Expressiv

An *Expressive* System for Individuals with Facial Paralysis

Jessica Feuston, MS-HCI
Melody Jackson, Advisor
Georgia Institute of Technology
April 2015

Introduction

Emotional expression is a critical aspect of human communication. For individuals who have significant difficulty producing speech and-or language, emotional expression can be limited or impossible. Augmentative and alternative communication (AAC), which can enable and enhance communication interactions for these individuals, seldom offers effective support for conveying emotion. Limitations in graphic and text representations of emotion as well as prosodic constraints of digitized and synthesized speech used in high tech AAC systems can make emotional expression unnatural and inadequate. While frustrating to individuals using these AAC systems directly, this unnatural and inadequate expression can also be confusing or misleading to those interaction with the system as conversation partners or third party observers. This research examines ways of portraying human emotion through real-time depictions of facial expressions during communication.

Background

Within AAC there are multiple ways to support emotional expression, including text or speech output and static graphics. These means may not be entirely representative of the expression of natural human emotion. As communication is “essential for patient happiness” (Gelinas, O'Connor, & Miller, 1998), these AAC options can “improve relationships, increase participation in family and community life, offer a greater sense of independence and help a patient make important medical decisions” (Brownlee & Palovcak, 2007). However, successful communication requires accurate interpretation by a communication partner who may or may not be competent when interacting with an AAC user (Kent-Walsh & McNaughton, 2005).

Despite emotion playing an important role in communication and the large body of work examining emotional expression and perception, portrayal of emotion has received limited research attention in the design of AAC and has primarily focused on improving expression by modifying synthesized speech (Murray & Arnott, 2008; Murray, Arnott, & Rohwer, 1996; Wülfing & Hoffmann, 2011). AAC devices with emotive synthesized speech could certainly provide more access to personal expression. However, speech output is not the only way emotion could be conveyed through AAC. Facial expressions play a significant role in the expression and perception of emotion, but have not been investigated as a potential technology-based output modality for AAC users.

Research involving facial expressions has largely shifted to computational algorithms used to detect faces and facial expressions for recognition purposes (Gong, Wang, & Xiong, 2014; Chang, Bowyer, & Flynn, 2006). Some uses of these algorithms are novel, such as identifying key features and generating new, computer-synthesized expressions (Wang & Ahuja, 2003). Application for AAC, however, has been limited, as much of this work is instead incorporated into marketing and government identification techniques. Despite the current emphasis on computationally analyzed or modified facial expressions, previous work demonstrates the importance of expression intensity, involved facial features, and amount and direction of feature movement to facial expression recognition accuracy (Coren & Russell, 1992; Bassili, 1979. Much of this work is built upon six, universal facial expressions: happiness, sadness, anger, disgust, fear, and surprise (Ekman, 1999).

Due to the well-defined nature of these expressions, work in emotion and expression recognition through brain signal has primarily focused on one or a combination of those involved (Pham & Tran, 2012; Makeig, Leslie, Mullen, Sarma, Bigdely-Shamlo, & Kothe, 2011). Methodology within these studies is not always clear, however, and emotion and expression recognition and production applications with brain-computer interfaces are largely unexplored and unavailable for public consumption. Promising work has begun to explore the potential of motor cortex signal in regard to facial movement and expression (Morecraft, Stilwell–Morecraft, & Rossing, 2004; Achaibou, Pourtois, Schwartz, & Vuilleumier, 2008), but the effect of neurodegenerative diseases on motor cortex signal or imagined movement has not been thoroughly investigated.

Design Alternatives

The initial literature review contributed to three plausible system designs drafted for further investigation. These designs are briefly reviewed here, and were discussed with participants throughout the duration of the user requirements portion of this research.

Figure 1: The first image shows the proposed wearable, facial expressive system. The second image shows the expressive system on a computer monitor. The third image depicts the final drafted design, a projection system.

Wearable Facial System

In Figure 1, the wearable facial system is represented by Brow Motion, a mechanical, or robotic, replacement for human eyebrows. Regardless of what feature or features the prosthesis could come to represent, the wearable device was designed to be worn on or around the face.

Computational Avatar

The computational avatar was a digital representation of a human face. Though pictured on a large, computer monitor, the benefit of the avatar would be in device flexibility. Particularly in regard to Windows machines, which have the same opperating system for computers, tablets, and phones. This means the computational avatar could feasibly be incorporated into a variety of technological form factors.

Facial Projection System

The final design option is an amalgam of the previous two. The concept was to use a projector to directly display specific facial feature animations on an individual's face. This design option would require use of a projector and computer (or tablet, or phone) system.

Facial Feature Neutralization Investigation

The initial investigation into this space focused on establishing the recognition accuracy of facial expressions of emotion depicted through modified images and videos. Results informed the design of an emotive system that conveys natural and dynamic human expressions that can be used with other high technology AAC systems, such as speech generating devices, or as a standalone component.

Methodology

Seventeen participants took part in this study (10 males, 7 females), with an age range from 22 to 71 years old (mean age = 33 years; standard deviation = 15 years). Participants did not report any vision impairment, or had corrected vision through glasses or contact lenses.

This study of facial expressions of emotion included six emotions (i.e., happiness, sadness, anger, surprise, fear, and disgust), three facial regions (i.e., top, middle, and bottom), and 10 specific facial features (i.e., eyebrows, eyes, forehead, cheeks, nostrils, nose, mouth, lips, teeth, and chin). To create the images and video, I used Version 2 (CK+) of the Cohn-Kanade AU-Coded Facial Expression Database (Lucey et al., 2010). Emotion-relevant images were randomly selected using a random number generator. These images were sorted into eight different categories: Image – Unmodified, Image – Modified Top, Image – Modified Middle, Image – Modified Bottom, Video – Unmodified, Video – Modified Top, Video – Modified Middle, and Video – Modified Bottom. Based on the category of the image or video, I modified the source material by neutralizing certain features (Figure 1) and asked participants several questions including:

1. What emotion is being expressed?
2. Which facial features did you use to determine this emotion?

Figure 2: The first image shows the actor's neutral expression. The second image shows the actor's face when labeled as "happy". The third image neutralizes the top part of the actor's face by replacing the top portion of the face from the second image with the top portion of the face from the first image. Images © Jeffrey Cohn.

This study used a within-subjects design to determine how neutralizing specific facial regions affects the recognition of facial expressions of emotion. Each participant viewed 48 images and 48 videos (approximately 5 seconds in length) twice. These images and videos were organized into three randomized sets and assigned to participants. Eight images and videos of each emotion were viewed by participants, and 12 different images and videos were used per neutralization category.

Results

The mean recognition accuracy for females was 77 % and the mean accuracy for males was 73% (sample accuracy = 75%). We also found similar accuracy percentages for images (75%) and videos (75 %). Tables 1 through 4 display the accuracy results found per media type for each emotion and neutralization category. All recognition percentages were larger than chance (17%); however, there were differences of recognition for the same emotion in different media types and neutralization categories.

Emotion	Total	Correct	% Correct
Happiness	136	109	80.15
Sadness	135*	115	85.19
Anger	136	78	57.35
Fear	136	91	66.91
Surprise	136	114	83.82
Disgust	135*	98	72.59

Table 1: Image accuracy per emotion.

* Indicates removed trial due to lack of response.

Emotion	Total	Correct	% Correct
Happiness	136	117	86.03
Sadness	136	101	74.26
Anger	136	80	58.82
Fear	136	81	59.56
Surprise	136	124	91.18
Disgust	136	111	81.62

Table 2: Video accuracy per emotion.

Neutral Face Section	Total	Correct	% Correct
Unmodified	204	183	89.71
Top	202*	124	61.39
Middle	204	160	78.43
Bottom	204	138	67.65

Table 3: Image accuracy per neutralization.

* Indicates removed trial due to lack of response.

Neutral Face Section	Total	Correct	% Correct
Unmodified	204	166	81.37
Top	204	164	80.39
Middle	204	165	80.88
Bottom	204	119	58.33

Table 4: Video accuracy per neutralization.

Happiness was most often confused for anger (51% of incorrect occurrences), as were sadness (44%) and disgust (82%). Anger, surprise, and fear were most often confused for disgust when they were answered incorrectly (37%, 39%, and 42%, respectively).

The three facial features used the most in determining a response were eyebrows (22%), eyes (21%), and mouth (22%). In regard to facial regions, features from the top part of the face were used to determine emotion for 49% of all trials (middle = 12%, bottom = 39%).

Conclusion

Despite a small sample (n=17), results from this investigation demonstrated that facial expressions of emotion differ as to whether they are aided or hindered by movement. Happiness, anger, surprise, and disgust saw accuracy improvements between image and video trials. For these facial expressions, these results suggest that movement is superior to static images in recognition accuracy, and possibly provide evidence that static images, such as those used in some current AAC systems, don't convey as much information as is needed to interpret and recognize emotion by communication partners.

I also found an increase in accuracy when using movement in two of the four types of neutralizations (top and middle). The top neutralization category saw approximately a 19% increase in recognition accuracy between image and video trials. This increase could demonstrate the importance of movement of the middle and bottom facial regions. These results demonstrated the projection facial projection system design alternative is particularly feasible due to synthetic facial movement on the bottom portion of the face.

User Requirements Research

The second investigation into this space was designed to understand how facial paralysis affects the communication of emotion, particularly in regard to or conjunction with synthesized speech. Stakeholders not only included individuals with facial paralysis, but also healthcare practitioners, friends and family members, and caregivers. This investigation solely involved participants with direct or indirect experience with ALS. This constraint decreases the likelihood of the external validity of the investigation; however, it seems likely results will be generalizable to other, similar neurodegenerative diseases, such as Huntington's disease and Parkinson's disease.

Methodology

Nine participants (3 men, 6 women) ranging in age from 25 to 67 years old were included in this portion of the research. Participants were organized into one of three categories: individuals with

ALS, healthcare practitioners, and family caregivers. (Note: caregivers are specified as *family* caregivers in this occurrence because all *interviewed* caregivers were or are directly caring for an individual in their family with ALS. I certainly do not expect all caregivers to be familial, or all family members to be caregivers.) Some individuals had eligible attributes for multiple categories. The final categorization for these participants was determined by the amount of content-per-category derived during each interview. None of the participants in this investigation had a comparable amount of information for two or more categories. The following list details this investigation's participants:

- **Jane*:** Female, 25 years old, physical therapist. Jane graduated with an advanced degree within the past year, but has professional and academic experience with individuals with ALS. In addition to planning care for a multitude of patients, she has a personal initiative to leave all of her patients smiling after their physical therapy session in complete.
- **Suzanne:** Female, 26 years old, speech pathologist. Suzanne's primary experience is with individuals with ALS. Her most poignant memory regarding her patients is a woman with locked-in syndrome she worked with for several months – many of which she did not initially realize the patient was cognizant due to the patient's inability to communicate because of severe pain.
- **Monica:** Female, 62 years old, physical therapist. Now retired, Monica's experience lies primarily in the pediatric realm with individuals with varying degrees of paralysis and physical constraints. The majority of Monica's patients were unable to speak, had limited mobility, and were unable to clearly express themselves via their facial muscles.
- **Marjorie:** Female, 57 years old, family caregiver. Nurse by trade, and ALS Care Clinic Nurse Director at the Alabama Chapter of the ALS Association. Marjorie's son, Sam, was diagnosed with ALS while he was an undergraduate. She and her husband, Randy, acted as his primary caregivers throughout the span of the disease process.
- **Randy:** Male, 58 years old, family caregiver. Mechanical engineer and CEO of Community Foundation Madison. Randy's son, Sam, was diagnosed with ALS while he was an undergraduate. He and his wife, Marjorie, acted as his primary caregivers throughout the span of the disease process.

- **Susan:** Female, 65 years old, family caregiver. Susan is her husband's, Ed, primary caregiver. As Ed is completely unable to speak, she often acts as his voice by reading aloud comments he's written down on paper or assistive, technological equipment.
- **John:** Male, 53 years old, diagnosed with ALS in 1988. John has very limited facial movement, and movement in general. He uses an EMG switch to communicate and interact with others in his environment.
- **Marie:** Female, 60 years old, diagnosed with ALS in 2012. Though limited in mobility in the lower portion of her body, Marie is still able to speak without technological assistive. This is, however, becoming more difficult and she is considering looking into buying a text-to-speech system.
- **Ed:** Male, 67 years old, diagnosed with ALS in 2014. Ed still maintains mobility in his limbs, but is unable to speak and has extremely limited movement throughout the lower portion of his face.

* All names have been changed to protect participant anonymity.

The primary initiative of this investigation was to determine design requirements for an expressive system. Discussion topics included: current technology usage (both to understand communication techniques and equipment necessary for comfort and support); facial mobility and expressions; preferences in the communication of information or emotion; and system form factor requirements, including comfort, size, and placement. Two interviews were conducted via email; six were over the phone; and one was in person. With the exception of the email interviews, in which total time was difficult to accurately measure, interview sessions lasted approximately fifty minutes. Only one session was conducted per participant.

Results

> "It's hard to get people to focus on me. I want them to look at me." – Ed

All three individuals with ALS were adamant about their preference for their conversation partner's focus during a conversation. It was important for them to be seen as contributing members of the discussion, and not simply a piece of technology. One individual described his frustration with conversation partners focusing on his text-to-speech device (an iPad) as opposed to him while conversing. The healthcare practitioners and caregivers interviewed admitted to this practice. Text-to-speech is a notably tedious process for all parties involved. For the actual user, it's physically cumbersome at times, requiring some form of muscle movement or brain signal to spell out individual words or select from images or phrases. It's time consuming; an aspect that does not escape notice of conversation partners, regardless of partner type (caregiver, physician, family member, etc.)

Two of the three individuals with ALS in this investigation utilize assistive technologies to communicate. John uses an EMG switch to type in all situations, as he has very limited movement overall. Despite facial paralysis, Ed is mobile and has less physical constraints than John. He prefers to text or write, and will often use a Boogie Board to convey messages to his friends, family members, and caregiver (Figure 3). Though this is certainly Ed's preferred method of communication, his wife and caregiver, Susan, finds this practice frustrating because none of Ed's writing devices incorporates audio options. In certain situations, such as dim lighting or while driving, Susan is unable to understand what Ed is trying to communicate because she is unable to look directly at what has been written. She would prefer for Ed to use one of the text-to-speech applications he has, such as Predictable. Ed, however, finds these communication technologies cumbersome and slow. He does not like navigating to these applications on his phone or waiting for the synthesized speech to speak for him.

Figure 3: A writing tablet called a Boogie Board.

"I would prefer [a facial system] if not invasive and complicated to use." – John

Both John and Marie conveyed their interests in a system that would be directly incorporated on the face. Ed preferred an option where the system would not be directly worn, but also mentioned the likelihood of him using any system was low due to his frustrations using higher technology AAC devices.

Comfort and ease-of-use were concerns for all individuals, though no specific concerns were directly mentioned by those with ALS. Healthcare practitioners and caregivers, however, emphasized the importance of the safety and impact of the physical system. Jane mentioned, "As long a system isn't obstructive, I don't see a problem with using it." Both Monica and Suzanne echoed this sentiment. Suzanne, as a speech pathologist, was particularly expressive in her discussion about any system's interaction with the mouth. She largely works on mouth and throat exercises, such as swallowing, and so stressed any proposed system could absolutely not hinder these regions or challenge the patient's breathing pathways. She felt a projection system, so long as it did not fully obfuscate the lower facial region, would work well in this regard. Her primary concern was comfort and safety involved with a physically wearable expressive application.

Despite the general agreement between stakeholder groups found in conversation focus and comfort, there was discrepancy in regard to which expressions were perceived to be the most important:

"I'm a really happy person. I want to be able to express that." – Marie

"I need to know if they're in pain, or are upset or afraid." – Suzanne

Marie's comment largely summarizes what all interview participants with ALS had to say in regard to the importance of expressions. Happiness, joy, and pleasantness were all mentioned and emphasized. Healthcare practitioners and, to a lesser extent, caregivers were concerned with being able to immediately identify patient discomfort, confusion, stress, anger, sadness, and pain. This way of thinking seems to be largely healthcare practitioner driven. Marie, John, and Ed didn't once mention a want to express anything other than a positive emotion until specifically asked about the ability to convey a sense of medical distress. While they did all agree that some negative expressions are important – such as an indication of pain – the fact remains that these individuals did not immediately consider negative expressions in regard to importance. In fact, anger and sadness were identified by individuals with ALS as two expressions they wouldn't mind missing from an expressive system.

This discrepancy is certainly interesting and may speak to individual experience. The healthcare practitioners I spoke with seek to identify and modify negative emotions within their patients. These same individuals are well-versed with alternative communication methods widely available in hospitals, including communication boards (Figure 4). Communication boards are often notably lacking in emotional content. The example pictured below seems to be an exception in that four emotional states (angry, afraid, frustrated, and sad) are represented. This may well not be the norm. Both Jane and Suzanne mentioned the limited design and expressive capability of the communication boards within their hospitals. Frequently, only two expressive images are symbolized: a happy face, and a sad face.

Figure 4: An example of an image-based communication board.

Despite the expressive capacity of a brick wall, communication boards are beneficial in that they can visualize facial expression. Current text and text-to-speech communication options are limited in this regard. (Text less so due to the popularity of emoticons.) Synthetic speech often sounds unnatural and computational. Decidedly unhuman, despite the amount of progress made over the past few decades. The Predictable text-to-speech mobile application is groundbreaking in that it includes an Emote menu option that allows users to select different emotions to incorporate in their speech (Figure 5). These emotions are loosely defined, however, and by no means exhaustive. Options include emotes like “yawn” and “laugh”. Susan, whose husband, Ed, occasionally uses Predictable described the Emote feature succinctly: “While it sounds good in theory, in practice it doesn’t add much to communication.”

Figure 5: A screenshot of Predictable.

In addition to divergent opinions on what expressions should be displayed or first considered, there were also varying perspectives on the functionality behind *how* expressions should be displayed. Voluntary versus involuntary expression control caused quite the stir.

While all healthcare practitioners interviewed discussed the importance of having an accurate picture of their patient's mental health and state of mind, only two out of three considered it a good idea to involuntarily display expressions during a treatment session. Both of these individuals included the caveat that, ethically, this may not be the most appropriate method, but that it would help in their handling of the situation and treatment plan. Out of these two individuals, Monica fluctuated on the idea, stating that she wasn't certain that it made her comfortable, and that, in a personal or non-professional setting, she'd never have even considered it. Jane, the one healthcare practitioner that felt all emotions should be voluntary, emphasized the importance of expression as a means of conveying personality. By automatically displaying expression across an individual's face, the system could potentially be shaping other's views of them without their expressed consent.

Caregivers and individuals with ALS could also not find agreement in regard to this concern. Marie, for example, couldn't even find agreement in regard to her own opinion. Her initial thoughts on the subject were that all emotion should be displayed involuntarily. She considers herself a fun, open, and easygoing individual – a person with honest expression who doesn't hide who she is. However, on further consideration, Marie worried about certain relationship interactions, such as becoming angry with an individual. She talked about how she wouldn't want to necessarily express her anger. Especially if she were in a public setting. This speaks to the difficulty of the question, certainly.

Some participants mentioned including both options in a fully functional system; leaving the user of the system in control of how to express themselves. (Note: Going forward, if a way to represent involuntary expressive brain signals becomes available, this question will need to be explored with more depth and care. Active testing for varying periods of time to understand the impact on individuals and within different environments is also advisable.)

"Technology changed as the situation changed. But he needed to reliably interface." – Randy

When speaking about the assistive communication technology his son, Sam, had used, Randy mentioned the difficulty of reliably interfacing with various equipment and programs. As Sam's disease progressed, his physical limitations also increased. What may have worked one week to communicate, may not have worked the next. As Sam developed facial paralysis, he initially adopted compensatory expressive behavior (Bogart, K. R., Tickle-Degnen, L., & Ambady, N., 2012), including a thumbs-up gesture to express approval or positive emotions. When he was no longer able to clearly communicate this gesture, Sam used more technology-driven communication equipment, such as a device Randy called a "pediatric rollerball", which essentially was a large mouse Sam was able to control with his foot. However, Sam did eventually lose all physical mobility. The last device both Randy and Marjorie, Sam's mother and caregiver, mentioned their son using was an eye tracking technology. This option became obsolete for Sam as his eye muscles lost their strength.

Changing physical constraints throughout the disease process largely contributed to the steady change in Sam's communication technology. This was frustrating for both Sam and his parents as all involved were frequently required to learn entirely new technologies. Randy specifically mentioned that, at these times, his focus during conversations with Sam was largely placed on the technology, as he strove to understand it and troubleshoot any problems that arose. This speaks to a need for a system that can remain consistent throughout the disease process, and is also easy to learn, set-up, and troubleshoot when the need arises.

Conclusion

Following the evaluation of the interview content, it became apparent that the system's stakeholders – individuals with ALS, their friends and family members, caregivers, and healthcare practitioners – had specific, and sometimes divergent, system requirements. However, all participants did agree on one thing: that a human-based, visually expressive system could supplement and enhance current AAC to create a more complete and authentic conversational experience. Participants, particularly healthcare practitioners and caregivers, were concerned about the positioning and additional equipment required to run such a system, and emphasized the importance of it not being obstructive. To a lesser extent, individuals with ALS were also concerned with comfort, as well as the invasiveness of the system.

While all groups of stakeholders were, to an extent, interested in the portrayal of positive emotions, particularly happiness, individuals with ALS seemed to favor this expressive possibility over more negative emotions, such as fear or sadness or pain. Healthcare practitioners and caregivers mentioned the importance of recognizing and understanding negative emotions so that they could immediately treat the problem. Individuals with ALS, however, all mentioned being relatively easygoing and wanting to portray a positive face to the world. A robust system should consider all of these opinions and incorporate them as best as possible.

Expressiv

Based on the gathered information from the user requirements research, as well as the Facial Feature Neutralization Investigation and literature review, I was able to build and partially implement a prototype for an expressive system – aptly called Expressiv. Due to the somewhat divergent opinions of individuals with ALS that I interviewed – particularly in regard to system placement – I decided to combine two of the previously mentioned design options, the computational avatar and the facial projection system, to create a projection system capable of facial projection and body projection. This enabled Expressiv to have two distinct, though similar, modes: the first of which was a projection of a mouth, to be projected onto the lower portion of an individual's face, and the second of which was a projection of a head, to be projected onto an individual's chest.

Figure 6: The set-up of Expressiv for an individual with ALS using a motorized wheelchair.

Expressiv has three main technological components:

- **Projector:** AAXA Technologies LED Pico Projector.
- **Projector Holder:** WoneNice Universal Flexible Long Arms Mobile Phone Holder.
- **Programing:** Processing, a Java-based programming language, was used to switch between expressions. During user experience testing, this code was run on a Windows

machine – a laptop, specifically. However, Processing can run on *any* Windows device, including tablets and phones.

It's worthwhile noting the importance of a pico projector, though not necessarily the AAXA Technologies LED Pico Projector. (Though I certainly did not have any complaints.) Not only can these devices be purchased for approximately 120 USD, they're also very light and portable. Individuals using this system may have limited mobility, but still move around their environment via specific tool or equipment, such as power wheelchairs. It's important that Expressiv does not cause additional restrictions to mobility, or act as a tether to a specific area. Additionally, the small form factor of the pico projector is less likely to obstruct other equipment, such as ventilators or text-to-speech technologies, or objects, such as trays or, as pictured above, dogs.

Expressiv currently covers a limited range of expressions: happiness, sadness, anger, surprise, fear, and disgust. These six universal emotions were selected due to their recognizable and associated facial expressions. Expressiv is currently controlled by button presses – which *could* be a feasible option for individuals able to complete the physical action. (Like Sam's example above, however, this feasibility would decrease as physical limitations increased.) Ideally, future iterations of Expressiv will find it controlled by one or multiple methods of brain-computer interaction. Because the six aforementioned emotions have such strong associated expressions, it's possible a generalizable brain signal or process could be utilized to control the system. (Note: In this regard it's important to understand how this could impact the customizability of the system.)

As mentioned, Expressiv contains two distinct and separate modes – a facial projection and a chest projection (Figures 7, 8, and 9). All figures demonstrate the "happy" expression as a means for discussion and comparison.

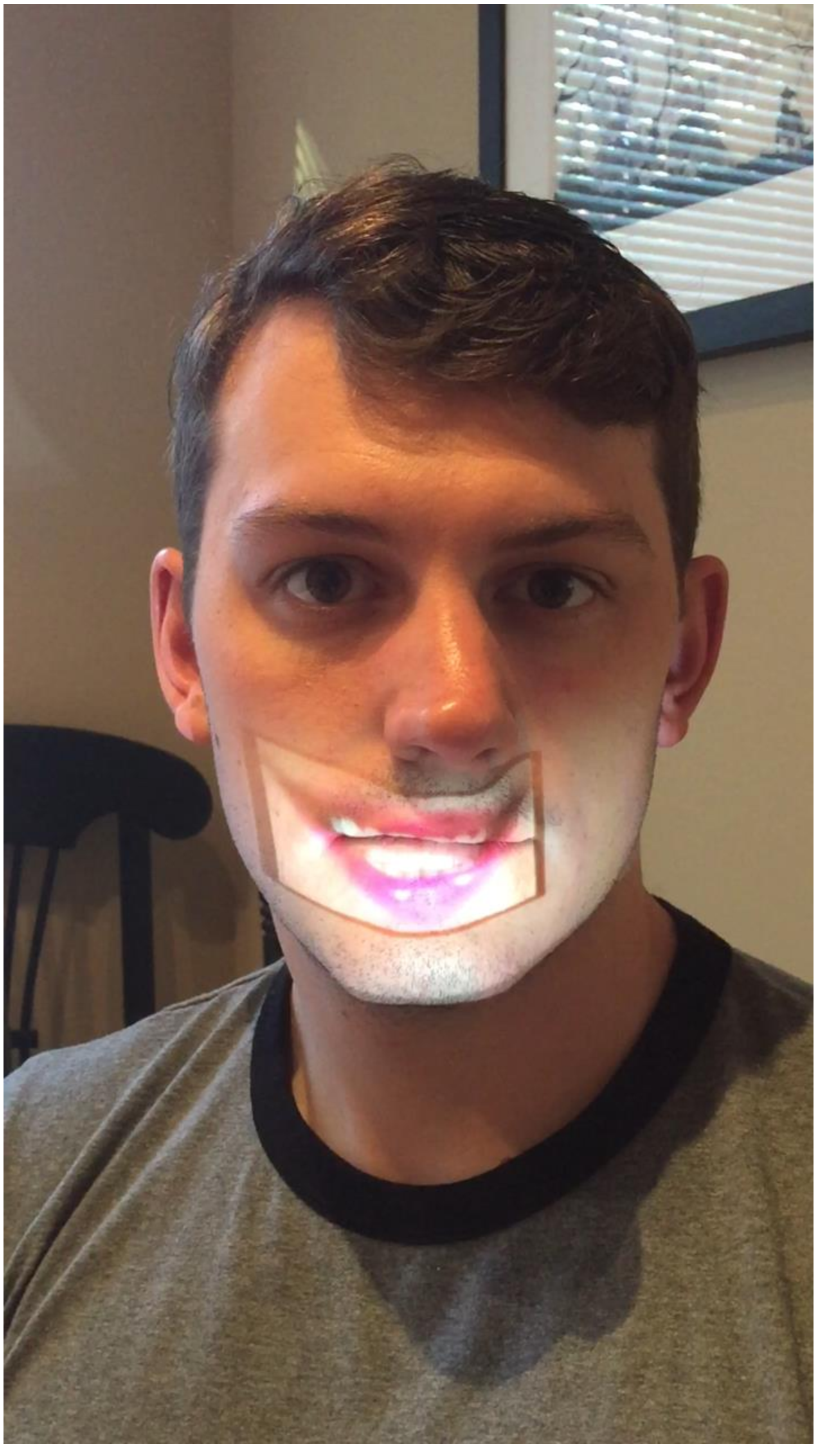

Figure 7: Expressiv's facial projection of a mouth projected directly onto the lower portion of an individual's face.

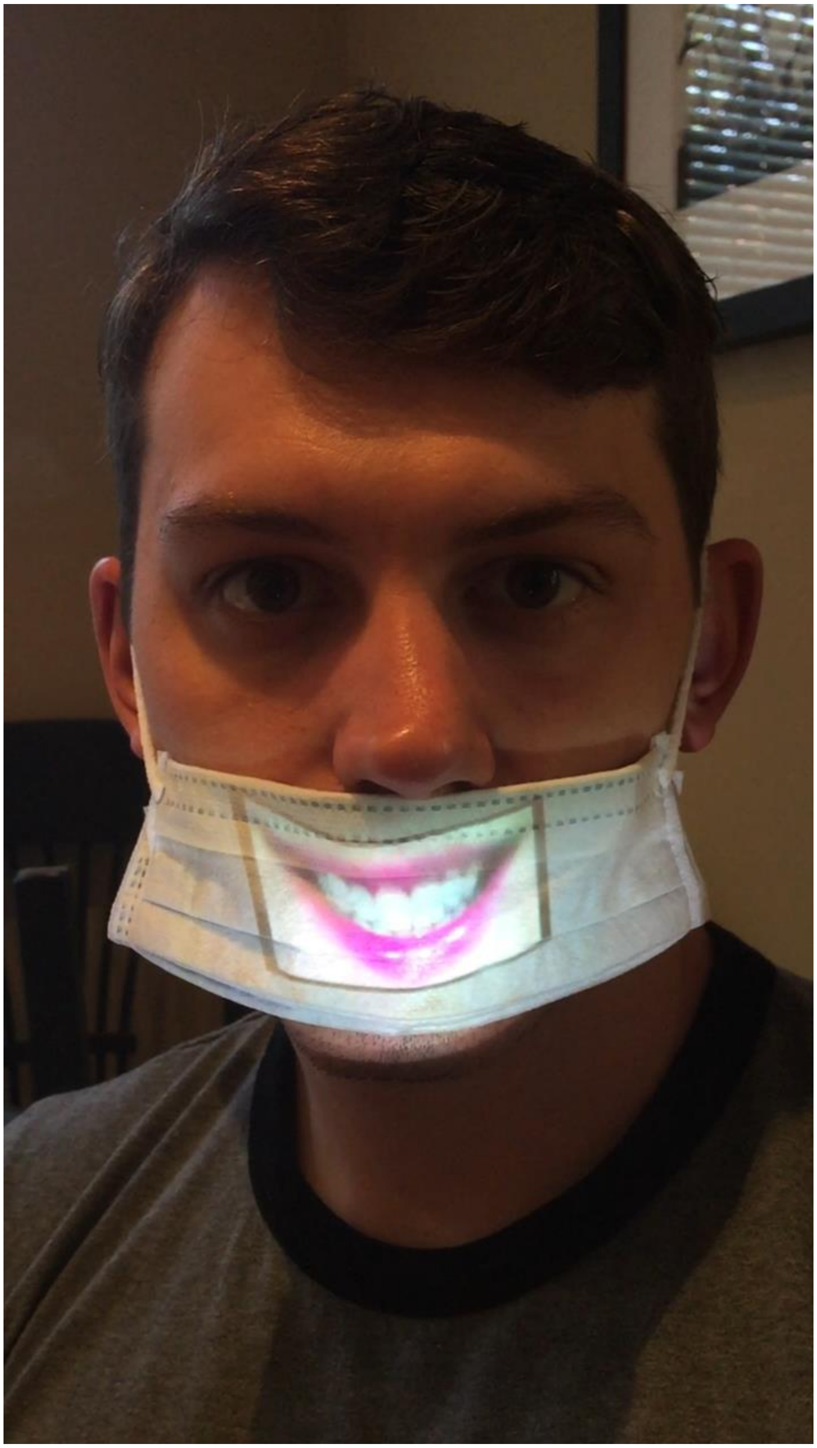

Figure 8: Expressiv's facial projection of a mouth projected onto a mask worn on the lower portion of an individual's face.

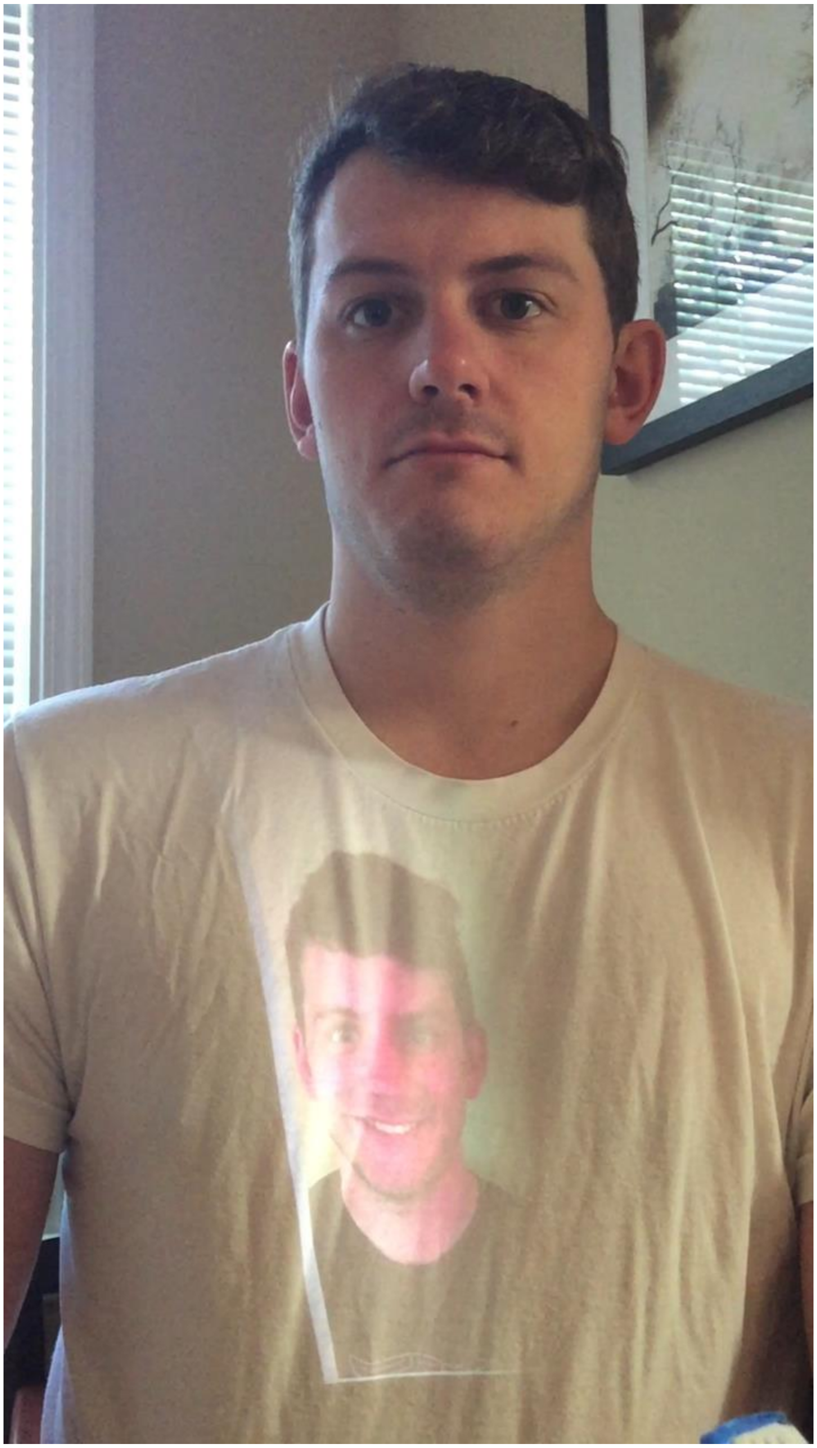

Figure 9: Expressiv's chest projection of a full head and partial upper body projected onto an individual's chest.

All images were taken in the same lighting. Figure 9 appears lighter than Figures 7 and 8 because the projector set-up is required to be further from the targeted projection area (the chest) for increased size of image and video. While still visible, it is worth considering the implications of lighting on a projection system. As Marie noted during the user requirements research, hospitals and doctor's offices are frequently lit with very bright, very unattractive fluorescent lighting. This current system was not evaluated in such an environment. Moving forward, real world scenarios, as opposed to laboratory run ones, will help with forward progress.

The two facial projection cases (mask and no mask) were developed due to the initial difficulty of mapping a projection to a curved, textured face. (Textured in that faces, particularly lips, are not flat.) The mask offers a clearer surface to directly project onto, which seems to increase the visibility, and possibly effectiveness, of the system.

Constraints

It should be noted that Expressiv has several constraints. The first, and perhaps most obvious, being the systems lack of portability. Certainly, when used in conjunction with a wheelchair or assistive mobility technology, Expressiv seems to work well, as it can be carried with the user. However, not all individuals with facial paralysis, even those with ALS, are limited in limb mobility. Ed, a participant from the requirements gathering study, though unable to move his facial muscles or speak, does not use anything to assist his mobility. Due to the nature of Expressiv, it is unlikely Ed, or individuals like Ed, would want to walk around carrying a projector and device equipped to run Processing – particularly as the projector would need to be directly pointed at the face or chest. (Note: Expressiv could be somewhat feasible for mobile individuals given certain situations, such as sitting down and watching television, or laying in bed.)

In addition to portability constraints, there are also visibility constraints. Projections, in general, do not show as well on dark surfaces as they do on light surfaces. For facial projections, this likely means that individuals with darker skin tones will be required to wear a lighter colored mask or cover over the lower portion of their face in order for the facial projection expression to be visible. For chest projections, individuals should wear lighter colors and avoid patterns, which does, of course, limit the expression of personality through fashion. While this may not be a large concern, it should be noted and discussed as a possible system detriment.

User Experience Exploration

Several methods of user experience testing were utilized to evaluate Expressiv. This is primarily due to the local and remote locations of the participants. In this last investigation, I sought to understand the impact of Expressiv on synthetic speech and conversation, as well as the perceived visibility of and confidence in the system. Comfort and safety were also evaluated to determine the feasibility of such a system in real life occurrences.

Methodology

Eight individuals (6 women, 2 men) participated in the user evaluation of Expressiv. Four of these individuals (Ed, Susan, Marie, and Monica) were also participants in the user requirements gathering research. The four remaining participants were recruited through fellow students interested in this research as conversation partners. (Note: The conversation partner category specifically refers to the fact that any person is a potential conversation partner for individuals using this system.) Participants identified as conversation partners included:

- **Kate:** Female, 30 years old. Kate is a graduate student at the Georgia Institute of Technology.
- **Mark:** Male, 26 years old. James is a game developer at an Atlanta-based company.
- **Lucy:** Female, 20 years old. Lucy currently works at a fast food restaurant and is in the process of returning to school to finish her undergraduate degree.
- **Yuki:** Female, 25 years old. Yuki is a graduate student at the Georgia Institute of Technology.

Due to the remote nature of some participants, user experience testing was conducted in person (four participants), over the phone (one participant), instant messenger (one participant) and via email (two participants). In person user experience testing, consisted of:

- **Video Survey:** Participants watched a video containing clips of synthetic speech expressing one of six emotions: happiness, sadness, anger, fear, surprise, and disgust. While watching the video, for each clip participants answered which expression they felt was being conveyed, how visible they felt the expression was (1: Least Visible to 5: Most Visible), and how confident they were in their selection of the expression (1: Least Confident to 5: Most Confident). The video clips had four different cases:
    - **No Expression:** Synthetic speech was used by an individual, but no facial expression or facial expression system was utilized.
    - **Facial Projection, No Mask:** Synthetic speech was used in conjunction with the facial projection system, by an individual who was *not* wearing a mask – meaning the projection was shown directly on their face.
    - **Facial Projection, Mask:** Synthetic speech was used in conjunction with the facial projection system, by an individual who was wearing a mask – meaning the projection was *not* shown directly on their face.
    - **Chest Projection, Avatar:** Synthetic speech was used in conjunction with the chest projection system.
- **Wizard-Of-Oz Roleplay:** Working with myself or another conversation partner, participants viewed the system in action – so to speak. Though brain signal functionality is not currently employed in Expressiv, I simply pressed buttons as individuals typed text to turn into synthetic speech. These roleplay conversations were discussed beforehand so individuals knew how to respond. Afterwards we discussed whether the emotional expression impacted the conversation in regard to both synthetic speech and natural conversation. Participants in this exercise also wore and used the system.
- **Discussion:** As mentioned, the impact of the system on conversation and synthetic speech was discussed. As was the system's comfort and ease of use.

Interviews conducted over the phone, instant messenger, and email had limitations in that holding conversations while using Expressiv was not plausible. So, instead, I acted out and filmed videos of Expressiv in action and shared these, as well as pictures of the system's set-up, with interview participants. The discussion revolved around the same questions as those posed during the in person user experience interviews.

Results

The first results gathered from the user experience testing were quantitative in nature and related strictly to the video survey.

Expression Type	% Correct	Visibility	Confidence
None	58	1	3.3
Chest Avatar	83	4.8	4.54
Mask	75	4.4	4.25
No Mask	70	3.0	3.9

Table 5: Results from the video survey questionnaire.

(Note: Only four individuals took part in the video survey. This aspect of the user experience testing was done as an exploratory study to determine whether a difference in expressive conversation content could be detected with Expressiv.) As demonstrated by the results in Table 5, the synthetic speech trial (no expression) of the video survey was the least successful in regard to participant recognition accuracy of expression (% correct), visibility, and confidence. All trials of Expressiv were more successful than solely the synthetic speech in all covered categories. The chest projection was the most successful in all areas - which is not surprising due to the projection's nature of displaying the entirety of an individual's face and, therefore, facial expression. Both of the facial projection cases, mask and no mask, did reasonably better than the trials only using synthetic speech. Though the statistical power is very low, these results are, nonetheless, optimistic and do seem to demonstrate a difference between synthetic speech with Expressiv, and synthetic speech without Expressiv.

"It's something you can respond to. It made me happy to see Yuki smile." – Kate

Qualitatively, the response to the system seemed just as positive as the quantitative results. All participants who reviewed the system preferred it over independent synthesized speech. Even Ed, the individual during the user requirements research who had mentioned his hesitance to use any high technology system, mentioned he would likely use the chest avatar. His wife and caregiver, Susan, further confirmed: "We both thought the avatar projected on the shirt works better. It really shows more expressions than just the mouth area by itself." Her comments certainly go in line with the quantitative results. However, I did find that some individuals in particular preferred the facial projection option over the chest projection.

Mark, a conversation partner, had this to say: "I prefer the projection on the mouth [with the mask]. It's more natural. It's also not distracting. Eyes flow in normal conversation." Mark brought up a very interesting and valid point about eye flow in conversation. During the user requirements investigation, questions designed to ask about visual focus during a conversation ("Where are you looking when…") attempted to understand conversation eye flow with individuals using assistive technology. Observing conversation partners in action, however, may lead to better understanding, as well as quantifiable data. Mark wasn't the only individual to prefer the facial projection. Marie, an individual with ALS who does *not* currently need AAC to communicate, talked about how she could see herself using the facial projection if her communication gets to a point where she's no longer able to accurately express herself.

In regard to the set-up of the system, I spoke again with Monica, a physical therapist. She was surprised by how small the system was (a sentiment shared by Marie) and expressed her belief that, "The set-up looks good. It doesn't get in the way of mobility, and seems like it could still work even if there were more equipment on the chair." This, of course, was extremely important to the success of the system, as Expressiv could in no way hinder other technology that would potentially be used by an individual with ALS. Marie, who "wore" the system, had this to say: "It wasn't uncomfortable, even when wearing the mask. I could see myself using it." Initial concerns had been particularly centered on the bright light associated with projection. Due to some videos being directly shone onto individuals' faces, it was a very possible reality that the light could shine into individuals' eyes and be very uncomfortable. The simple remedy to this concern, however, was to tape a small piece of foam over a small portion of the top of the projector light. This foam blocks the glare from the light, but still enables the projection to show clearly in both the facial mode and the chest mode.

Another trend of comment that stuck out during the user experience testing was customizability of expressions. Participants liked the ability to express themselves more than current AAC enabled, but voiced preference for more personal expressions as a means of showing personality. Conversation partners, as well as system users, shared this opinion.

Conclusion and Future Directions

Overall, I do believe Expressiv was a success. It's a step forward in an area that hasn't been widely or deeply explored. Though the final system may look worlds different from the current one, the user requirements gathered, and trends detected during the user experience evaluation, could certainly lend themselves to use in other designs seeking to explore the same space.

Future work should strive to incorporate brain signal functionality, or an actual interface where individuals with facial paralysis can physically control the system. (With something, perhaps, that is more intuitive than specified keyboard buttons.) Additionally, future work should seek to make such systems more customizable. While six expressions are certainly better than none, it's a shame for any system to be considered "better than nothing". Using Expressiv as a baseboard, more progress and growth can be made in regard to designing human-based, visual expressions. Throughout this project, a clear need has been shown for systems such as these to enhance conversation opportunities for both individuals with ALS, and others with facial paralysis.

References

Achaibou, A., Pourtois, G., Schwartz, S., & Vuilleumier, P. (2008). Simultaneous recording of EEG and facial muscle reactions during spontaneous emotional mimicry. *Neuropsychologia*, *46*(4), 1104-1113.

Bassili, J. N. (1979). Emotion recognition: the role of facial movement and the relative importance of upper and lower areas of the face. *Journal of personality and social psychology*, *37*(11), 2049.

Bogart, K. R., Tickle-Degnen, L., & Ambady, N. (2012). Compensatory expressive behavior for facial paralysis: adaptation to congenital or acquired disability. *Rehabilitation psychology*, *57*(1), 43.

Brownlee, A., & Palovcak, M. (2007). The role of augmentative communication devices in the medical management of ALS. *NeuroRehabilitation*, *22*(6), 445-450.

Chang, K. I., Bowyer, W., & Flynn, P. J. (2006). Multiple nose region matching for 3D face recognition under varying facial expression. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, *28*(10), 1695-1700.

Coren, S., & Russell, J. A. (1992). The relative dominance of different facial expressions of emotion under conditions of perceptual ambiguity. *Cognition & Emotion*, *6*(5), 339-356.

Ekman, P. (1999). Facial expressions. *Handbook of cognition and emotion*, *53*, 226-232.

Gelinas, D. F., O'Connor, P., & Miller, R. G. (1998). Quality of life for ventilator-dependent ALS patients and their caregivers. *Journal of the neurological sciences*, *160*, S134-S136.

Gong, E., Wang, T., & Xiong, J. (2014). Hidden Emotion Detection through Analyzing Facial Expression.

Kent-Walsh, J., & McNaughton, D. (2005). Communication partner instruction in AAC: Present practices and future directions. *Augmentative and Alternative Communication*, *21*(3), 195-204.

Lucey, P., Cohn, J. F., Kanade, T., Saragih, J., Ambadar, Z., & Matthews, I. (2010). *The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression.* Paper presented at the Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on.

Makeig, S., Leslie, G., Mullen, T., Sarma, D., Bigdely-Shamlo, N., & Kothe, C. (2011). First demonstration of a musical emotion BCI. In *Affective Computing and Intelligent Interaction* (pp. 487-496). Springer Berlin Heidelberg.

Morecraft, R. J., Stilwell–Morecraft, K. S., & Rossing, W. R. (2004). The Motor Cortex and Facial Expression:: New Insights From Neuroscience. *The neurologist*, *10*(5), 235-249.

Murray, I. R., & Arnott, J. L. (2008). Applying an analysis of acted vocal emotions to improve the simulation of synthetic speech. *Computer Speech & Language*, *22*(2), 107-129.

Murray, I. R., Arnott, J. L., & Rohwer, E. A. (1996). Emotional stress in synthetic speech: Progress and future directions. *Speech Communication*, *20*(1), 85-91.

Pham, T. D., & Tran, D. (2012, January). Emotion recognition using the emotiv epoc device. In *Neural Information Processing* (pp. 394-399). Springer Berlin Heidelberg.

Wang, H., & Ahuja, N. (2003, October). Facial expression decomposition. In*Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*(pp. 958-965). IEEE.

Wülfing, J.-O., & Hoffmann, L. (2011). *What does it mean to communicate (not) emotionally?* Paper presented at the Proceedings of the Second Workshop on Speech and Language Processing for Assistive Technologies.