results that we are hoping for (Section 5).

2 BACKGROUND

Before describing the material and software solutions that we propose to design, implement and evaluate, we describe some previous results concerning the use of digital live enrichment applications in comparative psychology (Section 2.1), the use of AAC techniques by OATH animals (Section 2.2), and the technologies available for designing new AAC technologies (Section 2.3).

2.1 Digital Life Enrichment Applications

The study of the abilities of OATHs and the use of life enrichment activities in general, and digital ones in particular, have been interconnected from their very beginning. General preoccupation for the welfare of captive OATHs is at least 150 years old. Kohn [10] dates the first legislation about zoo animal welfare to 1876, with the “Cruelty to Animals Act” in the “Criminal Code of Canada”. Since then, the list of duties of such institutions has grown to include not only the basic welfare tenets of adequate feed, water, shelter, sanitation and veterinary care of their OATH residents, but also higher level concerns such as the handling and training of the OATH residents, their psychological well-being, the design of their enclosures, the preservation of their species, issues of environmental and conservation, and programs to breed captive OATHs. Kohn [10] mentions (in 1994) the “emerging field of psychological well-being in captive animals”, incorporating physical health, normal and captive behavior, and interactions with the enclosure environments and mentioning how environmental enrichment is an important component of this issue. He goes on to list innovations in life enrichment such as specialized toys and puzzle feed boxes (but no digital applications).

Yet, the use of digital applications to measure OATH abilities seems to predate Kohn’s report [10] by at least 10 years. In his discussion of the impact of game-like computerized tasks designed to promote and assess the psychological well-being of captive OATHs, Washburn [18] refers to a three decade old history in 2015, placing the beginning of such use sometimes around 1985. In 1990, when Richardson et al. [15] describe a Computerized Test System to measure some abilities in a population of rhesus monkeys, they mention that “the animals readily started to work even when the reward was a small pellet of chow very similar in composition to the chow just removed from the cage”, and that “the tasks have some motivating or rewarding of their own”.

Furthermore, OATH subjects seem to choose to participate in cognitive studies involving digital applications over other activities. Washburn [18] describes, among various other anecdotes, how game-like application for apes were developed as early as 1984, and how the subjects “chose to work on joystick-based tasks, even though they did not need to perform the game-like tests to receive food”, and “opted for computer task activity over other potential activities that were available to them”, a behavior described ever since 1963 [9] and christened as “Counterfereeloading” [11] in 1977. Lastly, he describes evidence that the subjects were not only motivated by food rewards, but also by the tasks themselves: when given a choice
between completing trials for pellets or receiving pellets for free but not being able to play the game-like tasks during the free-pellet period, the subjects chose to work for their reward.

2.2 AAC uses by OATHs

In the last decade, various solutions allowing OATH animals to communicate in human words by pressing buttons associated to recorded voice have been developed, promoted by Cristina Hunger [7] in the case of dogs and OATH animals of similar sizes (Section 2.2.1) and by Jennifer Cunha [3] for parrots (Section 2.2.2).

2.2.1 AAC for medium sized OATHs. In 2018, Cristina Hunger [7] described in her book [6] how the dog Stella successfully learned to use Augmentative and Alternative Communication (AAC) to better communicate. In the last four years, the book inspired a whole community of guardians to replicate their feats of communication (see Figures 1 and 3 for pictures of a dog with their AAC buttons), and various companies to sell sets of Recordable Buttons along with instructions about how to progressively introduce them to OATH subjects.

Currently, the research group How They Can Talk [5] is studying the best way to measure and validate the depth of communication performed between humans and OATHs using such technology, and the best way to teach both humans and OATHs to use those technologies to communicate. A key hurdle on this path will be to measure an adequate metric to measure the success of such communication methodology.

Also, there is precedence in lexigram keyboard usage for ape cognition and communication research [16], which sets a precedence for spatial cues in AAC devices, and evidence of dogs learning associative positions in the field of comparative psychology [1].

2.2.2 AAC for parrots. Even though parrots are known by the general public for their ability to learn and repeat spoken words, the ability to truly “speak” varies greatly among individuals of the same species, and among species [13]. Compensating such variability, some human guardians successfully taught some parrots to communicate via a touch screen interface: Cunha and Rhoads [3] describe how

An enculturated Goffin’s cockatoo (Cacatua goffiana) was taught to use a commercially available, android-based communication board. Through associative conditioning, the subject learned to press pictures representing items in categories for foods, beverages, activities, objects, and interactions.

See Section 2.3.2 for a description of the software they used, and Figures 2 and 6 for pictures of a monk parakeet subject using the CommBoards App [17] to ask for Water, and Figure 7 for a picture of the same subject using the first prototype developed for this research proposal to ask for a treat.

2.3 Input Device Technologies Available

We describe here the various technologies available to implement AAC with OATHs, both physically (Section 2.3.1) and digitally (Section 2.3.2), each with its own advantages and drawbacks, and each adapted to specific species of OATH subjects.

2.3.1 Physical Technologies. While some researchers in ACI designed, created and evaluated devices specifically designed for dogs, based on the way they naturally interact with their environment through biting, pulling and tugging, or simply touching with their nose [2, 8, 8], Cristina Hunger [6, 7] used classical AAC
devices, which are round buttons with a diameter of 10 cm, which play a previously recorded voice when pressed (see Figure 1 for a picture of a dog using such buttons). Albeit they were originally designed for humans, they seem to be usable by medium sized OATHs with enough strength to press the buttons: the users of the forum of How They Can Talk website [5] list a wide range of species such as dogs, cats, bunnies, rats, guinea-pigs, horses and peafowls. The buttons come in various colors to differentiate them, but it is recommended to glue them on a fixed board so that OATH subjects can use their relative positions to better identify them, as their color perception might not be the same as that of their human guardians.

The company Fluent Pet, in collaboration with the researchers from the How They Can Talk research group, is designing and selling specialized buttons to help people to replicate their method. Their buttons are smaller, and require less strength to be pressed (see Figure 3 for a picture of a dog facing two such voice buttons), making them appropriate for smaller OATHs such as small rodents, but are still too hard to press for some smaller OATHs such as small birds. They are of a uniform color, but come with brightly colored pads, so that OATHs can use a combination of their relative position on each pad and of the color of the pad to distinguish between buttons.

Getting data about the frequency with which each button is pressed either requires taking some manual notes at each press of a button, or to record on video all the interactions, and spend time analyzing such interactions afterwards. Such a process is cumbersome and time consuming (see Figure 4 for an example of set-up to take manual note at each interaction, and Figure 5 for an example of the resulting manual log). Being tied by such a time consuming process is likely to restrict the amount of such usage data, which can be gathered. The company is currently designing a product partially automatizing such logging, Fluent Pet Connect1, with a separate speaker for each tile of 6 buttons, and able to connect to a cell phone via WiFi. Originally scheduled to be released in the summer 2022 (in the North Hemisphere), the product was not yet available at the date of redaction of this work (August 2022).

2.3.2 Digital Technologies. Some guardians of parrots use the highly configurable CommBoards App. This application was designed by and for humans, the authors mentioning on their website the following:

---

1 https://fluent.pet/products/fluentpet-connect
CommBoards was born as a side-project meant to help our close friend whose child was diagnosed with ASD. We decided to create an app tailored to the child and parent’s needs, and share it with the world.

It has been used with success by some guardians of OATH trained in the use of touch interface [3]: see the pictures of a monk parakeet parrot using this application in Figures 2 and 6.

The CommBoards App application displays a set of visual cards arranged on a grid. Each card is associated to a short sequence of words (e.g. “I”, “Yes”, “No”, “Want”), and to an audio recording of a voice pronouncing such words. Some cards (e.g. “to play”) are associated to a separate grid of words (e.g. “Garden”, “Ball”, etc.) to be used to complement the original sequence of word. The sequence of words corresponding to the sequence of cards pressed is composed on the top of the screen. Selecting the special “settings” icon prompts for a challenge (e.g. typing a sequence of 4 digits), which in turn yields access to the settings of the application, where one can create, move and delete cards from the grids.

3 SYSTEMS PROPOSED

We propose to design, implement and validate a system composed of two components. The first component is a configurable software to be used directly by small sized OATHs via the use of touch screen interfaces (see a picture of a monk parakeet parrot using a first prototype of such an application in Figure 7). This component will be the one in charge of emitting sounds, to record a log of the input events, and eventually to record additional information such as input from the guardian or experimenter, or pictures or videos captured by the computing device (to complement the automatically generated log with some context). We describe this software component in Section 3.2. The second component is a physical input device to be used by OATHs unable to use a touch screen, but able to press keys designed according to their morphology (in terms of size and pressure amount required to activate each key), consisting in a set of input devices,
Fig. 8. Illustration of the process to produce a keyboard for OATHs with few keys: removing most plastic keys from existing keyboards, and gluing the remaining ones to larger panels.

Fig. 9. First prototype of a keyboard for OATH. The panels glued are made of light material, making the keyboard adequate for cats, but inadequate for chickens or dogs.

Fig. 10. Second prototype of a keyboard for OATH. The panels are made of poly-carbonate and both glued and screwed to the keys, making the keyboard adequate for chickens or dogs.

which can be connected to a computing device running the software described as the first component above. We describe our proposal and preliminary work on this component in Section 3.1.

3.1 Physical Keyboards

Dog and cat paws are too large to press the keys of keyboards designed for humans, and keyboards designed for humans are not sturdy enough to sustain being used by chickens or dogs. As there does not seem to be any keyboard with very large keys (e.g. 3 by 3 cms) which could be connected to a computer device available on the market, we propose to produce such devices ourselves, first by recycling existing keyboards designed for humans into keyboards more suitable for OATHs with a few keys (Section 3.1.1), then by building one from bare electronic components with a larger quantity of keys (Section 3.1.2).

3.1.1 Reconstructed USB Keyboard. To test the hypothesis that a USB keyboard could replace the boards of AAC buttons currently used in Augmentative Interspecies Communication [AIC] experiments, we designed and implemented two simple keyboards by merely removing most plastic keys from existing keyboards, and gluing the remaining ones to larger “keys”, to make it easy to press by OATHs (see Figure 8 for an illustration of the construction process).

The material of the large keys must be light enough so that the springs of the original keyboard actually raise back the key: for instance, keys made of wood would be too heavy. For the first prototype (see Figure 9 for a picture), we used “EVA rubber”, a light material left over from the packaging of the AAC buttons from Fluent Pet buttons. The material is light enough for the springs of the original keyboard to push the keys back in place, but too fragile to support the repeated impact of the chickens’ beaks or of dogs’ paws for any reasonable length of time. We plan to test the use of such a keyboard with cats. The second prototype (see Figure 10 for a picture), we used “poly-carbonate”, a light material commonly used for construction. The material is light enough for the springs of the original keyboard to push the keys back in place, and resistant enough to resist repetitive presses from dogs and chicken. We reinforced the glue with screws.

3.1.2 Built from Scratch Grid Keyboard. Various projects are described online, which aim to build a USB or Bluetooth input device from scratch. Arduino Keyboard [4] (if none with such large keys). The technology
most recommended for such projects seems to be MacroPads [12]. The amount of keys, whose pressing can be recorded, seems relatively large, especially if one does not need to record the pressing of too many keys at once (e.g. playing chords on a piano), but rather a single modifier key (e.g. the Shift key, potentially for guardians to press when modeling the use of the keyboard to a learner).

3.2 Digital Application

Whereas the CommBoards App [17] is highly configurable and produces a text sequence of the words produced by pressing its buttons, it lacks several options required for the purpose of this study, automatizing as much as possible the creation, gathering and analysis of usage logs of the use of ACC techniques:

- gathering the usage history on a remote server, to avoid the cumbersome and error prone process of gathering such usage history from distinct devices used;
- allowing the guardian to label such log usage (either while interacting with the application or device, or later on) to differentiate between key presses resulting from the modeling by the guardian or from the “real” usage by a subject, for instance a special, human specific, modifier key (such as the Shift, Alt or Ctrl keys of classical human keyboards) for guardians to be able to label in real time an interaction with the device(s) as one performed by the guardian rather than the subject;
- activating the buttons both through a press on the screen and through keyboard presses;
- running on both traditional computers and modern cell phones and tablets.

As such we propose to design, develop and validate a sequence of software solutions which mimic some of the features of the CommBoards App while complementing them with features required by our study. We propose to develop the following features:

1. playing sounds on buttons and keyboard presses (see Figure 11 for a screenshot of such a first prototype);
2. configure the number and positions of the buttons, as well as the keys and sounds associated to them;
3. configure a server to which the usage information is automatically sent;
4. configure the system to take a picture on each press of a button;
5. configure the system to record a short video after each press of a button;
6. configure the system to record a video stream continuously, but save short extracts corresponding only to the interval starting before and finishing after each press, for a configurable length of time, so that to save storage space and transmission bandwidth to the central repository of logs.

Pictures and Videos are needed to give a context to the communication being recorded, so that researchers can evaluate the intentionality of the communication attempt. For instance, the intentionality will be rated differently if the video shows the subject as waling on the buttons rather than pressing it with a beak or paw.

A first version of such a software, able to emit sounds, but not yet allowing the configuration of buttons nor to gather a log, was successfully tested by humans and OATHs: see a picture of a monk parakeet parrot using it in Figure 7.
Fig. 11. First Prototype of loggable communication board. The buttons are organized by categories: the first row corresponding to answers, the second one to actions and the third one to locations commonly used by the subjects. All the words which appear in the communication board are terms used by the guardian outside of the communication board, which the subjects were able to hear (and sometimes utter) before the introduction of the communication board, and with it the ability to emit such words. The words “COUCOU” and “OTRA” in particular are already pronounced by the particular subjects for whom this early prototype of communication board was designed: the word “COUCOU” was modeled to be uttered as a greeting and as an answer in a game of hiding labeled “peekaboo”, and the word “OTRA”, meaning “OTHER” in Spanish, was modeled to be uttered when asking to repeat an action, mostly when receiving a reward, but also to repeat an exercise.

4 EXPERIMENTS PLANNED

The experiments planned aim to validate the most basic hypothesis behind the design of the software and physical products developed, and to guide the ongoing design of more advanced versions.

4.1 Experimental Protocol

Each iterated phase of validation by OATHs will be preceded by an extensive validation phase by human subjects, to avoid any frustration of the OATHs subject, both for ethical reasons and for the success of the experiment, as such frustration would anyway taint the results of the experiments. We plan to execute each experiment with some medium size OATHs such as dogs, cats and chickens; and some small size OATHs such as monk parakeets, but will welcome the participation from any willing guardians and their protegees from other species:

1. Presented with a keyboard with a single button, the subject is expected to press it to get attention or access to a toy (for dogs and cats) or food (for chicken);

2. Presented with a keyboard with two, three or four buttons, the subject is expected to press one to get access to the corresponding toy (for dogs and cats) or food (for chicken).

Two is the minimum amount of buttons required to measure the ability to choose between two options. Increasing the amount of buttons will allow measuring how many associative positions the subject can learn, and facilitate measuring the intentionality of even short sequences of interactions, as opposed to sequences
of binary choices, which can be difficult to distinguish from random ones. Space placement is likely to be a factor in the usability of such keyboards. Having an extensive log of usage on keyboards with distinct space placements will allow researchers to emit and validate hypothesis on the key parameters of space placement on such keyboards for various species. Even though many more experiments can be designed using such devices and software, those experiments will be enough to validate our hypothesis, which we describe in the next section.

4.2 Ethical Guarantees

We are currently applying to the local Institutional Animal Care and Use Committee (IACUC) for the authorisation to run such experiments, for such application to be processed during the initial phase of development of both software and devices, and during the phase of testing with human subjects.

The following measures are embedded in the experimental protocols to guarantee the subjects’ welfare and to minimise their frustration:

1. Each version of each product (electronic and digital) is separately tested by two human subjects (its author on one hand, and the project coordinator on the other) before being tested with OATH subjects.
2. Each subject is offered the use of the products to select the activity (or toy) with which they wish to play for the next period. There is a “default” activity, less desirable, being offered as a substitute if the subject refuses to use the product.
3. No subject is deprived of food or water, and neither food nor water is used to motivate the use of the products by the subject.
4. The emotional state of the subjects will be measured by their willingness to continue participating in the activity, as the activities can be interrupted at any time by the subject, by simply going away;
5. The potential frustration and negative welfare of the subjects will be mitigated by always offering an alternative to the activity being evaluated: for instance, if a dog is required to press one of two keys to specify which of their favorite toys they want to play with, we plan to always have a third toy available for direct access (in addition to the option for the dog to just go away).

5 RESULTS HOPED FOR

By performing such experiments, we hope to validate

1. the usability of the application by OATHs to validate
   • the possibility of reproducing behaviors observed by medium sized OATHs with traditional buttons, even though the experimental set-up introduces an increased distance between the button and the speaker playing the sound; and to validate
   • the possibility of reproducing behaviors observed by small sized OATHs with the CommBoards App; to
2. the usability of the application by guardians, and particular their ability to
   • configure which button corresponds to which sound, to
   • indicate which key presses correspond to them modeling the desired behavior, as opposed to OATHs’ key presses, and to
annotate the OATHs’ key presses with some additional information, such as whether this key press is perceived as intentional or not;

(3) the usability of the website presenting the usage information to researchers, and in particular their ability to

- visualize basic statistical information about such logs, such as the frequency of usage of terms associated to buttons and the evolution of such frequencies in time, to
- select and download such usage information in a format that they can use to perform more advanced statistical analysis.

6 CONCLUSION

We have described our research proposal and preliminary work concerning the design, implementation and validation of a loggable and configurable communication board with OATH of various sizes. This work is still in its early stage, and we welcome any kind of criticism or feedback which could allow us more success with this project.

ACKNOWLEDGMENTS

REFERENCES

https://doi.org/10.3758/BF03203132.

