

Word recognition and content comprehension of subtitles for television by deaf children

Soledad Zarate, University College London

Joseph Eliahoo, Imperial College London

ABSTRACT

This project explores how deaf children read subtitles on television. The participants – recruited from years 3 to 6 of a mainstream school with a hearing impairment unit – were exposed to both broadcast and enhanced subtitles and their performances were compared. In particular, the focus is on identifying enhancements that can help children to understand subtitle content and to recognise new or difficult words. Among the enhancements introduced were repetition and highlighting of new or difficult words through the use of a bigger and different typeface, use of longer reading times, text reduction and careful spotting. This pilot study provides some useful information for future empirical experimental research on subtitling for deaf children.

KEYWORDS

Subtitling, deaf children, word recognition, content comprehension.

1. Introduction

The research aims, in a broad sense, at exploring how deaf children read subtitles on television. The approach is empirical and consists of practical observation conducted in the classroom through the use of short screenings of two subtitled episodes of the cartoon Arthur, followed by questionnaires aimed at assessing the children's comprehension of subtitles and their ability to recognise new or difficult words.

Three relevant studies motivated the sharpening of the initial question to assessing content comprehension and in particular recognition of new vocabulary. Firstly, Neuman and Koskinen (1992) examined how “comprehensible input,” as intended by Krashen (1985), in the form of subtitled television, influences incidental vocabulary learning in a second language (L2). Krashen argues that children learn L2 incidentally, through exposure, by focusing on the meaning rather than the form or grammar of the message. Students stretch their knowledge when they are provided with and receive “comprehensible input,” i.e. information that goes slightly beyond the students’ actual knowledge (Díaz Cintas and Fernández Cruz, 2008:203). Neuman and Koskinen (*ibid.*) conducted a study with 129 Southeast Asian and Hispanic bilingual hearing children in grades 7 and 8, aged 12 and 13, living in the US and having English as L2. Four different formats of a children’s television science production, 3-2-1 Contact, were considered: subtitled TV, TV without subtitles, reading along and listening to text, and textbook. They concluded that students incidentally learned more words from subtitled television than from any of the other three formats and also acquired content. Similarly, Koolstra et al. (1997)

conducted a three-year panel study with a sample of 1,050 Dutch hearing children in grades 2 and 4 and observed that vocabulary was identified as the only sub-skill that profited from watching subtitled programmes. It was also suggested that the development of decoding skills may be promoted since reading subtitles provides an opportunity to practise word recognition. Subsequently, Koolstra and Beentjes (1999) conducted a study with Dutch children aged 9-10 and 11-12, using a 15-minute documentary. The children were exposed to three different versions: 1) programme about grizzly bears with original English soundtrack and Dutch subtitles; 2) same programme with original English soundtrack and no subtitles; 3) programme about prairie dogs in original Dutch language (control). The subtitled version proved to be the one that benefited the acquisition of foreign words the most.

Another relevant study on the impact of subtitles on vocabulary recognition was conducted by d'Ydewalle and Van de Poel (1999), who presented a short subtitled cartoon to Dutch-speaking children (8-12 years old) with Danish and French subtitles. The fact that Danish is more similar to Dutch than French is to Dutch affected acquisition scores of Danish positively. In both the visual and auditory parts of the vocabulary test, acquisition effects emerged when Danish was available in the sound track; when Danish was present only in the subtitles, there was only acquisition in the visual part of the vocabulary test. In the French vocabulary test, no acquisition was apparent, except in the auditory test when the sound track contained the French language. This study was conducted with hearing children but it is partially relevant to deaf children as it showed that visual acquisition of vocabulary occurs with none or limited access to the auditory channel. Having collected evidence that subtitles encourage the acquisition of new vocabulary, the focus shifted towards assessing whether the introduction of certain techniques, not currently used in broadcast subtitles, could be more facilitative in the task of word recognition, here intended as the ability to recognise a word by sight without needing to apply word analysis skills. In accordance with evidence gathered by Ewoldt et al. (1992), the techniques chosen to facilitate word recognition were mainly repetition and highlighting (through the use of a bigger and different typeface), combined with the use of longer reading times, text reduction and careful spotting.

As far as reading comprehension is concerned, there are two contrastive views, i.e. the bottom-up (text-based) model and the top-down (reader-based) model. The bottom-up model, which begins at the bottom with letters and ends up at the top with comprehension, depends on the ability to name letters quickly and accurately, and to associate sounds with these letters. The top-down model, which begins at the top in the reader's head and ends up at the bottom with text, coincides with the story or the whole book approach adopted by the Leicestershire Service for Hearing-Impaired Children. Children are encouraged to make use of all reading cues: their knowledge of the world, the book, the characters, the language and the pictures. Words are considered in a holistic manner that goes beyond

phonics and takes into account not only their sound but also their shape and sight. The decoding process required by the bottom-up model in order to associate sounds to letters is laborious and does not seem to go hand in hand with the reading of subtitles, which are by nature immediate as they constantly appear and disappear from the screen. This suggests that they are unlikely to be used as a tool of reflection on text since the viewer, in a normal viewing context, has no control over the reading time at disposal. Different is the case of printed text as it allows the viewer to look at the text, think and analyse it at his/her own pace. However, the shape and appearance of the words may be noticeable in the subtitle reading process. This suggests that the top-down model is more likely to be applied in reading subtitles than the bottom-up model. Reading the subtitles is only one of the tasks involved in the decoding of the entire semiotic apparatus. The familiarity with the programme, the knowledge of the characters, and of course the understanding of the subtitles, placed in the moving images, all contribute to the comprehension of audiovisual programmes.

Part of the study question of this research concerns the reading comprehension of subtitle content. Different techniques were used in the enhanced subtitles with this purpose in mind: lower reading speeds, text reduction and consideration given to spotting and line-breaks. The choice of these techniques was made in line with recommendations by Ewoldt et al. (1992: 354) about the importance of text style as a device that facilitates comprehension and with good practices in subtitling (Ivarsson and Carroll 1998; de Linde and Kay 1999).

2. Materials

2.1. Clips

The cartoon Arthur (www.bbc.co.uk/cbbc/shows/arthur), which follows the adventures of Arthur Read, an eight year-old aardvark, was chosen for this study. It is broadcast by CBBC, the BBC channel that presents programmes tailored for school children aged between 6 and 12 years. The choice of the cartoon was dictated by the broadcast availability of the shortest episode for the intended age group. While the duration of the clip needed to be as limited as possible in order to maintain the children's attention spans, the clip also needed to be self-contained and intended for the age group in question. Two 12-minute clips were randomly selected: "A Portrait of the Artist as a Young Tibble" (Clip 1) and "War of the Worms" (Clip 2). Two clips were needed in order to expose the children to both broadcast subtitles and enhanced subtitles and to assess whether the enhancements introduced made a difference in the comprehension of subtitles and particularly in the word recognition task.

2.2. Selection of words for word recognition task

The process of selecting which words to include in the word recognition task was based on the following criteria: (1) word acquisition taken from the first 1,000 words (Fry et al. 2003) and (2) a computerised database of printed word frequencies as read by children aged between 5 and 9 (Stuart et al. 1993-1996, updated in 2003). A database of words which appear in books for children in the first two years of primary school was compiled and used to develop stimuli for experimental work investigating the literacy acquisition of young children. The authors state that

[f]or the first time researchers interested in the empirical investigation of the development of printed word recognition skills will have access to an up-to-date source of stimuli. This will allow stringent experimental control over variables such as word frequency, orthographic neighbourhood size and spelling-sound consistency at both grapheme-phoneme and rime levels. Teachers and other practitioners will be able to discover which words children need to know (and be taught) in order to read at a given level. The database will also allow the development of literature for children with reading difficulties with age-appropriate content presented in the highest frequency, earliest learnt vocabulary (Stuart et al 2003:3)

The clips were watched and a list of words likely to be new with some degree of complexity for deaf children aged between seven and 10 were selected for each episode. The number of words selected for each episode was dictated by the number of questions included in the test. Nine questions out of 13 were devoted to word recognition. The 18 words selected, listed in Table 3, had very low frequency (ranging between 0 and 162 in a million words) and did not appear among the first 1,000 words acquired by children as compiled by Fry et al. (2003). Researchers such as Caselli (1983), Gardner and Zorfass (1983), Schlesinger and Meadow (1972), and Stoloff and Dennis (1982) suggest that deaf children at the age of five have acquired 500 words as part of their vocabulary. This is the main reason why the first 1,000 words acquired by children according to Fry et al.'s database (*ibid.*) were excluded from the word recognition task, in order to ensure that the words considered were likely to be unknown by the children.

Word	Frequency (in a million words)
Artist	32
Coins	5
Cupcake	5
Educational	3
extraordinary	0
hey	89
Hoax	0
Lemonade	27
Painting	154
Scraps	5
Slime	19
Squad	8
Tomatoes	27
Tractor	30
Unique	0
Weird	8
Wings ¹	141
Wonderful	162

Table 1: Word selected for the word recognition task

2.3. Font

The font chosen for the enhanced subtitles is the sans-serif typeface Arial, size 30, as it is the closest – as confirmed by the publisher through personal correspondence – to a trademarked font developed by the independent published Barrington Stoke (www.barringtonstoke.co.uk) to make reading easier for reluctant readers. The use of sans-serif typefaces in printed material, be they books or newspapers, is rather unusual, while it is common on websites as it is believed that they work well in low-resolution computer screens. Below is an example of Arial typeface:

a b c d e f g h i j k l m n o p q r s t u v w x y z A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
--

Whenever a word is highlighted to encourage word recognition, the typeface and size change. The slab-serif Lucida Fax, size 33, is used on these occasions, as recommended on <http://screenfont.ca>. The peculiarity of slab-serifs is that the serifs – the small features at the end of the strokes within letters – are long and sit at right angles to the underlying strokes. They work well in low-resolution environments, where the image lacks sharpness. Below is an example of Lucida Fax typeface:

a b c d e f g h i j k l m n o p q r s t u v w x y z A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
--

The typeface used on digital television in the UK, as recommended by Ofcom (1999), is the Tiresias Screenfont, specifically designed in 1998 for screen display, with characters that are easy to distinguish² from each other (as it can be noted in the typefaces above, confusion can arise in Arial between lower-case 'l' and upper-case 'I'). The two episodes of Arthur with broadcast

subtitles used in the main experiment were recorded from digital CBBC, which used Tiresias Screenfont. Below is an example of this font:



The use of Tiresias Screenfont in the enhanced subtitles was not considered as it is particularly expensive to license and therefore the use of other typefaces (Arial and Lucida Fax) was researched.

2.4. Enhancement introduced for word recognition: repetition

One of the techniques adopted in the enhanced subtitles is the repetition of words whenever they are repeated in the soundtrack. Broadcast subtitles occasionally exclude a word that is repeated in the soundtrack, especially if it is not crucial in understanding the plot. For instance, the theme tune of the cartoon repeats 'hey' several times and sometimes the word is being sung by the character on screen. The broadcast version occasionally leaves 'hey' out, while enhanced subtitles reflect all repetition, especially since the word 'hey' is among those selected for word recognition.

Figure 1 shows how the broadcast subtitles use a descriptive label (last screenshot) to indicate the disagreement between the two characters, whereas the enhanced subtitles include the dialogue in a way that reflects the repetition of the word tractor which is among the ones included in the word recognition task. Also, the reading time has been increased where possible and, for instance, the last enhanced subtitle has a reading time of 51 wpm and stays on screen for 05:03 (5 minutes and 3 frames) whereas the last broadcast subtitle has a reading time of 61 wpm and stays on screen for 01:19. Note that the use of a bigger font in the enhanced version will be discussed in section 2.5. Also, in this specific example the padding expression I think has been sacrificed in order to allow for a calmer reading speed. Note that editing text to gain reading time is one of the techniques used for the purposes of word recognition and is discussed in section 2.6.

Broadcast

Enhanced



00:12:22:15 – 00:12:24:23 186 wpm

00:12:22:13 – 00:12:25:13 120 wpm



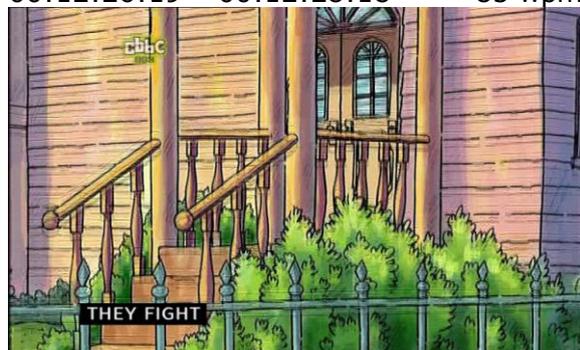
00:12:24:24 – 00:12:26:18 190 wpm

00:12:25:15 – 00:12:27:20 120 wpm



00:12:26:19 – 00:12:28:18 85 wpm

00:12:27:22 – 00:12:29:11 107 wpm



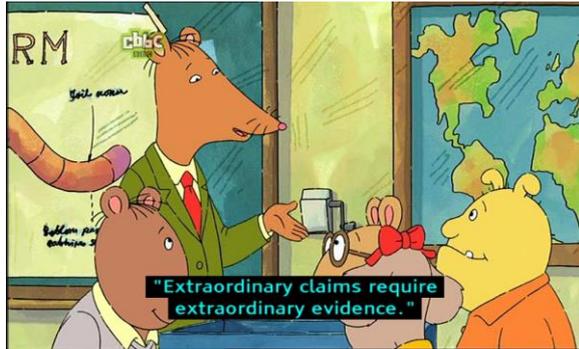
00:12:28:19 – 00:12:30:13 61wpm

00:12:29:13 – 00:12:34:16 51 wpm

Figure 1 Enhancement: repetition over reformulation³

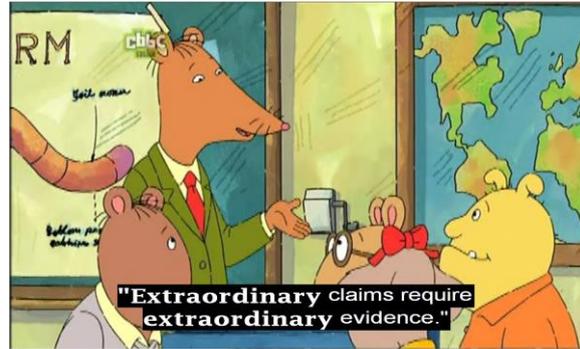
Enhanced subtitles are presented with the same layout and syntactical distribution whenever possible so as to maintain the visual similarity. Compare the broadcast subtitles with the enhanced subtitles in Figure 2:

Broadcast

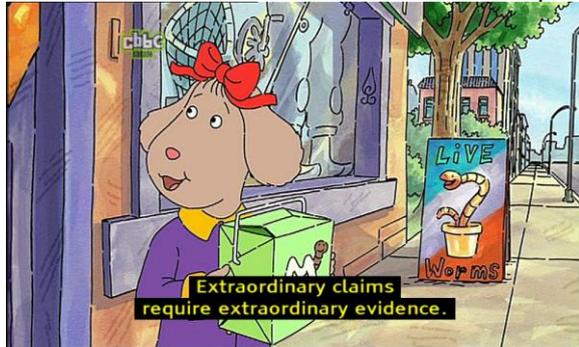


00:04:15:01 – 00:04:19:02 148 wpm

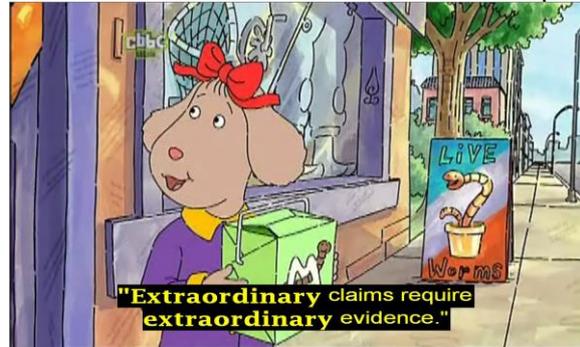
Enhanced



00:04:17:12 – 00:04:21:19 140 wpm



00:04:38:12 – 00:04:42:01 161 wpm



00:04:39:03 – 00:04:42:23 157 wpm

Figure 2 Enhancement: visual repetition

2.5. Enhancements introduced for word recognition: highlighting

Another technique used in the enhanced subtitles to support word recognition is highlighting of words. This is done by using: (1) a different typeface, that is Lucida Fax instead of Arial; (2) a bigger font size, that is increasing the font from size 30 to size 33; and (3) bold formatting. This technique is often accompanied by the use of longer reading times, which in this specific case is of 32 wpm for the first enhanced subtitle, made possible by leaving the subtitle on screen for the six-second maximum duration:

Broadcast



00:02:17:15 – 00:02:20:15 92 wpm

Enhanced



00:02:18:14 – 00:02:24:14 32 wpm



Figure 3 Enhancement: highlighting of words

2.6. Other enhancements introduced for word recognition: careful spotting, text reduction (omission) and longer reading times

While highlighting is the only new enhancement introduced for word recognition, the clips are subtitled in their entirety following good practices in subtitling (Ivarsson and Carroll 1998; de Linde and Kay 1999). As discussed in section 2.4., the words selected for the WR task are also intentionally repeated if repetition occurs in the soundtrack. This is the case for four of the words selected, as illustrated in Table 2 (Figure 1 shows one example).

Some of the subtitles that contain the words selected for the WR task present more careful and adequate spotting compared to the broadcast versions and/or longer reading times (both aspects are quantified in Table 2).

Also, in order to gain reading time for word recognition, some elements that are not necessary for understanding content (Ivarsson and Carroll:86) are omitted (in four occasions, as specified in Table 2). Díaz Cintas and Remael (2007:146) distinguish between partial reduction, achieved through condensation, and total reduction, achieved through omission of lexical items. Figure 4 shows an example of total reduction, where the conjunction 'and' and the adverb 'very' are omitted in order to gain reading time in favour of the word unique, selected for the word recognition task.

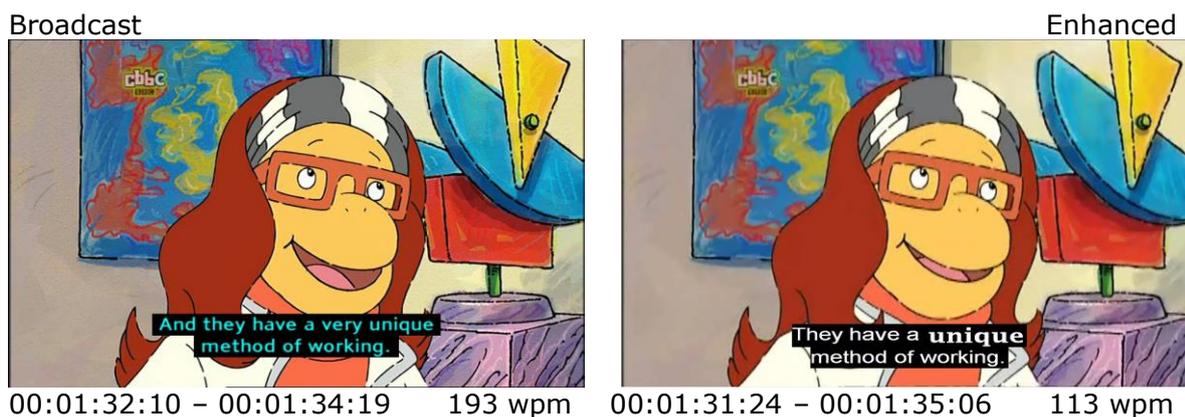


Figure 4 Enhancement: omission to gain reading time

Similarly, as illustrated in the first screenshot in Figure 1, 'I think' has been omitted to allow for longer reading time in favour of the recognition of the word 'painting.'

2.7. Enhancements introduced for content comprehension

Word recognition has been the main focus of this pilot study, mainly because this variable, unlike content comprehension, has been identified by researchers (d'Ydewalle and Van de Poel 1999; Koolstra *et al.* 1997; Koolstra and Beentjes 1999; Neuman and Koskinen 1992) as the one that can benefit from watching subtitled programmes. The strong rationale provided by literature for word recognition supported the introduction of new techniques, as discussed earlier, in favour of this variable. The case of content comprehension is different as there is no compelling evidence that subtitles can be used as a tool for comprehending text. Nevertheless, four questions (out of a total of 13) were included in the questionnaires with the intent of exploring a methodology to be developed in future studies to further research this aspect. Content comprehension is not achieved by the recognition of words or isolated expressions but by the ability of making sense of what is being read.

The technique more often used to facilitate content comprehension is text reduction (for instance, omission of interjections, i.e. 'hmm', 'I mean,' 'er' or of other lexical items), performed if the speech rate is high for the reading speed set (140 wpm) or if there is a complex or new expression that requires longer reading time than the one set. As reported in Table 2, seven of the eight cases selected for content comprehension present text reduction in the form of omission. Generally the omission of text determines an increase in reading time.

More careful spotting of subtitles is performed in six of the eight cases selected for content comprehension, as reported in Table 2. Figure 5 illustrates one example. As recommended by Ivarsson and Carroll (1998:77) lines are divided in such a way that "words intimately connected by logic, semantics or grammar are written on the same line wherever possible."

Broadcast



00:07:18:16 – 00:07:21:24 112 wpm



00:07:30:00 – 00:07:33:02 116 wpm



00:07:33:03 – 00:07:35:03 120 wpm



00:07:35:04 – 00:07:38:02 102 wpm

Enhanced



00:00:07:18:14 – 00:07:21:21 98 wpm



00:07:29:20 – 00:07:31:07 113 wpm



00:07:31:09 – 00:07:35:01 123 wpm



00:07:35:03 – 00:07:39:21 104 wpm



00:07:38:03 – 00:07:39:24 117 wpm

Figure 5 Enhancement for content comprehension: spotting

2.8. Summary of enhancements introduced

Table 2 sums up the number of enhancements introduced in the enhanced subtitles for each category. Repetition and highlighting were used only for the word recognition task, while careful spotting, text reduction (in the form of omission) and longer reading times were used to facilitate both word recognition and content comprehension tasks. Note that each questionnaire included nine word recognition questions and four content comprehension questions.

	Cases	Enhancements				
		Repetition	Highlighting	Text reduction	Careful spotting	Longer reading time
WR	18	4	16	4	6	6
C	8	-	-	7	6	7

Table 2: Enhancements

2.9. Questionnaires

Questionnaires were used as the evaluation method of the studies. In designing the questionnaires, one of the major concerns was easy comprehension by the children. Closed questions seemed to be suitable for the purpose of word recognition and content comprehension. The use of open-ended questions would have required writing abilities and would have constituted a major challenge to analyse, unnecessary for the purposes of this study. Multiple-choice questions were preferred over yes/no questions.

The questions were given four possible answers and were accompanied by colourful screenshots to make the activity amusing and to also help the children contextualise the questions by associating them with the video. The 'not sure' option was always given to discourage the children from guessing the answer. It was explained that the activity was not a test and that the not sure answer was as valid as any of the others.

Each questionnaire included nine questions aimed at testing word recognition (WR) and four aimed at testing the comprehension of subtitles (C). In this last instance, a successful performance did not depend on the ability to perform WR, but rather on the ability to understand the subtitle content. The higher number of questions on WR reflects the intention to look into vocabulary more than comprehension, as research suggests that this reading sub-skill benefits from watching subtitles (d'Ydewalle and Van de Poel 1999; Koolstra et al. 1997; Koolstra and Beentjes 1999; Neuman and Koskinen 1992). Questions were kept to a minimum in order to maximise the children's performances in terms of focus, especially since many deaf children are characterised by lack of concentration (Marschark 1993)⁵.

3. Participants

The participants were recruited from a mainstream school based in inner London and with a hearing impairment unit in place attended by approximately 70 deaf children (aged 3 to 11 years) that constitute one sixth of the entire school population. This choice was made with the intent of recruiting participants that were representative of the majority of deaf children. The selection of this type school was justified by evidence that the majority of deaf children – around 80% – are enrolled in mainstream schools (NDCS 2003) that use a monolingual auditory oral approach in their teaching, that is a method where children – with hearing aids and/or cochlear implants – are expected to develop listening skills and speech through the use of English and without the support of sign language or finger-spelling. Due to the nature of the study, which aims at studying the subtitling reading skills of deaf children, children enrolled from Year 3 to Year 6 were considered. The children recruited had a chronological age that varied between 7 and 10 years and had a reading age that varied between 64 and 126 months. Children enrolled in years below Year 3 were not considered as they were more unlikely to be able to read in a proficient way.

The children were divided into two groups that overall had a similar level of literacy for the purposes of counterbalancing. The study had a repeated measures design where all subjects were exposed to both broadcast and enhanced subtitles. Therefore all subjects in both groups needed to attend two studies. Children who only attended one study were subsequently excluded from the sample. While this was a disadvantage for sample size, this type of filtering of participants was used to remove individual differences between children as a potential confounding variable. The final sample included N = 11 for Group One and N= 9 for group Two. Two methods were in place in the school to assess the children's literacy: 1) the Salford Sentence Reading Test⁶ and 2) the Progress with Meaning (PM) benchmark⁷. The children were administered the Salford Sentence Reading Test yearly and the PM benchmark every term. Through the use of these two tests, the children's reading ages were determined. Based on the results of these two tests, children were allocated to one of the two groups.

The average reading age of children in Group One was 99 months and 103 months in Group Two.

Generally speaking, the population of deaf children is small in size. Out of 13 million children in the UK, there are more than 45,000 who are deaf (Action on Hearing Loss, 2011), that is three in one thousand. Children with some additional needs were not excluded from the sample on the basis that 40% of all deaf children have some extra health, social or educational need (NDCS, 2011). These children constitute a considerable proportion of the entire population and therefore need to be taken into consideration. Two children in the sample had visual impairment, corrected by the use of spectacles. Two other children had unspecified learning difficulties and one other child had suspected learning difficulties. One child had difficulty in focussing and one other had suspected language disorder.

4. Study design

The participants were presented with two clips, one with broadcast subtitles and one with enhanced subtitles. An assessment was made as to whether the enhancements introduced improved the two units of analysis, namely word recognition (WR) and comprehension of subtitle content (C). In specific, the study aimed to examine whether and to what extent the newly introduced techniques favoured WR and C.

To control for differences between film clips and order effects, the tasks were counterbalanced. By using this counterbalancing measure, the participants' possibility of learning something in the first task that could help them to perform better in the second task was neutralised. The study had a repeated measures design, that is the children received both levels (broadcast and enhanced) of the independent variable (subtitles) but in an inverted order. Also, two different episodes of the same cartoon – Arthur – were chosen so that the participants, having already seen the clip, were not at an advantage in performing the second task. Below, Table 3 shows how the study was designed. Note that the viewings are numbered chronologically:

5. Findings

A quantitative approach was chosen as it allows more statistical analysis to be used to answer relevant research questions:

- 1) Do enhanced subtitles improve the subjects' performance as far as the word recognition (WR) task is concerned?
- 2) Do enhanced subtitles improve the subjects' performances as far as the content comprehension (C) task is concerned?
- 3) Taking into account the subtitle condition (broadcast versus enhanced), how does reading age affect the performance?

The first two research questions demand a descriptive answer. The third question requires a quantitative study as it is testing the hypothesis that the higher the reading age, the better the performance, and is also looking at any potential differences between broadcast and enhanced subtitles.

Stata version 10 was used for the analysis.

Group One	Clip 1B	Broadcast subtitles	Viewing One
	Clip 2E	Enhanced subtitles	Viewing Three
Group Two	Clip 1E	Enhanced subtitles	Viewing Two
	Clip 2B	Broadcast subtitles	Viewing Four

Table 3: Word selected for the word recognition task

5.1. Order of presentation and clip effects

Group One was administered broadcast subtitles followed by enhanced subtitles, while Group Two was administered enhanced subtitles followed by broadcast subtitles. The Mann-Whitney test was used to determine whether the order of presentation had an effect on the performance of the word recognition (WR) task by comparing the total WR scores for broadcast and enhanced subtitles. No evidence of a difference in WR total broadcast scores between orders of presentation ($p=0.465$) was noted and only marginal evidence of a difference in WR was noted for total enhanced scores between orders of presentation ($p=0.094$).

Fisher's exact test was used to compare content comprehension (C) results against order of presentation. There was no evidence of an association between C total broadcast scores and order of presentation ($p=0.670$) or between C total enhanced scores and order of presentation ($p=0.406$).

In view of the results obtained, the data for the two groups was merged and the focus set on differences in the performances with broadcast and enhanced subtitles.

5.2 Enhanced subtitles versus broadcast subtitles

The Paired Sign test was used to compare differences between broadcast and enhanced subtitles total scores for WR. No evidence of a difference was found ($p=0.2379$).

McNemar's test was used to look at differences between broadcast and enhanced subtitles total scores for C. No evidence of a difference was found ($p=0.375$).

As far as word recognition is concerned, the median WR total score for enhanced subtitles is 6.5 against 5.5 for broadcast subtitles.

5.3 Reading ages versus performance scores for word recognition

It was considered appropriate to look at whether there were any differences in the way children performed with broadcast and enhanced subtitles depending on their reading age. Spearman correlation was used to look at relationships between the total WR scores and the reading age. The median reading age is 99 months; the range goes from 64 to 126 months.

There is strong evidence of a relationship between WR total broadcast scores and reading age ($r_s=0.705$, $p=0.0005$) and between WR total enhanced scores and reading age ($r_s =0.707$, $p=0.0005$).

5.4 Reading ages versus performance scores for content comprehension

The two-sample Wilcoxon rank sum (Mann-Whitney) test showed evidence of a difference in reading age between high and low C total enhanced scores ($p=0.004$) and strong evidence of a difference in reading age between high and low C total broadcast scores ($p=0.001$).

6. Discussion

The original research question intended to look at the role of subtitles in relation to two main reading variables, that is word recognition (WR) and content comprehension (C).

The research on WR in a subtitling context is limited to the field of second-language acquisition by hearing children (d'Ydewalle and Van de Poel 1999; Neuman and Koskinen 1992; Koolstra et al. 1997; Koolstra and Beentjes 1999). However, this research is to a certain extent also applicable to deaf children on the grounds that the setting is the same – watching subtitles on a screen – and that the language, for different reasons, is somehow unfamiliar. There is evidence that subtitles may facilitate vocabulary (Neuman and Koskinen 1992) and improve the development of word recognition (d'Ydewalle and Van de Poel 1999; Koolstra et al. 1997; Koolstra and Beentjes 1999). As Koolstra et al. (1997) explain, the evidence that subtitles benefit the acquisition of vocabulary cannot however be extended to the task of reading comprehension because subtitles do not provide practice in comprehending coherent texts. Reading subtitles is a different task from conventional reading. The immediacy of the subtitles that appear and disappear from the screen in a limited space of time, the reading rate and the segmentation of text are elements that do not favour the comprehension of coherent texts. With these premises in mind, WR was the main focus of the research but space was also given to C, where specific

techniques were introduced to test hypotheses generated by previous research.

The expected strong evidence of a relationship between reading age and reading performance for both the WR and C variable is applicable to both broadcast and enhanced subtitles. However, there is no evidence that the enhanced scores are higher than the broadcast ones, but only a tendency is shown for the WR variable.

In order to understand the results, it needs to be noted that while the participants had chances of having been exposed to broadcast subtitles previous to the study, their exposure to enhanced subtitles was a complete novelty. While some of the enhancements – repetition of words, use of longer reading times, careful spotting – were less noticeable, the highlighting of new or difficult words and the switch to a different typeset of bigger size, was a complete novelty in the subtitling practice. This could have distracted the participants and could have interfered with their WR and C performances. The participants were not given any training on what the enhancements meant, so they were left to work it out for themselves. A potential consequence was that the enhancements were in fact a distraction and so did not contribute significantly positively to scores, despite a tendency for the enhanced scores to be higher than the broadcast score for the WR variable. If the participants were alerted to the significance of the enhancements, it would be interesting to look at whether they would make a better use of them in their reading performances. It would be useful in future studies to test whether once the novelty factor is neutralised, the impact of the enhancements on the performances is greater.

This pilot study is a first attempt into conducting empirical experimental research on subtitling for deaf children. A power calculation was conducted in order to determine how much larger a sample would be needed for the results to be valid. A sample size of 44 will have 90% power to detect a difference in means of -1 (e.g. a First condition mean, m_1 , of 5.5 and a Second condition mean, m_2 , of 6.5), assuming a standard deviation of differences of 2, using a paired t-test with a 0.05 two-sided significance level.

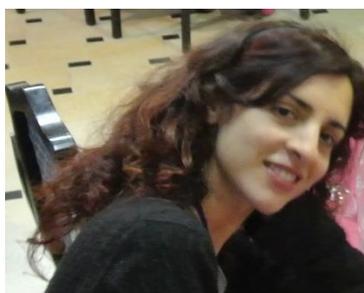
Bibliography

- **Action on Hearing Loss** (2011). *Facts and figures on deafness and tinnitus*. London: The Royal National Institute for Deaf People.
- **Bookbinder, Geoffrey E.; Denis Vincent and Mary Crumpler** (2002). *Salford Sentence Reading Test* (revised). London: Hodder Education.
- **Caselli, M. Cristina** (1983). "Communication to language: Deaf children's and hearing children's development compared." *Sign Language Studies* 38, 1-23.
- **De Linde, Zoé and Neil Kay** (1999). *The Semiotics of Subtitling*. Manchester: St. Jerome.
- **Díaz Cintas, Jorge and Marco Fernández Cruz** (2008). "Using subtitled video materials for foreign language instruction." Jorge Díaz Cintas (ed.) *The Didactics of Audiovisual Translation*. Amsterdam/ Philadelphia: John Benjamins, 201-14.
- **d'Ydewalle, Géry and Marijke Van de Poel** (1999). "Incidental foreign-language acquisition by children watching subtitled television programs." *Journal of Psycholinguistic Research* 28(3), 227-244.
- **Díaz Cintas, Jorge and Aline Remael** (2007). *Audiovisual Translation: Subtitling* (Translation Practices Explained, 11). Manchester: St. Jerome
- **Ewoldt, Carolyn; Neita Israeilite and Ron Dodds** (1992). "The ability of deaf students to understand text. A comparison of the perceptions of teachers and students." *American Annals of the Deaf* 137, 351-361.
- **Fry, Edward Bernar; Jacqueline E. Kress and Dona Lee Fountoukidis** (2003) *The Reading Teachers Book of Lists*. Third edition. Englewood Cliffs, NJ: Prentice Hall.
- **Gardner, Judith and Judith Zorfass** (1983). "From sign to speech: the language development of a hearing impaired child." *American Annals of the Deaf* 129, 20-24.
- **Ivarsson, Jan and Mary Carroll** (1998). *Subtitling*. Simrishamn: TransEdit.
- **Koolstra, Cees M.; Tom H.A. Van der Voort and Leo J. Th. Van der Kamp** (1997). "Television's impact on children's reading comprehension and decoding skills: a 3-year panel study." *Reading Research Quarterly* 32(2), 128-152.
- **Koolstra, Cees M. and Johannes W.J. Beentjes** (1999). "Children's vocabulary acquisition in a foreign language through watching subtitled television programs at home." *Educational Technology Research and Development*, 47(1), 51-60.
- **Krashen, Stephen D.** (1985). *The Input Hypothesis: Issues and Implications*. New York: Longman.
- **Marschark, Marc** (1993). *Psychological Development of Deaf Children*. New York: Oxford University Press.
- **Neuman, Susan B. and Patricia Koskinen** (1992). "Captioned television as comprehensible input: Effects on incidental word learning from context for language minority students." *Reading Research Quarterly* 27(1), 95-106.

- **NDCS** (2003). *Statistics on Childhood Deafness in the UK*. London: The National Deaf Children's Society.
- **NDCS** (2011). *Deaf Children with Additional Needs*. London: The National Deaf Children's Society.
- **Nelley, Elsie and Anne Smith** (2000). *PM Benchmark Kit Teacher's Notes (Progress with Meaning)*. Cheltenham: Nelson Thornes.
- **Ofcom** (1999). *ITC Guidance on Standards for Subtitling*. London: Office of Communications.
www.ofcom.org.uk/static/archive/itc/itc_publications/codes_guidance/standards_for_subtitling/index.asp.html (consulted 14.10.2013)
- **Schlesinger, Hilde S. and Kathryn P. Meadow** (1972). *Sound and Sign: Childhood deafness and mental health*. Berkeley, CA: University of California.
- **Silver, Janet; John Gill; Christopher Sharville; James Slater and Michael Martin** (1998). *A New Font for Digital Television Subtitles. Hearing Concern*.
www.tiresias.org/fonts/screenfont/report_screen.htm (consulted 14.10.2013)
- **Stoloff, Lynn and Dennis Zona** (1982). "Matthew." *American Annals of the Deaf* 123, 452-459.
- **Stuart, Morag ; Masterson, Jackie and Dixon, Maureen**. (2003). *Children's Printed Word Database. Manual and documentation*.
<http://www.essex.ac.uk/psychology/cpwd/documents/CPWD%20manual.pdf>
(consulted 12.12.2013)

Biographies

Soledad Zarate is currently completing a PhD on subtitling for deaf children at University College London. She works as a freelance subtitler and as visiting lecturer, teaching audiovisual translation and media accessibility. She is also a puppeteer and has work for a wide range of theatres in London.



Joseph Eliahoo is a Statistical Consultant at Imperial College London and has compiled statistics for this article. He can be contacted at j.eliahoo@imperial.ac.uk

¹ The frequency of the words selected for Clip 1 is slightly higher (total value = 507) than that of Clip 2 (total value = 459). To compensate for this, Clip 2, unlike Clip 1, includes one word – *wings* – that appears among the first 1,000 words acquired. None of the other words selected appear among the first 1,000 words acquired since the words introduced for the recognition task needed to be new words with some degree of complexity.

² This typeface was adopted following research conducted by Silver et al. (1998) with visually impaired subjects and hearing impaired subjects. The visually impaired people (N = 35; average age = 60) were presented with a sentence printed in 1) Standard AlphaMosaic (used in analogue television), 2) Tiresias (first version), and 3) Times New Roman. The hearing impaired subjects (N = 48; average age 62) were presented with a short video using a later version of the Tiresias Screenfont typeface (with improvements to the kerning) in four sizes: A (30 lines); B (20 lines); C (24 lines) and D (26 lines). The subtitles appeared in white on a black strap at the bottom of the screen. The majority of visually impaired viewers expressed a preference for Tiresias Screenfont. The preferences expressed by the hearing impaired were unfortunately not reported. The research basis of Tiresias Screenfont's legibility claims have been called into question (<http://screenfont.ca/fonts/today/Tiresias>).

³ Screenshots from *Arthur* provided with permission from WGBH Educational Foundation and DHX Media Ltd. Characters (including 'Arthur' and 'D.W.') and underlying materials are the copyright and trademarks of Marc Brown.

⁴ Note that in the broadcast subtitles *that* is emphasised through the use of upper-case as an indicator of intonation. The indication of paralinguistic features has not been the focus of this research, but it is worth mentioning that in this specific case it was not considered necessary to emphasise the word *that* as it does not make the subtitle any clearer and also does not add anything to the content.

⁵ Two small pilot studies were conducted prior to the main pilot study. The purpose of the pilot studies was to help design the children's questionnaires and to assess whether they were able to cope with the task. They helped mainly in improving the design of the questionnaire and they also shed light on logistic issues that needed to be addressed before the main experiment. The questionnaire was the part of the initial material most heavily changed for the main pilot study. The number of questions was reduced from 17 to 13 as some children did not seem to cope well with the length of the questionnaires.

⁶ The Salford Sentence Reading Test is a popular individual test of oral reading for 5 to 10 year olds (Bookbinder *et al.* 2002). The test is performed orally on a one-to-one basis and can take as little as four minutes per pupil. It is ideal for use with less able readers from about age six.

⁷ The PM Benchmark assess students' instructional and independent reading levels using fiction and non-fiction texts ranging progressively from emergent levels to reading age 12 (Nelley and Smith 2000).