
# The Effect of TV Captions on the Comprehension of Non-native Saudi Learners of English

**Mubarak Alkhatnai**
E-mail: S0794005@sms.ed.ac.uk
The University of Edinburgh (United Kingdom)

## Abstract

*This paper investigates the effectiveness of closed captioning in aiding English comprehension of non-native Saudi speakers. Research was carried out in a qualitative manner, and participants were 12 Saudi students pursuing their studies at Indiana University of Pennsylvania, USA (IUP). Participants in the study were asked to compose a narrative after viewing a 5-minute film segment, both with and without captioning. Their responses were then analyzed, and results indicated that while captions may aid one in comprehension, they also tend to limit one's interpretations, reaffirming the nature of written language as an authoritative source of information.*

## Research Question

Modern trends in second language teaching emphasize the importance of language as a tool for communication, where students most effectively learn through meaningful use of language, not by rote memorization. The increased availability and lowered cost of technologies such as television, computers, video players, and video recorders has allowed instructors to more easily bring authentic communication into the L2 classroom. Various researchers have found strong evidence pointing to the effectiveness of multimedia in providing rich, authentic, comprehensible input for students (Khalid, 2001). Multimedia, as its name implies, is not only a visual medium – but an aural one as well, and when combined with traditional print through the use of captions, it becomes an even more valuable asset.

The primary focus of this study is to investigate the effects of closed captioning on Saudi students with English as their second language (L2). Specifically, how does captioning affect the overall comprehension of speakers of English as a second language?

## Literature Review

Of particular relevance to the present study is the work of Paivio (1971), who developed a theory of dual coding connecting imagery (mental or pictorial) with language (oral and written) to explain comprehension. A study performed by Sadowski et al. (1991) supported the dual coding theory in the first language context, and extended it to the bilingual contexts. These studies provided clear evidence regarding the importance of captions in facilitating second language students' comprehension, and concluded that captions improved students' reading ability, and also increased vocabulary (Paivio and Lambert, 1981; Koolstra and Beentjes, 1999).

Also relevant to this study is the work of Markham, Peter, and McCarthy (2001), who examined the effects of Spanish captions, English captions, or no captioning on the L2 comprehension of students in a

university-level intermediate Spanish as a foreign language course. Of the 69 participants in this study three groups were created, each of which viewed the same DVD passage with Spanish, English, or no captioning present respectively. After viewing the passage, participants were asked to provide a written summary of it, and also complete a 10-question multiple-choice quiz. Results revealed that English captioning was a significant aid to students' comprehension; students in the English captioning group performed at a considerably higher level than either the Spanish, or no captioned groups. Additionally, students in the Spanish captioning group performed higher than those with no captions at all, illustrating that there is some pedagogical value in captioned video in the L2 classroom.

In a similar study, Neuman and Koskinen (1991) advocate the use of captions as sources of comprehensible input, and also researched the ability of captions to stimulate incidental word learning for language minority students. Participants consisted of 129 primarily Southeast Asian seventh and eighth grade students, who were 2-3 years below their expected proficiency levels. Students viewed various television segments targeted toward 8 to 12 year olds, and were divided into three different groups. The first group viewed segments with standard closed captioning, while the second group had no captioning at all. For the third group, captioning was not provided on-screen, but in a separate script format. As in the previous study, the captioned group scored higher than non-captioned participants, as well as those with separate scripts. These results suggest that on-screen representations of words aid learners in the acquisition of vocabulary and meaning when coupled with comprehensible input of different types, such as video and sound.

## History of Captioning

Captions refer to subtitles or translations of a video's original audio track. Captions were originally designed so that hearing impaired viewers, or speakers of other languages, could follow along with a film's dialog. Captions are displayed by embedding code into a regular television signal, which is then converted to on-screen text the viewer can read, usually appearing in the bottom portion of the screen. Captions are generally white text against a black background, and placed at either the top or bottom portion of the screen so that they won't obscure any parts of the picture. The National Captioning Institute and ITT Corporation have developed a caption-coding microchip in 1989, which can be installed in any television. In 1990, US Congress passed the Television Decoder Circuitry Act, which mandated that by the middle of 1993 all televisions with displays larger than 13 inches must be capable of displaying closed captioning (Linebarger, 2001). The National Captioning institute (NCI) estimated in 1993 that by the end of 1994 there would be 40 million households with caption-ready sets (National Captioning Institute, 1993). According to the same source, the quantity of captioned programs increased from 400 hours per week to 800 hours a week in 1994.

Because the popularity of captioning is growing, it subsequently can be of great use in language learning. Also, the types of programs that feature captions are usually popular genres, such as news, dramas, documentaries, sitcoms, children's programming, and sports. Carolyn (1994) notes that captioning is important to language learning since, " . . . the level of language used, age appropriateness, sophistication, and overall quality of these programs vary widely" providing viewers with a very unique, multileveled input. For instructors, closed captioning provides "a rich resource and provides new options for instruction" (para. 3).

## Captioning and the Language Classroom

Research has been conducted regarding teachers' uses of captioned video in designing, developing, and implementing a number of different lessons. Koskinen et al. (1991) performed a study involving 45 learning impaired students, where teachers developed supplemental readings using captioned situation comedies, cartoons, and science-fiction films. There were many interesting findings in this study. First, both teachers and students were very satisfied with the lessons; student motivation was high, as was the teachers' level of enthusiasm while designing the lessons. Secondly, teachers reported that captioned TV helped to improve students' vocabulary development. The instructors also suggested that captioning aided the development of other skills, such as predicting the outcome of a plot, character analysis, and sequencing. Finally, teachers reported that students seemed to stay on-task while completing the activities, an indication of high motivation.

Tim Rees (1993) implemented CCTV with Chinese and Japanese students of ESL at the International Language Institute of Massachusetts, where he used captioned TV news programs and situation comedies to improve vocabulary, increase cultural awareness, and encourage class discussion. A key component of his lesson was the transcription of captions with word processors; printouts from transcripts were then used in class for discussion, homework activities, and in-class reading. Captions helped to improve students' listening and comprehension skills, and provided them with a widened understanding of the language's cultural intricacies, while at the same time keeping them up to date with current events.

Todd Ellsworth (1992) also utilized captions while conducting a study at the Benjamin Institute in Mexico. Using captioned TV programs received via satellite, he divided his class into three groups: the first viewed a program without captions, the second with captions, and the third with only audio. By using information gathered from a discussion afterward, Ellsworth designed a lesson on grammar and vocabulary, discovering that students who had captions available used the language with more ease and confidence.

Interestingly, Webb, Vanderplank, and Parks (1994) even suggest using closed captioned children's programs such as "Sesame Street," "Reading Rainbow," and "3-2-1 Contact" with adult ESL learners, and argue that the content, speed of captioning, and vocabulary make these programs suitable for use in an adult ESL classroom. Additionally, they provide some example activities that can be designed around these programs.

## Research Method

Background knowledge plays an important role in the way one interprets information, including life experiences. With that in mind, it seemed important to find participants who shared similar educational backgrounds. The participants in this study were Saudi males who are Ph.D. candidates in the graduate English program at IUP. All of the participants share similar language backgrounds, speaking Arabic as their first language.

This qualitative study attempted to answer the question, "Do TV captions contribute to the comprehension of non-native Saudi speakers?" The researcher was interested in determining if, or to what degree captions affect the English comprehension of Saudi students. In order to carry out this study, 12 participants were asked to watch a 5-minute segment of the film "The Day After Tomorrow," a film about the potentially disastrous effects of global warming on the world. This film, and the segment from it, was chosen because of its non-offensive nature; there was nothing in the clip the researcher believed might have offended any viewers. Participants were divided into two groups (six students each) and asked to

write a short narrative of the segment. The narrative form was chosen over a multiple-choice quiz since it provided participants with more latitude to respond, and allowed the researcher to better compare subtle differences in the respondents' answers. Members of Group A viewed the film with no captioning, while members of Group B viewed the film with captioning. Data was gathered and analyzed over the course of approximately three weeks. Rather than conducting the study in a cold classroom, it took place at one of the participant's homes, creating a relaxing atmosphere where respondents could more easily focus. Also, there was no need to rent or acquire equipment, as is the case when using some classrooms, since many homes already have televisions and DVD players.

## Data Analysis

After data collection was finished, responses from both groups were separately examined. Sentences in each response were contrasted against a detailed written description of the original segment, which took note of any visual, verbal, or written data that the participants may have seen, as well as each element's juxtaposition. Participants' protocols were analyzed to determine the quantity and also which thoughts were identical to the original work, and what new thoughts may have been added that were not originally present. Once this process was completed, the results were analyzed to determine how many identical ideas were repeated, and why certain new ideas seemed to appear. Other questions included whether the new ideas shared any common characteristics, and what kinds of descriptions were most prominent: visual, verbal, or written? Lastly, the spelling accuracy of both groups was analyzed.

## Results

As shown in Table 1, the non-captioned group provided an average of 110.6 words per narrative, while the captioned group gave an average of 83 words per narrative, exemplifying a tendency to be less descriptive in their analyses. Analysis of the data revealed that members of the non-captioned group tended to provide information that could not be directly correlated to the original segment. For example, they were more likely to give personal elaborations, explaining events from their own point of view. Members of the non-captioned group were more likely to express personal sentiments, such as one member whose narration of a scene where a boy praises a female colleague during a school competition was interpreted as flirting.

S= student

| | Number of words per narrative | | | | | | Group Totals | Group Averages |
|---|---|---|---|---|---|---|---|---|
| Captioned group | S1 | S2 | S3 | S4 | S5 | S6 | 498 | 83 |
| | 71 | 101 | 88 | 95 | 62 | 81 | | |
| Non-captioned group | S1 | S2 | S3 | S4 | S5 | S6 | 664 | 110.6 |
| | 124 | 95 | 86 | 115 | 99 | 145 | | |

Table 1: Quantity of ideas recalled

The captioned group in contrast, provided descriptions reflecting a less-personal analysis, relying more on the captions' textual descriptions of events. For example, one participant in the captioned group identified wolves in a cage as dogs, while participants in the non-captioned group believed they were either seals or dogs – two drastically different things. Members of the captioned group tended to focus their attention on

factual items, such as characters' names, and specific numbers; these were less important to the non-captioned group. The captioned group outperformed the non-captioned group in recognizing verbal materials. One specific example of this was one member of the captioned group who was able to transcribe the lyrics to a song that briefly played.

Both groups correctly identified lightning over Washington, DC – however the non-captioned group was keener in recognizing other visual cues; an average of 13.66 out of 18 visual prompts were recognized by the non-captioned group. One member mentioned in his response that he preferred to have the captions off, since reading and following the film was too difficult to do at once, possibly explaining why the non-captioned group excelled at the retention of visual prompts (see Table 2).

Table 2: Replication of visual, verbal, and written ideas

| | Original | Captioned group | Non-captioned group |
|---|---|---|---|
| **Visual** | 18 | Tally<br>S1: 11<br>S2: 9<br>S3: 13<br>S4: 8<br>S5: 15<br>S6: 14 | Tally<br>S1: 14<br>S2: 13<br>S3: 17<br>S4: 9<br>S5: 18<br>S6: 11 |
| Average | | 11.66 | 13.66 |
| **Verbal** | 43 utterances | Tally<br>S1: 33<br>S2: 19<br>S3: 21<br>S4: 9<br>S5: 28<br>S6: 16 | Tally<br>S1: 14<br>S2: 25<br>S3: 20<br>S4: 13<br>S5: 8<br>S6: 11 |
| Average | | 21 | 15.16 |
| **Written** | 128 | Tally<br>S1: 113<br>S2: 68<br>S3: 44<br>S4: 90<br>S5: 109<br>S6: 78 | Tally<br>S1: 74<br>S2: 51<br>S3: 47<br>S4: 73<br>S5: 69<br>S6: 82 |
| Average | | 83.6 | 66 |

The major difference between the two groups is clear when comparing their written scripts. Although the non-captioned group produced a higher quantity of writing than the captioned group (Table 2), they were not as precise as the captioned group in the exact retention of the original script, reaffirming that their attention was not focused on the film's elements as a whole, but primarily on its captioning, thus distracting them from other sensory items.

Table 3: Deviation from the original script

| | Original | Captioned group | Non-captioned group |
|---|---|---|---|
| **Visual** | 18 | Tally of non-existent ideas<br>S1: 2<br>S2: 1<br>S3: 3<br>S4: 1<br>S5: 1<br>S6: 0 | Tally of non-existent ideas<br>S1: 0<br>S2: 1<br>S3: 2<br>S4: 1<br>S5: 0<br>S6: 1 |
| Average | | 1.33 | 0.83 |
| **Verbal** | 43 utterances | Tally of non-existent ideas<br>S1: 4<br>S2: 2<br>S3: 0<br>S4: 2<br>S5: 1<br>S6: 1 | Tally of non-existent ideas<br>S1: 0<br>S2: 3<br>S3: 4<br>S4: 2<br>S5: 1<br>S6: 1 |
| Average | | 1.66 | 1.83 |
| **Written** | 128 | Tally of non-existent ideas<br>S1: 0<br>S2: 1<br>S3: 5<br>S4: 3<br>S5: 0<br>S6: 0 | Tally of non-existent ideas<br>S1: 1<br>S2: 5<br>S3: 3<br>S4: 4<br>S5: 6<br>S6: 2 |
| Average | | 1.5 | 3.5 |

The captioned group recollected an average of 1.33 non-existent visual cues per narrative, while the non-captioned group recalled an average of .83 non-existent cues (Table 3). Verbally, the two groups did not show a significant deviation from the original script, however in the written form the non-captioned group outnumbered the other group in the average of new materials added. The average of the non-captioned group was 3.5 new materials compared to 1.5 new materials for the captioned group.

The non-captioned group's lack of written accuracy does not necessary detract from their overall accuracy, especially considering that their focus was more universal. The full, and natural combination of visual and aural senses allowed for a more thorough analysis of the segment. In other words, it allowed them to become more critical of the film as a whole, not just its script. To say the captioned group's interpretation of the segment is superior because they were able to more faithfully reproduce it textually is false; it is the researcher's belief that they were subconsciously limiting themselves to the text only, in an effort to follow the captioning, and subsequently saw (or missed) a portion of the film that the other group did not.

Table 4: Spelling accuracy

| | Captioned group | Non-captioned group |
|---|---|---|
| **Total number of words** | 498 | 664 |
| **Total number of spelling errors** | 33 | 48 |
| **Percentage of spelling errors** | 6.6% | 7.2% |

Comparatively, the non-captioned group had more spelling errors than the captioned group. Once again, this seems to be a natural result, since they were focused more on the film's written script. The exact spelling of words were likely fresh on the participant's mind, since they had appeared only shortly before transcription. Spelling mistakes not made of scriptural transcriptions could be explained in two ways. First, viewers of the non-captioned segment may have been writing in a stream of consciousness, more eager to express their ideas, and less focused on pedantic accuracy. Secondly, the very presence of captioning may have put some viewers in an entirely different state of mind; that is, a written, finalized state of mind where spelling and grammar errors are unacceptable. While visual and aural elements have no specific rules governing them, textual elements are governed by grammar, and the respondents may have subconsciously been more aware of this fact. The accuracy of spelling reflected in the captioned group, in the researcher's opinion, does not strongly suggest students exposed to captioning will become better spellers, though this is a possibility. It does suggest, however, that it may aid short-term spelling memorization. For more results concerning spelling accuracy see Table 4.

## Discussion

At face value this study seems to reaffirm the findings of others: that captions aid one's comprehension of video. However the researcher is not sure that this is true. What is known for sure from this research is that participants in the captioned group were less likely to go outside the bounds of the original work; they were less likely to be interpretive, which is not necessarily a positive thing. This study seems to reiterate the idea of text being a finalized, unquestionable source of information for many individuals; the text in this scenario limited the perceptions of the viewer to something very finite – though not necessarily more accurate. In that respect I believe one respondent's comment that he found the captions distracting to be very relevant. To move one's attention from the millions of colors, actions, sounds, and other visual and audio intricacies is distracting indeed. The fixed text in this scenario may in fact destroy the work, since it is filtering so many complexities through only one sense.

Furthermore, though one can assume captions probably aid one in acquiring L2 vocabulary, this experiment did not really gauge that effectively, since most participants were already fluent speakers, with high proficiency. Nonetheless, the researcher does not believe that takes any value away from this study since there were still important findings, though from an unexpected viewpoint.

## References

[1] Ellsworth, T. (1992, October*). Integrating subtitled video into your teaching*, "English Teaching Forum".
[2] Khalid, A. (2001). *The effect of multimedia annotation modes on L2 vocabulary acquisition: A comparative study*. Language Learning and Technology, 5(1), 202-232.
[3] Koolstra, C. & Beentjes, J. (1999*) Children's vocabulary acquisition in a foreign language through watching subtitled television at home*. Educational Technology, Research, and Development, 47, 51-60.
[4] Koskinen, P. S., Wilson, R. M., Gambrell, L. B., & Jensema, C.J. (1991). *Captioned video technology and television-based reading instruction*. In S. Clewell (Ed.), Literacy: Issues and practices (pp. 39-47). Bethesda, MD: Yearbook of the state of Maryland International Reading Association.
[5] Markham, P., Peter, L., & McCarthy, T. (2001). *The effects of native language VS. target language captions on foreign language students' DVD video comprehension*. Foreign Language Annals, 34(5), 439-445.

[6] Neuman, S. & Koskinen, P. (1992). *Captioned television as a comprehensible input: Effects of incidental word learning in context for language minority students*. Reading Research Quarterly, 27, 95-100.
[7] Parks, Carolyn (1994). *Closed Captioned TV: A Resource for ESL Literacy Education*. ERIC Digest. Retrieved January 25, 2005 from http://mirror.eschina.bnu.edu.cn/Mirror/ed.gov/www.ed.gov/databases/ERIC_Digests/ed372662.html
[8] Paivio, A. (1971). *Imagery and deep structure in the recall of English nominalization.* Journal of Verbal Learning and Verbal Behavior, 10, 1-12.
[9] Paivio, A. & Lambert, W. (1981). *Dual coding and bilingual memory.* Journal of Verbal Learning and Verbal Behavior, 20, 532-39.
[10] Rees, T. (1993). *"Closed captions in the classroom*. Unpublished manuscript. Northampton, MA: International Language Institute of Massachusetts.
[11] Sadowski, M., Kealy, W., Goetz, E. & Paivio, A. (1991). *A critique of schema theory and a dual coding alternative*. Reading Research Quarterly, 26, 463-84.
The National Captioning Institute, Inc., 1900 Gallows Road, Vienna, VA 22182. (703) 917-7600. (For information about decoders and research studies)
[12] Webb, M., Vanderplank, R., & Parks, C. (1994, April). "*Developing ESL materials for closed-captioned video programs*." Paper presented at the Teachers of English to Speakers of Other Languages annual convention, Baltimore, MD.