

Ian Spangler  
iaspangler@yahoo.com  
Thesis Document: Chap. III/ First Draft  
March 22, 2007

## Human Avatar Project: Universal Emote

### **3.1 Overview of the Design Process**

As described in Chap. I, *Universal Emote* (working title) is a web-based application that visualizes collective thought patterns in the form of emotional expressions on an anonymous human's face. Designing and developing this project involved numerous iterations. The role of the user, the nature of user interaction, the form of the interface, and the conceptual framework were all variables that changed considerably in the process. In the project's refinement from a self-contained exploratory virtual experience to a customizable widget built around validated psychological data and external web-based content, qualities of rationality, approachability, and real world contextualization superseded philosophy, exploration, and poetry.

In developing prototypes of the application, a variety of technological design tools were employed including Macromedia Flash/ Actionscript for the interface and user interaction, Macromedia Coldfusion and Microsoft Access for the database, a video camera and blue screen for recording the avatar's facial expressions, and Adobe After Effects for compositing the facial components and preparing them for manipulation in Flash.

### **3.2 Representing Thoughts as Words**

At the outset, the goal was to create a virtual experience that externalized the notion of a thought process, highlighting and bringing awareness to certain universal patterns of the human

mind by making the user a conscious “thinker”. The relationship between the user’s thought patterns and emotions was to become a focal point of the interaction. In particular, the aim was to provide an accessible model of the *reflective* component of the human mind, the part of the mind that evaluates, judges, and conjures up or strengthens emotions in the body, leading to varying levels of neurosis. I have found that representing thoughts in the form of words is a valid approach in this case for a number of reasons. One is that most people would say that a large part of their thought processes, particularly higher level reflective thinking, occurs in the form of words in their native language. Secondly, through the development of verbal abilities in humans over thousands of years, language has become very intertwined with the realm of thought and has spawned the great meta-domain of psycholinguistics. A great deal of research has been conducted in the relationship between thought and language. Third, from a technical standpoint, using words in an interface allows for a lot more flexibility and variability in user input and computer analysis than does using images or some other form of visual symbolism.

### **3.3 Thought-Formation Mechanics**

I tested a variety of mechanics for inputting, or “forming” thoughts in the experience, gradually seeking to make the system more open and generative while preserving the psychological integrity of the simulation. For instance, one of the principles that I sought to imbed in the system was that a positive thought would more likely be followed by another positive thought than a negative one, and vice versa. Having a word pool from which users could create thoughts would allow the system to change the word pool according to every thought that was formed, increasing the relative numbers of one type of word and decreasing the number of other types. This manipulation was based on the premise that users would be looking to make

things “happen” in the experience and would not interact with the system in a way that is truly reflective of their thought patterns. It was also based on a desire to convey to users the notion that their thought processes are *not* as free as they might believe they are. Under actual life circumstances, certain personal habits may impede positive thinking or subconscious agents can take over. Moreover, as already mentioned in Chap. I with reference to Carl Jung, our thoughts are not as individual as we may think. We are limited by certain universal predispositions.

### *3.3.1 Prototyping Technology*

Based on a metaphor of the collective unconscious, I had planned to make the application a simultaneous multiple user experience – many users affecting one avatar with their thoughts. However, before dealing with the complexities of involving multiple users, and because I wanted the system to be as entertaining for a single user as it would be for many at once, I sought to quickly simulate a single user’s interaction. To prototype the thought-formation mechanic, I chose Macromedia Flash as an authoring tool. Its native programming language, Actionscript, is ideally suited for developing richly interactive and dynamic online multi-user experiences like this one.

### *3.3.2 Active Thought Formation: Dragging and Clicking on Words*

In the first prototype I devised, the user and the computer take turns dragging a word or phrase from a moving stream of words into the universal man’s (avatar’s) “mind bubble”. The computer checks the set of words currently in the bubble against a list of meaningful sentences (stored in a database) and as soon as it finds a match, it quickly displays the words in the appropriate arrangement to form the matching sentence. Afterward, the next word thrown into

the bubble clears the sentence and starts a new thought. At this early stage, I was entertaining the possibility of the experience being a game in which users would try to produce positive thoughts while avoiding negative ones. The involvement of the computer here was intended to introduce a bit of a challenge by making it more difficult for users to ensure that a positive thought would form in the mind bubble. In testing the mechanic, seeing the computer suddenly shoot a random word into the mental mix proved to be too disruptive and disconcerting.

In addition to the above, it became apparent from analyzing this prototype that users would want a quicker result and more flexibility in the way that they combine words. A later version eliminates the set of prescribed sentences and allows any combination of words to be used to form a sentence up to a two-line limit. To encourage users to form grammatically sound verse, I categorized the set of test words by part of speech and wrote a simple set of grammar rules in an XML document for Actionscript to parse and check for basic noun-adjective-verb agreement. Words that grammatically follow previous ones in a sentence are highlighted in the word stream while those that do not are faded, providing a word-by-word guide for users to follow in forming sentences. The program also contains rules to detect when a complete sentence has been formed so that the mind bubble can be cleared for the next thought. In addition, I changed the mouse mechanic from “drag-and-drop” to a simple click, which feels less tedious and more intuitive.

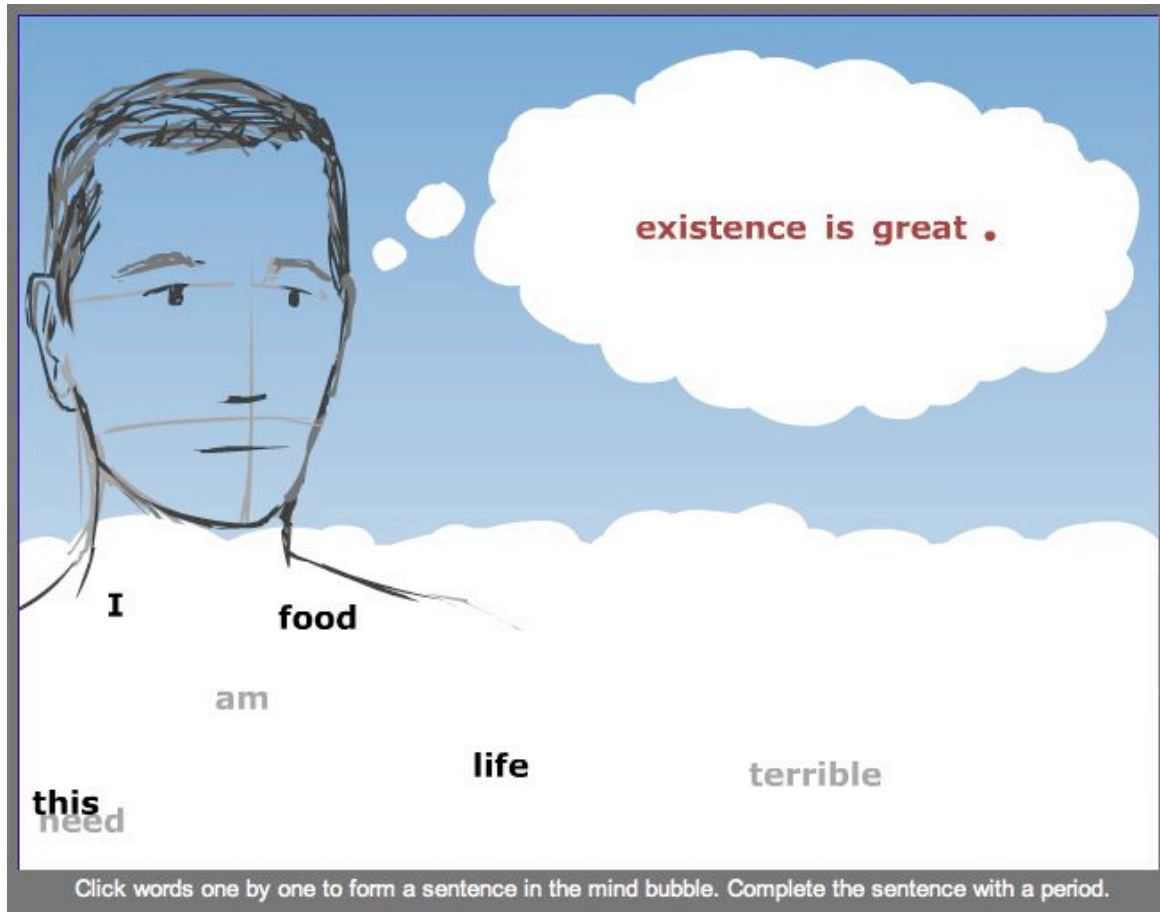


Fig. 1 Sketch prototype of active thought formation based on grammar rules

User testing of this prototype was premature in that it did not yet incorporate code to bring in new sets of words after each thought was formed, but nonetheless I took into consideration the sense of restrictiveness and repetitiveness that users felt when forming thoughts. I thought about ways in which the system's vocabulary could grow on its own. I searched for dictionaries and thesauri in web-based formats with the intention of writing code that would add new words dynamically to the database, giving them the same emotional associations as their synonyms. With this method, I could allow users to contribute thoughts in a free-writing form prior to run-time or extract content from blogs on a regular basis. In one scenario, I envisioned users clicking through a thought formation guide based on constantly

evolving collections of sentences from these forms or blogs. Testing this mechanic in Flash revealed its many flaws, the primary one being that the user often becomes locked into finishing a sentence that he does not want to finish. In addition, treating written sentences as being indicative of natural thought patterns, even coming from a so-called “stream-of-consciousness” blog, is problematic.

### 3.3.3 *Automated Thought Formation*

It became clear that getting this type of experience to be approachable, intuitive and engaging would be very difficult based on its philosophical underpinnings. When given a lot of ability to interact with a system, the user’s primary disposition is often to accomplish a goal or make something very tangible happen rather than to engage in a mental discourse with the interconnected elements of a system. Hence a pivotal step in my process was deciding to significantly alter the role of the user. I opted to change the experience from an immersive “game” in which the user is actively involved, to a largely automated spectacle. In this way, users are encouraged to take more of an observational stance rather than becoming frustrated with the results of their actions.

In the subsequent iteration of *Untitled Human Avatar Project*, the words intelligently combine into thoughts by themselves based on a set of grammar rules. Once a “full sentence” forms, it enters the avatar’s head. The user interacts simply by typing the URL of a web page, such as a blog, into a text field. The system extracts the emotional words from the web page, along with some random pronouns, auxiliaries, and universally relatable nouns such as job, family, and money, and places them into a chaotic mix where they quickly combine with other words whose part of speech is in accordance with their own. A pleasing aspect of the experience

is that neither the user nor myself has direct control over the outcome. The user intuitively adds URLs that he believes will contribute to a certain emotional response by the avatar, and derives pleasure from the confirmation of his expectations or from a surprise outcome.



Fig. 2 Prototype for a semi-automated stream of consciousness virtual experience

### 3.3.4 *Words without Thoughts*

In the prototype I am currently developing, the words take on a less overt and externalized role. I have opted to eliminate actual thought-formation from the application for several reasons. One is that the thoughts tend to produce very predictable and repetitive results on the avatar's face. The user does not get an *overall* sense of a web site's emotional effect on the avatar, and in many cases the expressions hover around neutral for long periods of time. Another reason is that designating emotions for single sentences based on keywords is problematic when context is not taken into account. (See 3.5 *Detecting Affect in Text through*

*Psycholinguistic Analysis*). Finally, developing flexible, but efficient grammar rules can be a difficult and unending task.

The current instantiation of the project is a customizable “widget” tentatively entitled *Universal Emote* — a micro-application that users can download to their desktops. It automatically connects to web pages that the users specify in the form of URLs such as [www.cnn.com](http://www.cnn.com) or the address of their own blog. The interface is designed like a TV with a remote control attached. The avatar’s face appears on a “screen” framed at the top and bottom by two bands, one showing the current URL selected, and the other functioning as a ticker of emotion words extracted from the web page. The words are color-coded to give users some confidence that actual linguistic analysis is being performed. Users can open or close the remote control tray in which they create and manage their own list of URL “channels” whose content the avatar processes emotionally, one URL at a time. Selecting “All URLs” triggers a cumulative expression based on all URLs currently in the tray. With the URL forward and back buttons located directly beneath the screen, the user can skip to the next or previous URL on the URL channel playlist. At regular intervals, the widget automatically checks for updated content on the web pages listed in the tray.

The notion that the avatar is processing the text of the web pages mentally is now much more implicit. Though it does not provide the same breadth of insights into the nature and mechanics of human thought as was originally intended, the design effectively focuses the user’s attention on a basic input and output effect and is more revealing in a way that ordinary Web users can relate to.

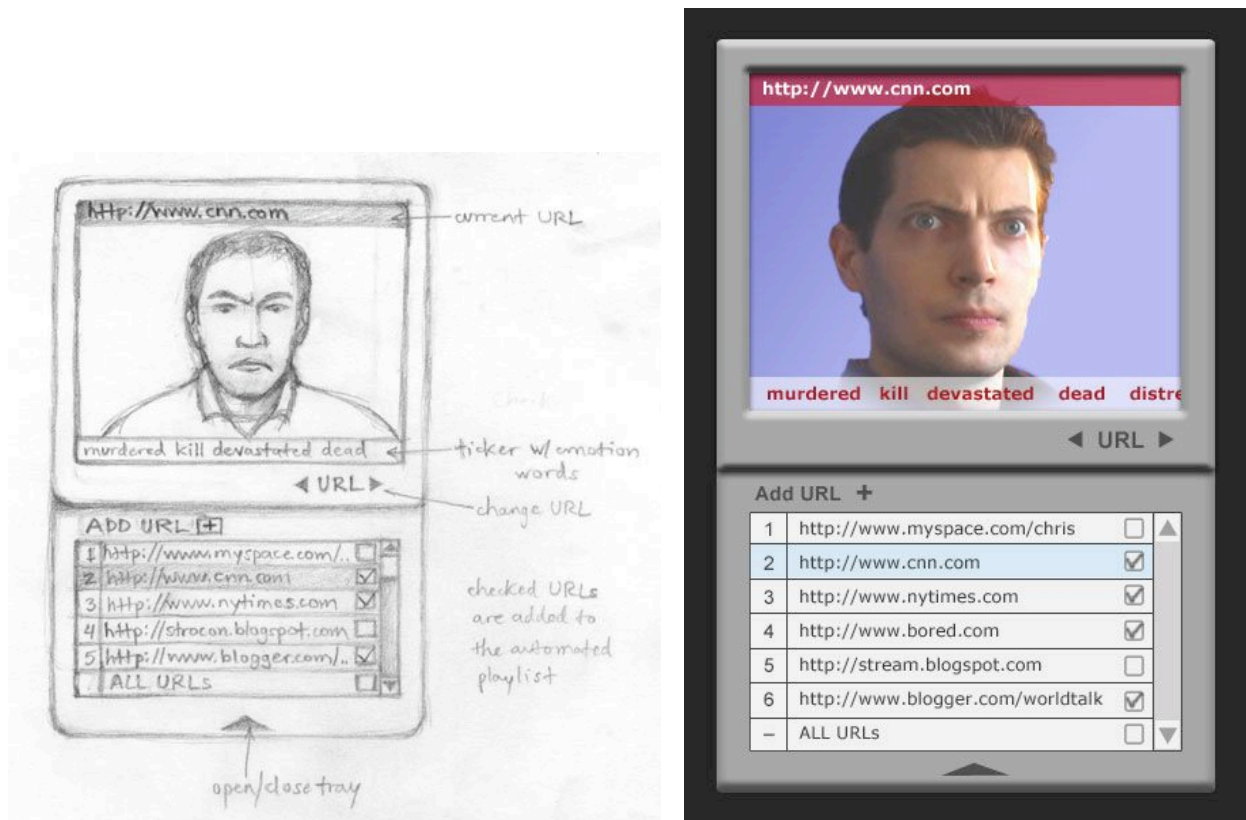


Fig. 3 A Sketch and Full Mockup of the *Universal Emote* interface

### 3.4 Detecting Affect in Text through Psycholinguistic Analysis

Through research and inquiry, I gained access to several of the psycholinguistic data resources introduced in Chap II. They are well regarded as tools for analyzing emotional affect in text and have become central to this project. L.I.W.C. and D.A.L. are both considered “word count” strategies for text analysis. (Pennebaker, 4) Word count programs are algorithmically simple and efficient in that they make psychological inferences based on cumulative keyword detection. (Pennebaker, 8) However, it is important to note that this keyword-based method is most statistically meaningful for large bodies of text as opposed to short ones. Word count strategies cannot account for irony, sarcasm, or the context in which a word is employed. (Pennebaker, 8) For instance, making the assessment that a single sentence such as “I have

beaten the odds” is inflected with anger because it contains the word “beaten” would be considered circumspect. However, if the word “beaten”, or semantically similar words, appear multiple times throughout a piece of text, it is more likely that a subtext of anger is involved.

#### 3.4.1 *L.I.W.C.*

Linguistics Information and Word Count designates 615 words as being indicative of one of five affective processes – positive feelings, optimism and energy, anxiety/ fear, anger, and sadness/ depression. In addition, it places certain words under the category of general “negative emotions”. With some guidance from my advisor, I parsed these desired word sets from individual text files into my database, mapping them to the designated emotions. (See 3.5 *The Word Database*).

#### 3.4.2 *A.N.E.W. and D.A.L.*

Ratings for “arousal” or “activation” and “pleasantness” or “pleasure” are much more commonly found in psycholinguistic studies of affect than actual emotions. However, taken in tandem, these values have been used to infer certain emotional states for some words. In D.A.L., for example, a word with a low activation rating and a low pleasantness rating is designated as being “sad”. A word with a high activation rating and a low pleasantness rating is designated as “nasty”. I am supplementing the LIWC data with A.N.E.W. and D.A.L. ratings in order to both significantly broaden the scope of words recognized by the system and to add another level of emotion data to the mix. (See Section 3.7 *Mapping Word Ratings and Designations to Facial Expressions*)

### 3.4.3 *Connecting to Web Pages*

Using the linguistic information in my database, I can have my application search through a text file, detect any and all matching words, and look up their corresponding emotional values and ratings. The idea of analyzing a web page is a good one for text analysis since web pages typically involve large quantities of text. I am using Coldfusion's <cfhttp> function to access web pages as streams of text. The only downside to the technology is that the text of some multimedia sites, such as ones created in Macromedia Flash, is inaccessible. Users will thus be limited to entering URLs of HTML-based pages.

## 3.5 **The Word Database**

Driving the interface of the project is a large word database created with Microsoft Access. In order to conduct word searches and to display words on screen, the Flash application connects to the database through Macromedia Coldfusion, a popular middleware language for Internet-based applications.

<u>word_id</u>	<u>word</u>	<u>emotion_id</u>			
			77731686	<a href="#">FRUSTRAT</a>	17
77731377	<a href="#">ACCEPT</a>	14	77731687	<a href="#">FUCK</a>	17
77731378	<a href="#">ACCEPTA</a>	14	77731688	<a href="#">FURIOUS</a>	17
77731379	<a href="#">ACCEPTED</a>	14	77731689	<a href="#">GODDAM</a>	17
77731380	<a href="#">ACCEPTING</a>	14	77731690	<a href="#">GREED</a>	17
77731381	<a href="#">ACCEPTS</a>	14	77731691	<a href="#">HARASS</a>	17
77731382	<a href="#">ACTIVE</a>	14	77731692	<a href="#">HATE</a>	17
77731383	<a href="#">ADMIR</a>	14	77731693	<a href="#">HATED</a>	17
77731384	<a href="#">ADOR</a>	14	77731694	<a href="#">HATEFUL</a>	17
77731385	<a href="#">ADVANTAGE</a>	14	77731695	<a href="#">HATES</a>	17
77731386	<a href="#">ADVENTUR</a>	14	77731696	<a href="#">HATING</a>	17
77731387	<a href="#">AFFECTION</a>	14	77731697	<a href="#">HATRED</a>	17
77731388	<a href="#">AGREE</a>	14	77731698	<a href="#">HOSTIL</a>	17
77731389	<a href="#">ALERT</a>	14	77731699	<a href="#">HUMILIAT</a>	17
77731390	<a href="#">AMAZ</a>	14	77731700	<a href="#">INSULT</a>	17
77731391	<a href="#">AMUS</a>	14	77731701	<a href="#">INTERRUP</a>	17
			77731702	<a href="#">INTIMIDAT</a>	17

Fig. 3 Sections of the “Words” table showing emotion words and word stems taken from LIWC

<u>emotion_id</u>	<u>emotion</u>
14	<a href="#">happiness</a>
15	<a href="#">sadness/ depression</a>
16	<a href="#">fear/ anxiety</a>
17	<a href="#">anger</a>

Fig. 4 The “Emotions” table showing the major emotions recognized by LIWC

### 3.6 Avatar Facial Expressions

The face of a “universal man”, an avatar, serves as a relatable and whimsical means through which the emotional content of text can be expressed. From the beginning, I have sought

to create a model of human facial expression that is both visually realistic and structurally modular, allowing for emotions to mix and gradate in unpredictable ways.

Using a video actor to represent the avatar as opposed to a 3D model or cartoon has several beneficial effects. One is that, in the context of this application, it heightens the element of the surreal. There is a certain spooky yet satisfying feeling in putting thoughts into the husk of an individual that is known to exist in the real world but whom has given up his identity in a virtual world. In addition, video can more easily convey and convince the user of effusive emotional states such as crying, laughing, and extreme rage.

### *3.6.1 A Simple Test in Flash*

Before shooting any video, I created a simple sketch rendering of the avatar's head in Flash and isolated the mouth and eyebrows of the sketch as separate movie clips so that they could be programmatically adjusted independently of one another to form a variety of facial expressions. (See *Fig. 1*) Rotating the corners of the mouth downward, for instance, produces a frown, while rotating them upward produces a smile. When tested, users enjoyed seeing the various combinations of anger, joy and sadness that could evolve from the automated shifting (in position and angle) of the eyebrows and corners of the mouth.

### *3.6.2 A Modular Approach to Interactive Video*

I devised a way to achieve a similar degree of modularity and emergence to that described above, with video-based facial expressions. The process involves videotaping three different sections of the face performing separate muscular movements and blending them together seamlessly within the application. Each facial component (eyebrows/ forehead, eyes,

and mouth/ lower face) is controlled separately by Actionscript and given separate instructions based on the particular emotion being processed. In devising the facial expressions, I referred to Paul Ekman's *Emotions Revealed*, a book that shows numerous photographic examples of universal facial expressions and what they signify. A wealth of different types of emotions can be conveyed primarily through subtle shifts in the relative positions of the eyebrows, eyes, and mouth.

### 3.6.3 *Developing the First Prototype*

The first prototype used myself as the actor. I framed my head against a hand-made blue screen and taped a smooth progression of mouth/lower face movements (smiling joyfully to neutral to lips pursed to mouth agape), a progression of eyebrow movements (brows knit to neutral to brows raised to brows angled upwards), and a simple eyelid progression (wide open to fully closed). It was important to make sure that I kept my head still for all of the takes so that all three components would stay consistently aligned with one another when assembled together in the computer. I used iMovie to capture the video footage and export the segments to Adobe After Effects. With After Effects, I keyed out the blue background in each clip, created a feathered mask to visually isolate the particular area of the face with which the clip was associated, and created a composite of the layers to tweak the blending and test out the animation. Finally, I exported each clip separately as a PNG image sequence, reducing the frame rate by an average of about 60 percent to minimize the total number of images. PNG is the most optimal image compression format that supports alpha channel blending, but the file sizes (40K each) still take a relatively large toll on Flash's processor when many images are involved. (Video formats such

as QuickTime interpolate the frames and yield a much more compact file size, but do not seem to support the necessary alpha channel blends. I plan to investigate this further).

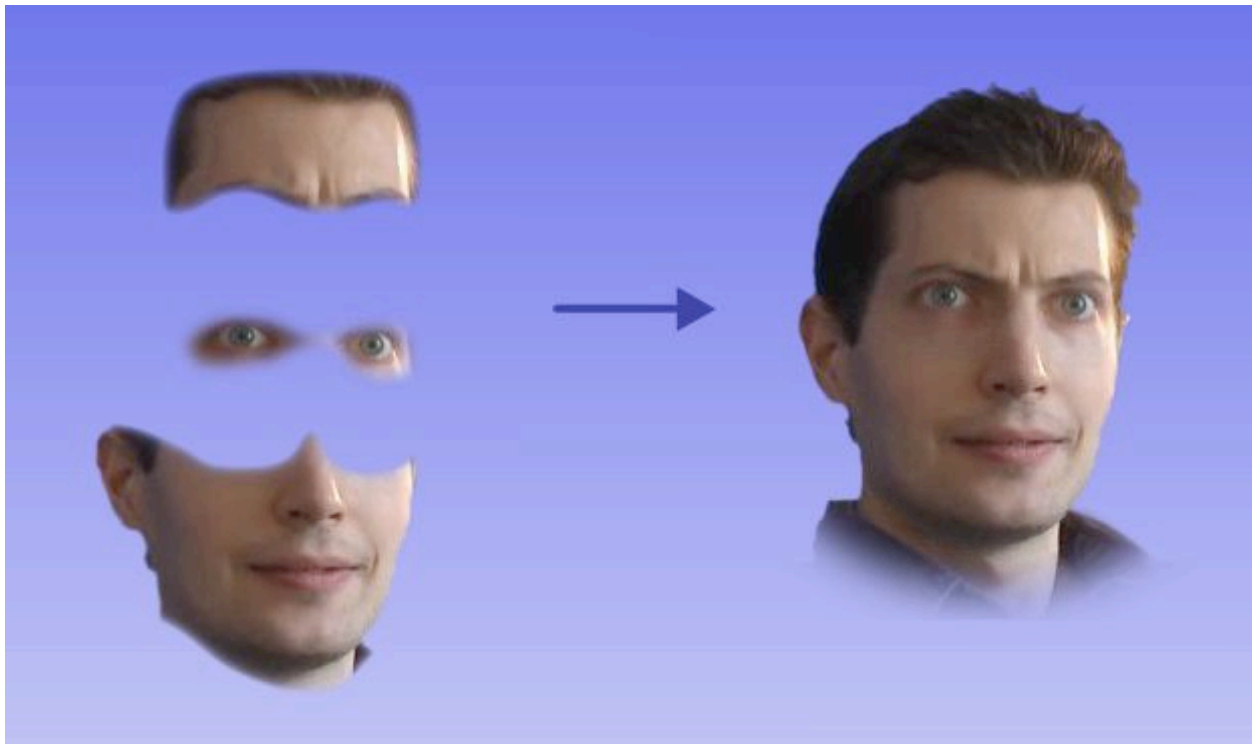


Fig. 5 My own form of video processing based on combining three separate facial components to form a variety of expressions

In Flash, I imported the image sequences into separate timelines and wrote a script to move backward or forward along each one, frame by frame, a particular number of frames at a time, depending on the emotion and intensity value of the particular word being processed. I developed a simple prototype in which a variety of emotion words stream behind the avatar's head and the user can click on a word to make it fly into the forehead and produce a reaction on the face. Each emotion cumulatively builds upon the previous one. An arsenal of angry words causes the avatar's brows to become increasingly knit up to a certain limit. A combination of happy words and fear-inducing words produces a curious expression in which the avatar appears to be masking his fear with a smile.



Fig. 6 A prototype test for facial expressions

Despite some of its technical successes, the test revealed that more work needed to be done in enhancing the variability and range of the facial expressions. It seemed that it would be necessary to incorporate a certain amount of “branching” to new video clips within each facial component.

#### 3.6.4 *New Prototype*

In a second prototype, I enlisted an older male actor to play the role of the “universal man”, looking to remedy some of the calculated and robotic feelings engendered by my own performance and to increase the overall entertainment value. However, trying to broaden the emotional breadth and allow for spontaneity in the performance is challenging in a case like this one, in which maintaining a high level of measured control is necessary to ensure continuity

between clips and to make the programming end manageable. Finding the proper balance is the key.

### 3.7 Mapping Word Ratings and Designations to Facial Expressions

The application maps emotions, arousal, and valence values to facial expressions according to the designations below. The number of frames within the sequence along which to advance or retreat the playhead depends on the valence value of the word.

Joy/ Happiness ->	<u>mouth</u> : move toward full smile or laugh/ back toward neutral
Sadness/ Depression ->	<u>eyebrows</u> : move toward lowered/ back toward neutral <u>mouth</u> : move toward full frown or pout/ back toward neutral
Fear/ Anxiety ->	<u>eyebrows</u> : move toward fully raised or angled up/ back toward neutral <u>mouth</u> : move toward mouth fully agape/ back toward neutral
Anger ->	<u>eyebrows</u> : move toward fully knit/ back toward neutral <u>mouth</u> : move toward full frown or teeth clenched/ back toward neutral
High Arousal ->	<u>eyes</u> : move toward fully open
Low Arousal ->	<u>eyes</u> : move toward almost closed

*When a specific emotion is not designated:*

High (Positive) Valence ->	<u>mouth</u> : move toward full smile/ back toward neutral
Low (Negative) Valence ->	<u>mouth</u> : move toward extreme of current negative emotion clip/ back toward neutral

Works Cited

Pennebaker, James W., Matthias R. Mehl, and Kate G. Niederhoffer. "Psychological Aspects of Natural Language Use: Our Words, Our Selves." Department of Psychology, University of Texas at Austin, 2003.