

EMOTIONAL SUBTITLES: A SYSTEM AND POTENTIAL APPLICATIONS FOR DEAF AND HEARING IMPAIRED PEOPLE

James Ohene-Djan¹, Jenny Wright¹ and Kirsty Combie-Smith²

¹Goldsmiths College, University of London,
New Cross, London, SE14 6NW
Phone: +44 207 919 7462, Email: j.djan@gold.ac.uk

²Deafax, UK.1 Earleygate
University of Reading, Whiteknights Road
Reading RG6 6AT (UK)
Phone: +44 [0] 870 770 2463 Email: kirsty@deafax.org

Abstract: Many Deaf and Hearing Impaired people use subtitles to gain access to audio content on television and film presentations. Although subtitles tell the viewer what is being said they fail to communicate how it is being said. This “emotional gap” experienced by viewer’s highlights a significant drawback to current subtitling especially when used for learning by the Deaf. In this paper we introduce a system that demonstrates the presentation of subtitles that depict the emotions behind the words used on screen. The system also provides viewers with the ability to personalize and adapt their interaction with subtitles, so as to assist them in their learning. Using this system we hope to conduct a series of surveys looking at how people receive and use subtitles. In conducting this research we aim to gain a comprehensive understanding of the issues associated with emotional subtitling and to provide guidance for future producers of subtitled materials.

Keywords: Emotional subtitles, Design principles, Deaf technologies, Hearing Impaired technologies.

1. Introduction

Subtitling is a popular communications technology with over 7.5 million users in the UK (Ofcom, 2006). However, a significant limitation of the present generation of subtitling techniques and technologies is their failure to express the emotional nuances of dialogue such as intonation and volume. Subtitles have been criticised for their lack of emotional content and inability to communicate the subtext of a rich dialogue. At present subtitles are limited to telling the viewer what is being said but not how it is being said. Often those that rely upon subtitles, such as the Deaf and Hard of Hearing, draw attention to the “emotional gap” that is generated, and the important emotional information that is lost. This paper presents a system for creating emotional subtitling. Section two provides background to our research. In section three we present our system and provide examples of emotionally subtitled content. In section four we place our research in context with that of related work. Finally in section five we draw some conclusions.

2. Background

The use of subtitling as an assistive technology to represent audio dialogue is well known and understood. (RNID, 2007, Friscolanti, 2004), In the UK between 80% and 90% of television programmes currently have subtitles (BBC, 2006a). Generally, the approach taken is to transcribe the audio dialogue into a succession of phrases that are shown more or less at the same time as the character vocalises them on screen (Robson, 2004). Such a series of phrases may be supplemented with visual cues that describe the locations of characters on screen. Visual cues are also used when multiple characters are speaking at the same time to indicate which character says which phrase. With the mainstream emergence of digital television, visual and hearing-impaired users have benefited from facilities to perform simple manipulations on subtitles to assist their viewing (BBC, 2006b). However, at present, such facilities are limited to allowing users to change the size of subtitles and their colour.

For the Deaf and Hearing Impaired viewers, subtitles provide only a limited communication representation. Although viewers can read the words that are said, they cannot determine how something has been said. For example, a character could say the phrase, "I will be there in a minute" in a menacing way or in a joyful way; yet in either case the subtitle is exactly the same. Similarly, whether the character says the line very quietly, or shouts it at the top of their voice, generally there is no difference in the way the subtitled phrase is displayed. The limitations of subtitles are further exaggerated when characters cannot be seen on screen, during action sequences, and when multiple characters are speaking at the same time.

In context of education and learning, Deaf and Hearing Impaired students use subtitles to read dialogue and identify characters. However they are disadvantaged when viewing teaching material such as filmed theatrical productions where, often, how a character delivers a line can be as important as the line itself (Kyle, 1992). In several interviews we conducted, Deaf and Hearing Impaired people indicated that what was lacking from subtitles was any degree of emotion. They found that, although they could follow the dialogue, it felt shallow, somehow 2- dimensional rather than 3-dimensional.

3. Emotional Subtitling

To create subtitles that depict emotions and tone, we propose a model in which subtitles are viewed as specifications of when to display content and how to render and format it. The model consists of a formal specification of subtitled presentations and a specification of how to annotate these specifications with user defined emotional semantics.

Conceptually, a subtitle presentation may be thought of as a sequence of phrases, which consist of one or more words, or chunks of text. Each phrase will have at least a Start Time Point (STP), more generally the point in time at which a phrase occurs within the film or television presentation, and an End Time Point (ETP), the point at which it is removed from screen. Each chunk of text may be given a set of attribute values which will determine how that particular chunk is displayed on the screen. A simple formal language, defined by a context-free grammar, is used to define the syntax of subtitle presentations. Figure 1 defines the grammar of the formal syntax of subtitle presentations using the classical Enhanced Backus-Naur form (EBNF).

```

Subtitle Pres ::= <SubPres> phrase* </SubPres>
Phrase ::= <Phrase> STP, chunk*, ETP </Phrase>
Chunk ::= <Chunk> S-content, Annotation </Chunk>
S-content ::= String
Annotation ::= <Ann> note* </Ann>
note ::= <attrib> attrib-assign</attrib>
attrib-assign ::= attribute:= attribute_value
STP ::= <STP> T-string </STP>
ETP ::= <ETP> T-string </ETP>
T-string ::= Formatted time string

```

Figure 1. Grammar for SPSs

In Figure 1 STP and ETP are defined as strings whose format is assumed to be that of a date time in a commonly understood implementation format. An S-String, which is the actual text of the subtitle, is defined to be a string implemented in a modern programming language. The * denotes a sequence. Annotations are used to give user defined semantics to a phrase. These semantics describe the emotional characteristics associated with a phrase or chunk of text. Possible attributes for a phrase are: volume, colour, text size and intonation. Possible values for intonation are: happily, sadly, sarcastically, excitedly, comically, fearfully, pleadingly, questioningly, authoritatively and angrily. The model we propose does not preclude the introduction of new emotional characteristics as the implementer sees fit. A more complete discussion on this formalisation can be found in (Ohene-Djan, and Rachel, 2006).

3.1 Creating Emotional Subtitles

The first step in creating emotional subtitles is to create the Flash movie that would control the subtitles and the associated emotion formatting. The Flash movie file consists of a media playback container for displaying the video file and a dynamic text field used to display subtitle text and the formatting instructions as text. Both text files are stored in separate XML files and retrieved dynamically at run-time. When a Flash movie is exported, the output is a .swf movie which can then be embedded into a web site. This particular file combines inputs from two both types of file; the film clip itself is .flv file, while the captions and formatting are stored as XML files. Figure 2 is an illustration of this, showing how these components come together to make the .swf file to be used in the web site.

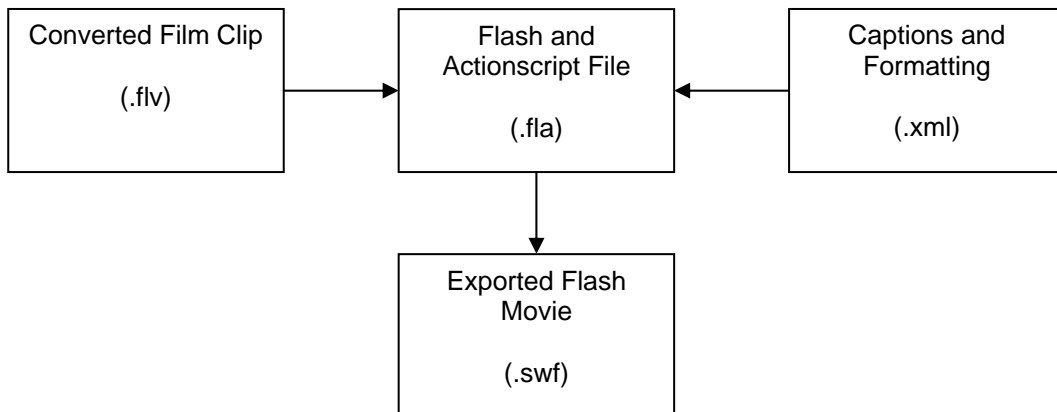


Figure 2: Diagram Showing the Components of a Flash movie swf File

To facilitate the easy creation of emotional subtitled videos an emotional subtitle editor was developed. This editor allows users to input subtitles and format them according to predefined schemes; changing the font only, changing the colour only, and changing the font and the colour. Users can also change the size of text according to volume. The form within the editor used for inputting subtitling schemes is shown in Figure 3.

Figure 3: Form for Inputting Subtitles Using Existing Schemes

3.2 Examples of emotional subtitling

Figure 4 and 5 shows examples of emotional subtitling. Figure 4 shows an extract from the Monty Pythons Holy Grail Film and Figure 5 shows an extract from BBC television programme, the Tweenies. Without changing any of the text, subtitles have been transformed into emotional subtitles using the following simple ideas. Each character has a different colour face. These colours are used to represent the speech of each character. In addition, lines which are delivered with a particular emotion are shown in a font that represents that emotion.



Figure 4 Pythons Holy Grail Film with Emotional Subtitles



Figure 5 BBC's Tweenies television programme With Emotional Subtitles

4. Related work

Researchers at Cambridge University, in conjunction with Red Green and Blue Co - a multimedia company – have produced a DVD-ROM entitled *Mind Reading: the Interactive Guide to Emotions*, the world's first encyclopaedia of emotions. (S Baron-Cohen, J Hill, O Golan, S Wheelwright, 2006). This research defined 412 emotions, which could be divided into 24 groups. The work reported here used these emotional groups to implement the emotional subtitling system.

In their paper *Emotive Captioning and access to Television*, (Fels et al, 2005), investigated the use of graphical captions to display emotions as contrasted with the work reported here which used colours, fonts and text sizes (see Figure 6). The results of which were met with a mixed reaction by the deaf and hearing-impaired testers of the system. Initially, neither groups had a positive reaction to the colours, but on second viewing the hearing-impaired users, on the whole, thought the colours added value to the video. Deaf users however, did not like the captioning shown to them as an alternative; even though they stated earlier that conventional captions omit emotive information. They did not find use of colour acceptable unless it could be part of the text.

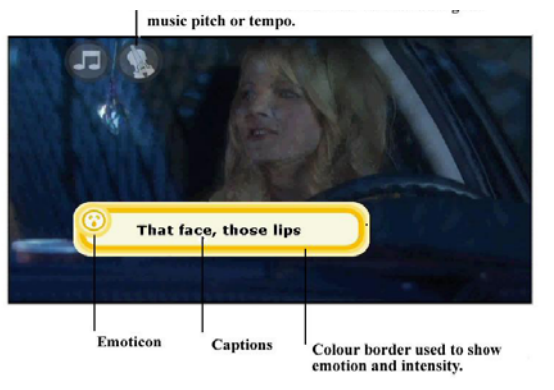


Figure 6 Emotive Captioning

Raisa Rashid, Jonathon Aitken and Deborah I. Fels, in their paper *Expressing Emotions Using Animated Text Captions* (Rashid, R. Aitken J. & Fels, D.I., 2006), investigated the use of kinetic typography to express emotion in subtitles. Kinetic typography is the concept of adding animation to text to reflect tone of voice and emotion. For example, one emotion they portrayed was intense fear, through text that increased rapidly in size and “vibrated” quickly. Their prototype has yet to be evaluated with users, and the decision to use certain animations to represent certain emotions was purely subjective. The paper was written to show what could be done, and they raise the question of how much animation is enough to improve understanding without distracting the user. The answer to this, I feel, is that any animation will distract the user, and from the examples shown in the paper, it does not seem to enhance understanding.

5. Conclusions

Subtitling is a well-established, popular form of providing audio description on television and DVD, aimed at aiding the deaf and hearing-impaired community. There are, however, limitations of the subtitling systems used at present; there is no way of expressing emotional nuances, and as such those users who are deaf or hearing-impaired cannot fully understand the media they are watching. There has been some research into using subtitles as a learning technology for teaching people to read, but little into how they can be used to aid other people in society. This paper presents an alternative to the static subtitles that exist at present; one that displays basic emotions in such a way that it will be universally suitable, but particularly beneficial to the Deaf and hearing-impaired communities.

If future work we will survey the fonts and colours used in emotional subtitling to find the most effective combinations and test them with a much wider audience, particularly with deaf and hearing-impaired users. We hope that the work reported here provides valuable knowledge to the wider academic community on current limitations of subtitling, ways of representing emotions through subtitling and potential future developments, such as personalised subtitling.

References

- Baron-Cohen, S., J. Hill, O. Golan and S. Wheelwright (2006). *Electronic Emotions Encyclopaedia*, Available from <http://www.autismtoday.com/articles/emotions.htm>
- BBC (2006a). *Policies: Subtitles and Audio Description on TV*, Available from: <http://www.bbc.co.uk/info/policies/subtitles.shtml>
- BBC (2006b). *The Joy of Subtitles*, Available from: <http://news.bbc.co.uk/1/hi/magazine/4862652.stm>
- Fels et al. (2005). *Emotive Captioning and Access to Televisions*. Available from: http://www.theclt.org/Papers_New/burnt_toast_omaha.pdf.
- Friscolanti, M. (2004). Deaf demand subtitles at all movie theatres, *National Post, Canada*, Oct 27, 2004 http://www.deaftoday.com/v3/archives/2004/10/deaf_demand_sub.html.
- Ivarsson, J. and M. Carroll (1998), *Subtitling*, TransEdit, Simrishamn, Sweden.
- Kyle, J. (1992). *Switched On: Deaf People's Views on Television Subtitling*, Centre for Deaf Studies, Univ. of Bristol, 235 p. (Report for the ITC and the BBC).
- Ofcom (2006). <http://www.ofcom.org.uk>.
- Ohene-Djan, J. and R. Shipsey (2006), E-subtitles: emotional subtitles as a technology to assist the deaf and hearing-impaired when learning from television and film, *6th IEEE International Conference on Advanced Learning Technologies (ICALT)*. Kerkrade, The Netherlands: IEEE Computer Society, Los Alamitos, CA, USA, July 9-11 2006, pp. 464–466, ISBN 0-7695-2632-2.
- Rashid, R., J. Aitken and D.I. Fels, (2006). Expressing emotions using animated text captions, *ICCHP 2006*, Linz.
- RNID (2007) Royal National Institute for the Deaf (RNID), *Subtitling*, http://www.rnid.org.uk/howwehelp/research_and_technology/communication_and_broadcasting/subtitling/.
- Robson, G.D. (2004). *The Closed Captioning Handbook*. ISBN: 0240805615, Focal Press.