Home > Deaf people's experiences, attitudes and requirements of contextual subtitles: A two-country survey

Deaf people's experiences, attitudes and requirements of contextual subtitles: A two-country survey

Marion Hersh [1]

University of Glasgow, Scotland

[3]

Abstract

Subtitles are text versions of the speech content of television and other audiovisual media. The paper discusses the results of a survey of deaf people in Poland and the UK on their experiences of, attitudes to and requirements for subtitles, including for the representation of emotions and contextual features. The results demonstrated the importance of subtitles, and that they considerably improved comprehension for the great majority of respondents. The overwhelming majority preferred verbatim subtitles, but attitudes to other suggested features were very varied, indicating the need for subtitle personalisation. This has not been considered previously and would now be feasible due to advances in technology.

Introduction

This paper presents the results of a survey of deaf people [5] in the UK and Poland about their experiences of, attitudes to and requirements for subtitles, including for the representation of emotions and contextual features. The methodology, results and conclusions of this survey are presented in Sections 2, 3 and 4 respectively. It is introduced by a brief overview of the literature in the remainder of this section.

Subtitles are verbatim or edited (and sometimes simplified) text versions of the speech content of a television programme or other audiovisual presentation. Closed or soft subtitles (called captions in the USA) are used to make audiovisual media accessible to deaf people, whereas the subtitles used to make foreign language materials accessible to both hearing and deaf people are generally open. Closed subtitles are presented as separate instructions from the programme and can therefore be personalised or omitted, whereas open or hard subtitles are an integral part of the video frames and cannot be turned off or personalised (Anon 2012a [6]). Closed subtitles can be produced either online or offline, whereas open subtitles are always produced offline.

In offline subtitling the subtitle file is prepared in advance and automatically cued into the programme using a timecode in the programme master tape or the subtitles may be encoded in the video signal. This allows the text to be checked and errors to be corrected and careful synchronisation of the text and speech. Online subtitles are used when there is insufficient time to prepare a subtitle file. They include realtime or live subtitling, in which a realtime stenographer uses a computer to transcribe stenographic input for nearly immediate presentation, and pre-prepared subtitles, which are prepared in advance and input using live manual cueing when the programme is transmitted. Speech recognition technology with revoicing (as speech recognition software is designed for use with a single voice) is frequently used in live captioning in the UK (Anon 2012a [6]).

The high rate of subtitling by the two main UK broadcasting companies at over 97.5% and 88% by the BBC and IT respectively (EFHOH 2011 [7]) and the over 1.5 million deaf subtitle users in the UK (Ofcom 2007 [8]) illustrate the importance of subtitling for deaf people. A small-scale survey of older deaf people (59-82 years old) found that the use of closed subtitles significantly increased their understanding of three different types of television programmes (Gordon-Salant & Callahan 2009 [9]).

Early subtitles in Europe and the USA were transmitted at 120-125 words per minute, two thirds the rate of speech, and in simplified language, based on assumptions about the needs of prelingually deaf people (Jensema & Burch 1999 [10]), (Schilperood et al. 2005 [11]). However, great care is required to ensure that editing and simplification do not change the meaning and facilitate rather than reduce understanding (Power & Leigh 2000 [12]), (Schilperood et al. 2005 [11]). Feedback from deaf viewers in both Europe and the USA has indicated a preference for verbatim subtitling (Jensema & Burch 1999 [10]), (Schilperoord et al. 2005 [11]). In many European countries subtitles comprise two lines of text (64 characters) which are displayed for a maximum of six seconds, giving 120 words per minute. Studies of subtitle reading rates with relatively large numbers of deaf and hearing people have obtained a most 'comfortable' reading speed of about 145 words per minute (Jensema et al. 1996 [13]), (Jensema 1998 [14]) and recommendations of not more than 180 words a minute (Ofcom 2005 [15]) and found (Braverman & Herzog 1980 [16]), (Tyler et al. 2009 [17]) no benefits in reducing subtitling speeds for children to 90 words a minute, despite recommendations of 60 words (Baker 1985 [18]). Many countries have converted from analogue to digital television, are in the process of doing so or plan to do so in the next ten years (Anon 2012b [19]).

Deaf people have criticised existing subtitling systems, largely on account of poor legibility and the lack of contextual features, including emotion in dialogue and background music, music tempo, mode or depth, non-speech items, speaker identification and the timing of jokes and puns (Bok?an-Cullen 2012 [20]), (Fourney & Fels 2008 [21]), (Lee et al. 2007 [22]), and appreciated the use of cues to indicate sarcasm, pauses and background noises (Ofcom 2005 [15]). The factors that reduce legibility include blur, inappropriate or poor formatting, poor colour contrast, spelling mistakes, poor synchronisation of the speech and text and too high a transmission rate (Bok?an-Cullen 2012 [20]), (ITC 1999 [23]), (Karamitroglou 1998 [24]), (Ofcom 2005 [15]), (Thorn & Thorn 1996 [25]). Good synchronisation is important as many deaf people lip-read and the synchronised subtitles and speech work together to aid comprehension, rather than the subtitles being the only source of information. High transmission rates may be a particular problem for older deaf people with reduced rates of information processing and/or visual impairments (Thorn & Thorn 1996 [25]).

Representing Emotions and Contextual Features

Existing subtitle systems increasingly provide information about background sounds, but rarely full contextual information, for instance that there is background music, but not its mood and other characteristics (Karamitroglou 2002 [24]), the texts of songs or the emotion in speech (Hersh et al. 2010 [26]), (Lee et al. 2007 [22]). Sound is important in conveying emotion in films, leading to the possibility of the misinterpretation of subtitles when this information is absent (Shilling et al. 2002 [27]). Most of the limited research on representing emotions in subtitles has been carried out by two very loosely defined groups, based in Canada (Fels et al. 2005 [28]), (Lee et al. 2002 [29]), (Lee et al. 2007 [22]), (Rashid 2008 [30]), (Rashid et al. 2006 [31]), (Rashid et al. 2008 [32]), (Silverman and Fels 2002 [33]), (Vy et al. 2008 [34]) and the UK (Hersh et al. 2009 [35]), (Hersh et al. 2010 [26]), (Ohene-Djan & Shipsey 2006 [36]), (Ohene-Djan et al. 2007 [37]).

Several approaches have been implemented on a trial basis in short audiovisual sequences to obtain feedback from deaf and sometimes also hearing people. Ohene-Djan et al. (2007 [37]) identified speakers by the use of different colours and represented emotions by different types of fonts and loudness by font size. Rashid and colleagues (Rashid 2008 [30]), (Rashid et al. 2006 [31]), (Rashid et al. 2008 [32]) have developed a framework relating seven animated text properties, including text size, duration and 'shaking', to four basic emotions, though others could be added. Larger text, faster movement and faster onset were used to indicate a stronger effect. Small-scale tests found that deaf and hearing people preferred the text to be in the standard location at the bottom of the screen and the version without strong shaking movements. Correct recognition of emotions was slightly, but not statistically significantly, greater for standard than animated subtitles (Rashid 2008 [30]).

Comic book conventions, including speech bubbles, colour, text styles, animation and icons, have been used to represent eight basic emotions and various speech features and identify speakers (Silverman and Fels 2002 [33]), (Fels et al. 2001 [38]). Positive feedback was obtained from a small number of viewers watching a short video clip, though several noted that they associated these conventions with children. A related approach (Fels et al. 2005 [28]), (Lee et al. 2007 [22]) used graphics and a coloured border round subtitle phrases to represent eight emotions and their intensity, and icons for music and sound effects. Small numbers of hard of hearing participants watching a short video clip liked the coloured captions, but Deaf participants did not, though age may also have affected preferences. Both groups did not like the moving captions, found they sometimes interfered with actors' faces and preferred knowing where the subtitles were than having to look for them. Other suggestions include the use of atmospheric pictures and photos of the faces of Deaf signers (Hersh et al. 2009 [35], 2010 [26]).

A very different approach involves vibro-tactile stimuli conveyed via a footrest with two vibrators and two variable speed fans for each foot (Degan & Fels 2001 [39]). Foot stimulation was chosen to keep the hands free, but could raise social acceptability issues. Tactile pattern, frequency and signal strength were used to represent four emotions, with parameter values based on guestimates. Small-scale tests found that participants had difficulty in recognising the different emotions.

Subtitling tools have been developed to support the use of additional features to indicate emotions and the context. The emotional subtitle editor (Ohene-Djan et al. 2007 [37]) allows the subtitle font, colour and/or text size to be formatted according to predefined schemes. The CapScribe tool (Boyce al. 2012 [40]) provides support for adding text styles and a second video window, which could provide sign language interpretation, additional graphics or animated features. The rendering engine tool (Fels et al. 2005 [28]) automatically creates graphical pop-on cartons using pre-designed image files associated with the different emotions. EmACT (Vy et al. 2008 [34]) can add text animation, style, formatting, colour and location.

Methodology

A survey of deaf people and their experiences, attitudes to and requirements of subtitles was carried out using a six-section questionnaire with a mixture of quantitative and qualitative questions. Both English and Polish versions of the questionnaire were produced.

The following information was sought:

- **Section A**: gender, age, education, employment, language (sign or spoken), ease of communication and reading.
- **Section B**: frequency of viewing and the availability and quality of television, DVD, video and cinema subtitles.
- **Section C**: ease of reading and understanding subtitles, whether they improve understanding and are displayed for long enough, and preferences for verbatim or edited subtitles.
- **Section D**: interest in additional information, including about speakers' emotions, features of speech, the text of songs, sounds and the atmosphere.
- **Section E**: evaluation of specific proposals for representing this additional information.
- **Section F**: the use of subtitles in educational programmes and lectures.

Methodologies for surveying deaf and other groups of disabled people are incomplete and there is not yet an accepted best procedure (Hersh 2010 [41]), (Hersh 2011 [42]). The smaller numbers and the fact that readily accessible public lists very rightly do not indicate disability status make this more complicated than surveying the general population. The questionnaire was sent to a large number of organisations and personal contacts, who circulated the questionnaire to other individuals and organisations, and posted on my web site. Information about the questionnaire was posted on several lists. When quoting comments from Polish respondents, I have tried to keep the translation close to the style of the original. Statistical significance at the 0.05 level was determined using Kirkman?s (1996 [43]) software for a contingency table *?2 test with five degrees of freedom or four in the case of zeros across a row.*

Results

The survey received 83 useable responses, though not all respondents answered all the questions. The calculated percentages for each response are based on the number of respondents for the question not the total. The respondents were approximately gender balanced, fairly evenly split between Poland and the UK and had a reasonable age distribution though the 26-40 age group was underrepresented (see Table 1). Both 'signers' and 'speakers' and users of both sign and spoken language were represented, though the percentage of signers was higher than in the deaf population. There was an approximately even division between respondents with secondary, vocational or further, and higher education and a few respondents with only primary education. The largest occupational group was retired (typical of the increase in deafness with age), followed by students, with just under a quarter employed full or part time and the remainder unemployed, on a disability pension or looking for a job.

It is difficult to totally avoid respondent bias in this type of sampling. However respondents were sufficiently varied on the main demographic variables to be representative of the various perspectives in the deaf community and there was no obvious direction of bias, making it unlikely there would be significant bias.

Gender (%)		Age Profile (%)					Country (%)		Language (%)		
Female	Male	16-25	26-40	41-60	61-70	71+	Poland	UK	Oral	Sign	Both
49.4	50.6	31.3	8.4	22.9	16.9	20.5	51.8	48.2	58.5	34.1	7.3

Table 1 - Respondent characteristics.

Viewing Behaviour and Interest in Subtitles

Watching television and DVDs or videos were important activities for the respondents with just over half of them watching TV for several hours a day and four fifths at least once a week, just over a quarter watching DVD or videos several times a week and 40% at least once a month. Respondents visited the cinema relatively infrequently, with only just over a third going at least once a month. One respondent who did not go to the cinema considered it 'too loud' and this exacerbated their tinnitus.

Although not all the respondents used subtitles, they were a (very) important component of viewing for the majority of them (see Table 2), with over a third only watching subtitled TV programmes, just over half only or mainly watching subtitled programmes and another 28.6% preferring subtitled programmes, but willing to watch others. The figures were similar for DVDs and videos. When subtitles were available only for some screenings of a good cinema film, nearly a quarter would only watch with subtitles and over another third would watch the subtitled version if the time was convenient. Nearly a fifth would not see the film and could be considered to have been prevented from doing so by the lack of subtitles. In all three cases only about a fifth considered the presence of subtitles unimportant.

Watching with subtitles	TV	DVD or Video	Limited cinema subtitling	%
Only subtitled programmes	35.1%	32.0%	Go when subtitles on	24.6%
Mainly subtitled programmes	15.6%	24.0%	See with subtitles if time convenient	36.9%
Prefer subtitles, will watch others	28.6%	26.0%	Watch unsubtitled film	20.0%
Unimportant whether subtitles	20.8%	18.0%	Not see film	18.5%
Response rate	92.8%	60.2%	Response rate	78.3%

Table 2 - Watching with subtitles.

Despite the commitment of the major UK TV broadcasting companies to subtitling, over a quarter and over 60% respectively of the respondents considered that a lot and at least some interesting programmes lacked subtitles (see table 3), though UK respondents noted the 'good coverage'. The figures for DVDs were similar, but slightly better. In both cases only about 15% of respondents considered that no interesting programmes lacked subtitles.

Programmes lacking subtitles	TV	DVD or Video
A lot	26.3%	20.5%
Some	34.2%	31.5%
A few	21.1%	26.0%
None	14.5%	15.1%
Unsure	3.9%	6.8%
Response rate	91.6%	88.0%

Table 3 - Programmes which lack subtitles.

Of the respondents, 39 commented on the quality of TV, 30 on DVD subtitles and 31 on cinema subtitles. In line with the preference for offline (pre-recorded) TV subtitles, most comments considered them 'generally good' or even 'usually excellent', 'unless transmission – delayed' and then they often bear little relationship to what is on screen. That is almost worse than 'useless'. Worryingly, the quality of live subtitles was considered 'not good enough', 'absolutely appalling', and 'too inaccurate and slow to enjoy watching' with 'too many misspelt words'. The introduction of revoicing to correct mistakes was considered to have made the situation worse. This may be due to analogous contextual features to those that make speech recognition less accurate in a class than an office (Wald 2006 [44]) and requires further investigation.

DVD subtitles were 'generally quite good' or 'excellent', though they were not 'descriptive like on TV', 'missed things such as door bangs outside – that help understand the ambience' and 'very often the extras are not subtitled'. 'DVDs without subtitles are generally old films often classics', but 'programmes broadcast with subtitles do not always have them on the DVD'. Many respondents were only interested in DVDs with subtitles and particularly irritated by 'the fact that many online sellers neglect to tell you whether or not the DVD has subtitles'. Several Polish respondents appreciated the ability to 'stop' the DVD to finish reading the subtitle if necessary and considered cinema subtitles 'too fast'. However, the main issue for cinema subtitles was availability rather than quality. One respondent noted that 'the subtitles do not work – five out of the six last times we have been' and another that cinema staff have to be 'remind[ed] – to turn them on'. Respondents also noted legibility problems, including 'light colour text – lost on a light coloured background', 'sometimes a line is missing', or the subtitles being 'out of tune with the dialogue?' i.e. not properly synchronised.

Subtitles made a significant difference to respondents' experiences of watching TV, DVDs and cinema films (Table 4), with nearly 60% and over 80% respectively understanding programmes a lot better and better with subtitles. This is paralleled by the just under 70% of respondents who were able to understand all or most of the text and the over 70% who found them (very) easy to read and considered that they were displayed for sufficient time to read and understand most of the text. However, about 10% of respondents found them (very) difficult to read and displayed for insufficient time, paralleling the 11% who found reading (very) difficult. Over 20% only understood about half and 5% very little of the text.

Understanding better with subtitles		Ease of reading subtitles	How much of subtitles understand			
A lot better	58.4%		Very easy	41.3%	Whole text	41.8%
Better	23.4%		Easy	31.3%	Most of text	26.6%
About the same	15.6%		Neither easy nor difficult	17.5%	About half	21.5%
Not sure	2.6%		Difficult	7.5%	Very little	5.1%
			Very difficult	2.5%	Depends	3.8%
					Unsure	1.3%
Response rate	92.8%		Response rate	96.4%	Response rate	95.2%

Table 4 - Understanding and ease of reading of subtitles.

The majority of respondents wanted verbatim subtitles (see Table 5), with nearly 60% always wanting verbatim subtitles and nearly 80% wanting verbatim subtitles either always or unless they would be very complicated. They considered it 'very patronising' when 'things are made simpler', wanted to be on an 'equal footing [with] hearing people' or 'to get to know the character' and what 'the use of the words show' about them. The problems lip-reading if 'what I am reading does not appear to match up with what they are saying' were also noted. Some respondents were 'happy to have a shortened version – less tiring' if what was being said was not 'very important', but wanted to 'see it all' if 'important to the subject matter'. Respondents who wanted a shorter version recognised that they were 'not able to understand the text' or needed 'time to both read the subtitles and see what was on the screen'. The approximately 9% wanting a shorter version was compatible with the 10% finding subtitles (very) difficult to read and the time too short. However, 30% wanted longer to read the subtitles, which is closer to the 25% who found the time sufficient to understand very little or at most about half of the text.

Sufficiency of available time		Preference for verbatim or edited subtitles		Want additional time to read subtitle text	
Read and understand all text	50.6%	Verbatim always	57.9%	Yes, always	30.8%
Read most text, understand sense	20.8%	Verbatim, unless very complicated	21.1%	Yes, unless miss part of speech	12.8%
Sometimes understand sense	18.2%	Shorter version	9.2%	No	37.2%
Often insufficient and not understand	10.4%	Unsure	11.8%	Unsure	19.2%
Response rate	92.8%	Response rate	91.6%	Response rate	94.0%

Table 5 - Sufficiency of time to read subtitles and preference for verbatim or edited version.

Interest in Contextual Features and Emotions and their Representation

Respondents were considerably more interested in subtitles showing information about the main and other characters' feelings than the atmosphere or the loudness or speed of their speech, and these were the only features a majority (over 60% and over 50% respectively) of the respondents considered information about (very) important (table 6). Slightly under half and about 40% considered information about the atmosphere and loudness of the main characters respectively (very) important. The 26 comments about information about emotions and 22 about the atmosphere expressed various perspectives.

Some respondents were interested in this information, as 'Normal heard speech is full of emotion, it conveys feelings that are lost when presented with the written word' and 'Tone of voice conveys a lot of information. A lot of jokes rely on other characters' reactions'.

Others felt it was already available to them, as 'emotions can usually be judged from the appearance of the character, comment in subtitles not really needed' and 'as a deafened person I am used to reading body language and how the speaker feels', 'should be obvious from their demeanour' or 'indifferent'.

Other respondents felt it 'depends on the context', wanted to know about 'whispers/shouts/mumbles' or felt that 'emotions of speakers in vision should be obvious' and that information was only needed for 'speakers off screen'.

A number of respondents found that ?stupid background?or ?that LOUD + GHASTLY AWFUL? music interfered with ?what little I can hear?. ?TURN IT OFF.? (their capitalisation). Others wanted ?anything that the music or sound conveys to hearing people that deaf people would miss?, ?something simple such as ?scream in the distance?? or considered that ?you get the atmosphere through the subtitle? or ? only if it isn?t obvious?. ?A more systematic set of symbol representations for ?atmosphere so I could learn what to expect? was suggested, with concerns about the use of ????? for ?all the types of background music, but no indications saying ?scary? ?romantic?, ?comic?, etc.?. Concern was also expressed that it was not possible ?to translate background music into words that will convey atmosphere. ? To say ? scary music? or ?romantic music? does not convey anything remotely like the emotional impact such music has on the human mind.? The 19 comments about information on speakers? loudness, speed and other features expressed interest in information about accents, tone of voice and a ?clearer indication of who is speaking, especially when there are lots of characters?, with the suggestion of a combination of colour coding and the name of the speaker.

Feature	Very important	Important	Indifferent	Not want	Really not want	Unsure	Response Rate
Feelings of main characters	26.0%	35.1%	22.1%	10.4%	3.9%	2.6%	92.8%
Feelings of other characters	18.2%	35.1%	27.3%	9.1%	6.5%	3.9%	92.8%
Atmosphere	12.3%	37.0%	28.8%	11.0%	6.8%	4.1%	88.0%
Loudness of main characters	16.9%	23.4%	44.2%	9.1%	5.2%	1.3%	92.8%
Loudness of other characters	15.1%	17.8%	38.4%	17.8%	8.2%	2.7%	88.0%
Speed of speech of main characters	21.6%	10.8%	40.5%	10.8%	8.1%	8.1%	89.2%
Speed of speech of other characters	13.5%	16.2%	39.2%	17.6%	9.5%	4.1%	89.2%

Table 6 - Importance of information about particular features of speech and the atmosphere.

Respondents were equally divided between those who did not want additional information in the form of a small picture and those who would like a picture under some circumstances, for instance if easy to understand or together with text. Nearly two thirds preferred text descriptions for background sounds over small pictures of the sounds, though a third wanted either small images or a combination of images and text. Some 23 respondents commented about the types and 14 about the details of background sounds they were interested in. Many of the comments were on the lines of ?only if important?, ?relevant?, ?essential to understanding the plot? or ?out of view?. Other respondents suggested particular types of sounds, such as dogs barking, door bells, phones, ?mumbling hmms, ahhs, radio chatter, animal noises, clocks ticking?. Respondents were interested in the text of songs, with well over a third wanting the text of all songs, even in the background, and another quarter wanting those that were part of the scene. Only just under a fifth did not want the text of any songs.

Attitudes to the different suggested representations of the atmosphere, emotions and features of speech were very varied (See Table 7). The highest mean score (maximum value 5) was 3.1, while several scores were under 2 and response rates ranged from 7 to 53%, indicating a lack of enthusiasm for all the representations. However, a significant minority of respondents were very interested in several of the options, as shown by the relatively high number of respondents with scores of 4 and over.

The aim of the graphical representations was to produce the same type of emotional and visceral reaction as occurs to atmospheric music and, to a lesser extent, strong emotions in speech. However, the highest scores were obtained for short text descriptions for the atmosphere and emotions and modifications of the text format (letter size in all cases, bold or italic for atmosphere and features of speech, different fonts for features of speech and colour for atmosphere), rather than graphical representations. Signers had higher preferences than speakers for graphical representations, in particular, atmospheric images (2.9 cf. 1.7) and atmospheric photos (3.6 cf. 2.0) and for different colours for the atmosphere, but the difference was only statistically significant for atmospheric photos (p=0.05). Unsurprisingly signers had considerably higher preferences for photos of signing hands (3.4 cf. 0.8), a combination of photos of signing hands, colour, size and font (3.5 cf. 1.0) for representing features of speech, with both differences statistically significant (p=0.000 and p=0.010 respectively), but their higher preferences for colours and font were not statistically significant. Signers had higher preferences for all the options for representing emotions, with the greatest differences for the combination of small photos of a signer?s face and colour (2.8 cf. 1) and small photos of a signer?s face (2.6 cf. 1.5), but the differences were not statistically significant.

There were 12 comments on representing the atmosphere, 10 on emotions and six on features of speech. Respondents noted the importance of ?balance ?. between information and the time we have to digest it?, that ?a few words are enough? and that ?we deaf people are all face watchers? and wanted ?less pictures of the back of people?s heads?, as ?you can also guess how people are speaking by the facial expressions?. Others wanted to ?grasp the emotion and feel what they are feeling?. There were suggestions of the use of emoticons and that symbols and words may be preferred to colour, with subtitle colours not being visible when watching programmes on vPlayer. The small number of additional comments showed a preference for text, for instance that ?pictures are patronising?. Another respondent noted the need for a consistent use of ?symbols/pictograms? which were ?internationally agreed? and questioned whether this was ?realistic?.

Type of representation	Atmosphere			Emotions			Features of speech		
	Mean score	No. at 4+	Answer Rate	Mean score	No. at 4+	Answer Rate	Mean score	No. at 4+	Answer Rate
Atmospheric paintings	2.0	12	51.8%	N/a	N/a	N/a	N/a	N/a	N/a
Atmospheric photos	2.4	17	53.0%	N/a	N/a	N/a	N/a	N/a	N/a
Small photo of face showing emotion	N/a	N/a	N/a	1.6	8	N/a	N/a	N/a	N/a
Photo of signing hands	N/a	N/a	N/a	N/a	N/a	N/a	1.5	6	48.2%
Short text description	3.1	20	50.6%	2.8	14	41.0%	N/a	N/a	N/a
Colour	2.5	15	54.2%	1.7	9	50.6%	2.4	10	42.2%
Letter size	3.0	3	7.2%	2.6	17	51.8%	2.6	12	50.6%
Bold or italic	2.8	13	27.7%	2.3	14	50.6%	2.7	14	50.6%
Different fonts	2.1	4	28.9%	2.0	9	50.6%	2.9	7	39.8%
Photo of face & colour	N/a	N/a	N/a	1.4	10	47.0%	N/a	N/a	N/a
Signing hand photo, colour, size & font	N/a	N/a	N/a	N/a	N/a	N/a	2.0	10	41.0%

Table 7 - Representation of the atmosphere, emotions and features of speech.

Respondents had greater experience of watching educational programmes, videos and DVDs than attending subtitled lectures or classes, though nearly twice as many had very frequently attended subtitled lectures or classes as very frequently watched educational programmes, DVDs and videos, 60% of respondents preferred subtitles for recorded materials, whereas the preferred option for lectures, though by a minority of 40%, was a combination of subtitles and sign language interpretation. A significant minority wanted sign language interpretation on its own in both cases.

There was limited interest in information on features of the lecturer?s speech or the context in both face-to-face and recorded educational materials, though a significant minority of between just over a quarter and just under a third was very interested in most of the features (Table 8). The greatest interest was in information about the lecturer?s and students? emotions, with similar average values and numbers of respondents who were very interested (scores of 4+) in both cases. Signers were considerably more interested than speakers in all the features. Their greatest interest was in the speed of speakers for both lectures (3.2 cf. 1.1) and recorded materials (3.1 cf. 1.1) and the differences in both these cases were statistically significant (p=0.017 and p=0.018).

One respondent noted that Open University programmes used to have subtitles, ?but most of them disappeared a few years ago. Other than that, subtitled educational material is virtually non-existent.? Additional comments included the importance of not ?dumbing down? subtitles and the importance of the ?largest group subtitles cater for? having ?correct access?, but personalisation to meet the different needs of different (groups of) deaf people is now possible.

	Mean score	No at 4+	Response rate	Mean score	No at 4+	Response rate
Loudness of lecturer	1.9	10	43.4%	1.8	11	46.7%
Speed of speech of lecturer	2.0	11	42.2%	1.8	9	42.2%
Background sounds	2.0	12	39.8%	1.7	6	43.4%
Background music	N/a	N/a	N/a	1.4	5	42.2%
Emotions of lecturer or students	2.6	14	38.6%	2.5	12	39.8%

Table 8 - Interest in information about the lecturer?s speech and the context.

Discussion and Conclusions

This paper has presented the results of a survey of deaf people in the UK and Poland on their experiences, attitudes to and requirements of subtitles, including for the representation of emotions and contextual features. It was introduced by a literature overview which demonstrated both the importance of subtitles to deaf people and that they have been criticised for not representing emotions and contextual features and poor legibility in some circumstances. The small number of previous surveys have generally focused on a small number of issue(s) and been limited to one country. A few techniques for representing emotions and contextual features in subtitles have been developed and the reactions of small numbers of viewers investigated using short video clips.

The survey resulted in 83 useful responses from respondents with sufficiently diverse demographic characteristics to give a good representation of the different perspectives in the deaf community. Differences due to country, gender and language (signing or speaking) will be discussed in a subsequent paper. The results confirmed the great importance of subtitles to deaf people. Watching television and viewing DVDs or videos, but not cinema films, were very important activities for the majority of them. Subtitles were (very) important in all three cases and (significantly) improved understanding for over 80% of respondents. In line with the literature, 80% wanted verbatim subtitles either always or unless they were very complicated.

Comments indicated that they considered simplification ?very patronising? and wanted to be ?on an equal footing with hearing people?. A significant minority of about 10% found subtitles (very) difficult to read, that they were displayed for insufficient time and wanted longer to read them. This is probably related to the fact that just over 11% found reading (very) difficult. A majority of respondents considered that at least some interesting TV programmes, DVDs and videos lacked subtitles though UK respondents noted the ?good [TV] coverage?. They were generally happy with the quality of pre-recorded and DVD subtitles, but considered that live subtitles needed to be improved, but not by revoicing, which had made the situation worse. Concerns were expressed about the lack of information about important sounds on DVDs, the lack of subtitling of extras and the lack of information from online sellers about whether a DVD has subtitles. Cinema-goers were concerned that advertised subtitles were not always turned on.

A majority of respondents wanted information about the feelings of the main characters and the text of either all songs or those that were part of the scene. Significant minorities were interested in information about the atmosphere, the feelings of other characters, the loudness of all characters and background sounds. The suggested graphical representations aimed to give equivalent emotional and visceral reactions to music. However, the highest average scores were obtained by short text descriptions and modifications of the text format, though signers showed greater interest in graphical representations than speakers. Although the average interest in the suggested representations was low, most of the representations were of interest to a significant minority of respondents.

Respondents had greater experience of subtitles on recorded than face-to-face educational presentations. The majority preferred subtitles for recorded materials. The largest preference by a significant minority was for a combination of subtitles and sign language interpretation for lectures. Interest in all the suggested contextual features was fairly low, but greatest for information about lecturers? and students? emotions. Signers were more interested than speakers in all the types of information, with their greatest interest in the speed of speech.

The varying degrees of interest expressed in receiving information about emotions and different contextual features, as well as in the different representations, and the minority of respondents who wanted a shortened version of subtitles with longer to read them indicate the need for research to develop personalisation systems for subtitles, including verbatim and edited versions. Personalisation will allow viewers who want them to benefit from graphics, colour and/or animation without irritating others or making the text more difficult to read for them. The increased flexibility and options provided by digital media have made personalisation feasible.

This leads to the following recommendations:

- Care in the use of (loud) music, particularly when characters are talking, to avoid making it difficult for deaf people to understand the speech.
- Separate instructions for background music, analogously to closed subtitles, to enable it to be turned on or off and the volume adjusted separately from the main programme.
- Subtitling of all cinema films and DVDs, including ?extras? and ensuring that advertised subtitles are turned on.
- Verbatim subtitles and the inclusion of the text of songs that are part of the scene as the default.
- Research, technical development and large scale end-user testing with deaf and hearing people and long video clips in various genres:
- Live subtitling systems with improved accuracy, synchronisation and quality. This should be considered a priority.
- Subtitle personalisation systems which allow various features to be easily turned on and off, including shorter text in simple language, the texts of background songs and various representations of emotions and contextual features.
- Suites of options for representing a wide range of emotions and contextual features.

Acknowledgements

References

Anon. 2012a. ?Subtitle (captioning)?. [Internet]. Accessed 23 May 2012. Available from: https://en.wikipedia.org/wiki/Subtitle_%28captioning%29 [45].

Anon. 2012b. ?Digital television transition?. [Internet]. Accessed 23 May 2012. Available from: https://en.wikipedia.org/wiki/Digital_television_transition#Transitions_completed [46].

Baker, R. 1985. ?Subtitling television for deaf children?. *Media in Education Research Series* 3: 1-46.

Bok?an-Cullen, Desmond. 2012. ?Sharing the feeling, deafened people and silent television?. [Internet]. Accessed 4 May 2012. Available from: http://www.livedescribe.com/wiki/live/shared/shared/clt2/Emily/ASID/CHAIR/Deafened_and_silent_TV.pdf [47].

Boyce, Michael et al. 2012. ?Online enhanced captioning guidelines?. [Internet]. Canadian Network for Inclusive Cultural Exchange, Accessed 31 May 2012. Available from: http://cnice.idrc.ocad.ca [48].

Braverman, Barbara; Herzog, Melody. 1980. ?The effects of caption rate and language level on comprehension of a captioned video presentation?. *American Annals of the Deaf* 125:143-148.

Degan, Singh; Fels, Deborah. 2001. ?Expressing non-speech information and emotion to the deaf and hard of hearing?. CD-ROM Proceedings of 2001 IEEE Systems, Man and Cybernetics Conference, Tucson, Arizona.

EFHOH 2011. ?Subtitling ? our door to information society, state of subtitling access in EU 2011?. [Internet]. Accessed 19 April 2011. Available from: http://www.efhoh.org/mp/db/file_library/x/IMG/30890/file/StateofsubtitlinginEU23March2011.pdf [49].

Fels, Deborah; Daniel, Lee; Branje, Carmen; Hornburg, Matthew. 2005. ?Emotive captioning and access to television?. Proceedings 11th Americas Conference Information Systems. 11-14 August; Omaha, Nebraska. Accessed 23 January 2013. Available from: http://pdf.aminer.org/000/239/965/towards_emotive_captioning_for_interactive_television.pdf. [50]

Fels, Deborah; Polano, Lorelle; Harvey, Terry; Silverman, Charles. 2001. ?Towards emotive captioning for interactive television?. *Universal Access in HCI: Towards An information Society for All: Human Factors and Ergonomics,* Stephanidis, Constantine (editor). CRC Press.

Fourney, David; Fels, Deborah. 2008. ??Thanks for pointing that out.? Making sarcasm accessible for all?. Proceedings Human Factors and Ergonomics Society 52nd Annual Meeting.

Gordon-Salant, Sandra; Callahan, Julia. 2009. ?The benefits of hearing aids and closed captioning for television viewing by older adults with hearing loss?. *Ear Hear* 30 (4). Accessed 28 January 2013. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2820302/. [51] http://dx.doi.org/10.1097/AUD.0b013e3181a26ef4 [52]

Hersh et al. 2009. ?Representing emotions and the context in subtitles?. Conference and Workshop on Assistive Technologies for People with Hearing and Vision Impairments. 20-23 April; Wroc?aw, Poland.

Hersh et al. 2010. ?Representing contextual features of subtitles in an educational context?. Tenth IEEE International Conference on Advanced Learning Technologies. 5-7 July; Sousse, Tunisia.

Hersh. 2010. ?Methodological issues in multi-country multi-language participative research with blind and visually impaired people?. SWIIS ?10. 27-29 October; Pristina, Kosovo.

Hersh. 2011. ?Participative research with diverse end-user groups: multi-language, multi-country blind and visually impaired people?. Eighteenth IFAC World Congress. 28 August ? 2 September; Milan, Italy.

ITC 1999. ?ITC Guidance on Standards for Subtitling?. Accessed 16 January 2013. Available from: http://www.ofcom.org.uk/static/archive/itc/itc_publications/codes_guidance/standards_for_subtitling/index.asp.html [53].

Jensema, Carl 1998. ?Viewer Reaction to Different Captioning Speeds?. *American Annals of the Deaf* 143 (4): 284-292. http://dx.doi.org/10.1353/aad.2012.0073 [54]

Jensema, Carl; Burch. Robb. 1999. ?Caption speed and viewer comprehension of television programs: final report for federal award number H180G60013?, Office of Special Education Programs, US Department of Education, Washington, District of Columbia.

Jensema, Carl; McCann, Ralph; Ramsey, Scott. 1996. ?Closed Captioned Television Presentation Speed and Vocabulary?. *American Annals of the Deaf* 141 (4): 284-292. http://dx.doi.org/10.1353/aad.2012.0377 [55]

Karamitroglou, Fotios, 1998. ?A Proposed Set of Subtitling Standards in Europe?. *Translation Journal* 2 (2). Accessed 23 January 2013. Available from: http://translationjournal.net/journal/04stndrd.htm [56].

Kirkman, T.W. 1996. ?Statistics to Use?. Accessed 25 December 2012. Available from: http://www.physics.csbsju.edu/stats/. [57]

Lee, Daniel; Fels, Deborah; Udo, John. 2007. Emotive captioning, *ACM Computers in Entertainment* 5 (2). Accessed 23 January 2013. Available from: http://delivery.acm.org/10.1145/1280000/1279551/a11-lee.pdf?ip=130.209.6.40&acc=ACTIVE%20SERVICE&CFID=173102091&CFTOKEN=94532277&__acm__=1358923220_e47b4a9976d7b7c278c81a [58]

Lee, Johnny; Forlizzi, Jodi; Hudson, Scott. 2002. ?The Kinetic Typography Engine: an Extensible System for Animating Expressive Text?. Fifteenth Annual Symposium on User Interface and Software Technology. 27-30 October; Paris, France.

Ofcom 2005. ?Subtitling ? an issue of speed?? Accessed 20 January 2013. Available from: http://stakeholders.ofcom.org.uk/market-data-research/other/tv-research/subt/. [59]

Ofcom 2007. Office of Communications. Accessed 3 December 2007. Available from: http://www.ofcom.org.uk/ [60]

Ohene-Djan, James; Wright, Jenny; Crombie-Smith, Kirsty. 2007. ?Emotional subtitles: a system and qualitative survey of potential applications for deaf and hearing impaired people?. Conference and Workshop on Assistive Technologies for People with Hearing and Vision Impairments, 28-31 August; Granada, Spain.

Ohene-Djan, James; Shipsey, Rachel. 2006. ?E- subtitles: emotional subtitles as a technology to assist the deaf and hearing-impaired when learning from television and film?. Sixth IEEE International Conference on Advanced Learning Technologies. 5-7 July; Kerkrade, Netherlands: 464?466. http://dx.doi.org/10.1109/ICALT.2006.1652472 [61]

Power, Des; Leigh, Gregory, 2000. ?Principles and Practices of Literacy Development for Deaf Learners: a Historical Overview?. *Journal of Deaf Studies and Deaf Education* 5 (1): 3-8.

Rashid, Raisa 2008. ?Representing emotions with animated text?. MSc Thesis, University of Toronto.

Rashid, Raisa; Aitken, Jonathon; Fels, Deborah. 2006. ?Expressing Emotions using Animated Text Captions?, Computers Helping People with Special Needs ? Tenth International Conference. 11-13 July; Linz, Austria: 24-31.

Rashid, Raisa; Vy, Quoc; Hunt, Richard; Fels, Deborah. 2008. ?Dancing with words: using animated text for captioning?. *International Journal of Human-Computer Interaction* 24(5): 505-519. http://dx.doi.org/10.1080/10447310802142342 [62]

Schilperoord, Joost; de Groot [63], Vanja; van Son [64], Nic. 2005. ?nonverbatim captioning in Dutch television programs: a text linguistic approach?. *Journal of Deaf Studies and Deaf Education* 10 (4): 402-416. http://dx.doi.org/10.1093/deafed/eni038 [65]

Shilling, Russell; Zyda, Michael; Wardynzki, Casey. 2002. ?Introducing emotion into military simulation and videogame design: America?s army operations and VIRTE?. Proceedings of GameOn Conference 28-30 November; Harrow, England: 151-154.

Silverman, Charles; Fels. Deborah. 2002. ?Emotive captioning in a digital world?. Computers Helping People with Special Needs ? Eighth International Conference. 15-20 July; Linz, Austria: 292-294.

Thorn, Frank; Thorn, Sondra. 1996. ?Television captions for hearing-impaired?. *Human Factors* 38: 452-463. http://dx.doi.org/10.1518/001872096778702006 [66]

Tyler, Michael; Jones, Caroline; Grebbenikov, Leonid; Leigh, Gregory; Noble. William; Burnham, Denis. 2009. ?Effect of Caption Rate on the Comprehension of Educational Television Programmes by Deaf School Students?. *Deafness and Education International* 11(3): 152-162.

Vy, Quoc; Mori, Jorge; Fourney, David; Fels, Deborah. 2008. ?EnACT: a software Tool for Creating Animated Text Captions?. Computers Helping People with Special Needs ? Eleventh International Conference. 9-11 July; Linz, Austria: 609-616.

Wald, Mike. 2006. ?An exploration of the potential of automatic speech recognition to assist and enable receptive communication in higher education?. *ALT-J* 14: 9-20. http://dx.doi.org/10.1080/09687760500479977 [67]

Endnotes

1. The term ?deaf? will be used, unless otherwise indicated, to indicate deaf, Deaf, deafened, hard of hearing and hearing impaired people and ?Deaf? for people who sign and consider themselves members of the Deaf Community.

Cite this article as: Hersh, Marion. 2013. ?Deaf people?s experiences, attitudes and requirements of contextual subtitles: A two-country survey?. *Telecommunications Journal of Australia* 63 (2): 23.1-23.14. Available from: http://dx.doi.org/10.7790/tja.v63i2.406 [68]

Cite this article as:

Marion Hersh. 2013. *Deaf people's experiences, attitudes and requirements of contextual subtitles: A two-country survey* TJA, Vol 63, No 2, Article 406. http://doi.org/10.18080/TJA.v63n2.406 [70]. Published by Telecommunications Association Inc. ABN 34 732 327 053. https://telsoc.org [71]

television	Deaf people and contextual subtitles	The Telstra-TJA Christopher Newell Prize papers	deaf people	Accessibility
[72]	[73]	[74]	[75]	[76]
contextual features	subtitles			
[77]	[78]			

n2/a406#Hersh_2011 [43] https://telsoc.org/journal/jtde-v63-n2/a406#Kirkman_1996 [44] https://telsoc.org/journal/jtde-v63-n2/a406#Wald_2006 [45] https://en.wikipedia.org/wiki/Subtitle_%28captioning%29 [46] https://en.wikipedia.org/wiki/Digital_television_transition#Transitions_completed [47] http://www.livedescribe.com/wiki/live/shared/shared/clt2/Emily/ASID/CHAIR/Deafened_and_silent_TV.pdf [48] http://cnice.idrc.ocad.ca/ [49] http://www.efhoh.org/mp/db/file_library/x/IMG/30890/file/StateofsubtitlinginEU23March2011.pdf [50] http://pdf.aminer.org/000/239/965/towards_emotive_captioning_for_interactive_television.pdf. [51] http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2820302/ [52] http://dx.doi.org/10.1097/AUD.0b013e3181a26ef4 [53] http://www.ofcom.org.uk/static/archive/itc/itc_publications/codes_guidance/standards_for_subtitling/index.asp.html [54] http://dx.doi.org/10.1353/aad.2012.0073 [55] http://dx.doi.org/10.1353/aad.2012.0377 [56] http://translationjournal.net/journal/04stndrd.htm [57] http://www.physics.csbsju.edu/stats/ [58] http://delivery.acm.org/10.1145/1280000/1279551/a11-lee.pdf?ip=130.209.6.40&acc=ACTIVE%20SERVICE&CFID=173102091&CFTOKEN=94532277&__acm__=1358923220_e47b4a9976d7b7c278c81a985f9f5c24. [59] http://stakeholders.ofcom.org.uk/market-data-research/other/tv-research/subt/ [60] http://www.ofcom.org.uk/ [61] http://dx.doi.org/10.1109/ICALT.2006.1652472 [62] http://dx.doi.org/10.1080/10447310802142342 [63] http://jdsde.oxfordjournals.org/search?author1=Vanja+de+Groot&sortspec=date&submit=Submit [64] http://jdsde.oxfordjournals.org/search?author1=Nic+van+Son&sortspec=date&submit=Submit [65] http://dx.doi.org/10.1093/deafed/eni038 [66] http://dx.doi.org/10.1518/001872096778702006 [67] http://dx.doi.org/10.1080/09687760500479977 [68] http://dx.doi.org/10.7790/tja.v63i2.406 [69] https://telsoc.org/copyright [70] http://doi.org/10.18080/ajtde.v63n2.406 [71] https://telsoc.org [72] https://telsoc.org/topics/television [73] https://telsoc.org/topics/deaf-people-and-contextual-subtitles [74] https://telsoc.org/topics/telstra-tja-christopher-newell-prize-papers [75] https://telsoc.org/topics/deaf-people [76] https://telsoc.org/topics/accessibility [77] https://telsoc.org/topics/contextual-features [78] https://telsoc.org/topics/subtitles