UNIVERSITY OF ALBERTA

HOW DOES VIDEO CAPTIONING IMPROVE LISTENING COMPREHENSION?

BY

JACOB B. SCHEFFER

A PROJECT REPORT

SUBMITTED TO THE DEPARTMENT OF EDUCATIONAL PSYCHOLOGY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS OF THE DEGREE OF

MASTER OF EDUCATION

IN

TEACHING ENGLISH AS A SECOND LANGUAGE

DEPARTMENT OF EDUCATIONAL PSYCHOLOGY

EDMONTON, ALBERTA

Spring, 2014

UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR: Jacob B. Scheffer

TITLE OF PROJECT: How Does Video Captioning Improve Listening Comprehension?

DEGREE: Master of Education

YEAR THIS DEGREE GRANTED: 2014

Permission is hereby granted to the Department of Educational Psychology, University of Alberta, to reproduce copies of this project, and to lend or sell such copies for private, scholarly or scientific purposes only.

The author reserves all other publication and other rights in association with the copyright of this project, and, except as herein before provided, neither the project nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.
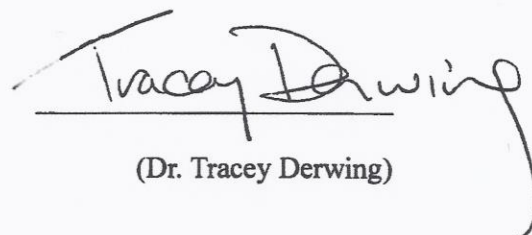
J.B. Scheffer
10823-130 Street NW
Edmonton, Alberta
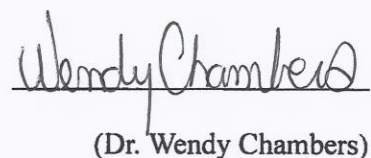T5M 0Z2 Canada

Date: April 22, 2014

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, a Project Report entitled "How Does Video Captioning Improve Listening Comprehension?" submitted by Jacob Scheffer in partial fulfillment of the requirements for the degree of MASTER OF EDUCATION in Teaching English as a Second Language (TESL).

(Dr. Tracey Derwing)

(Dr. Wendy Chambers)

Date: April 22, 2014

Abstract

The question, 'How does video captioning improve listening comprehension (LC)?' is discussed from the perspective of the value captioned video brings to the adult English as a second language (ESL) learner and from how they can be effectively used in the ESL classroom. Listening comprehension is a multifaceted construct made up of and influenced by numerous factors. Captions and subtitles can vary in quality and only through judicious use will they maximize language proficiency for the language learner (LL). The idea that captioned video improves listening comprehension is presented through numerous studies that indicate the benefits and pitfalls of caption use. The major elements that influence listening comprehension when using captioned video include learning strategies and proficiency level, caption type, and video type. Each factor is discussed and the means to manage them is described.

Danan (2004) poses the question 'How can we utilize the undervalued language learning strategies of captioning and subtitling?' There is a large body of research that indicates these tools are of significant benefit to language learners (LLs) (Baltova, 1999; Bianchi & Ciabattoni, 2008; Chang, Tseng & Tseng, 2011; Danan, 2004; Garza, 1991; Grgurović & Hegelheimer, 2007; Markham, 1999; Price, 1983; Robin, 2011; Rubin, 1995; Vandergrift & Goh, 2012; VanderPlank, 1988, 2010, 2013; Winke, Gass, & Sydorenko, 2010). A Google scholar search of captions and English as a second language (ESL) revealed over 2000 papers; and a similar search of captions and English as a foreign language (EFL) revealed over 4000 papers discussing this topic. Sampling these studies indicate that a large majority show captions can be used beneficially.

With such strong evidence in favor of captioning, the current study will suggest practical use of captions in the adult ESL classroom. I will investigate research on the benefits and limitations of captions in the classroom. Specifically, I will identify and recommend how to control those factors that most significantly affect listening comprehension.

Overview

This study comprises a literature review to address the research questions and a sample lesson and instructions for use in the language classroom (see Appendices 2-3). To restrict the information found in the vast number of articles on video captioning, a few criteria have been established:

1) Frequency of citation – this criterion encompasses articles of historic significance (citation frequency in brackets) including, Baltova, 1999 (#58); Bird & Williams, 2002 (#75); Garza, 1991(#174); Guillory, 1999 (#83); Markham, 1999 (#72); Price, 1983 (#52).

2) Currency – articles that are relatively recent additions to the field of study, for example, Danan (2004, 2010); Winke, Gass and Sydorenko (2010); Grgurović and Hegelheimer (2007); Chang, Tseng, and Tseng (2011); Vanderplank (2010, 2013); Vandergrift and Goh (2012).

When faced with a different pedagogical tool, a question any language teacher might have is, 'Does it work? And if it does work, how do I use it?' Research suggests that captioning does facilitate language learning given the right conditions. The questions posed by this paper are an attempt to narrow down the specifics of effective caption use:

1) What influence do the combined modalities of audio/video (AV) and captioning have on listening comprehension?

2) How do video captions improve listening comprehension?

These questions will be addressed by outlining the research that 1) defines the terminology, and 2) considers the interplay of video captions with a variety of other independent variables that affect listening comprehension, including text, interlocutor, task, listener, and process characteristics (Rubin, 1994). Furthermore I will identify research that clarifies how teachers and students can use them to achieve the greatest language gain.

Literature Review

Listening comprehension

Definition

This study on listening comprehension (LC) must begin with a clear understanding of the term. Very commonly, researchers have defined LC in a result-oriented manner as spoken text input that becomes uptake (Danan, 2004; Markham, 1999; Price, 1983; Vanderplank, 2013). Other authors (Rost, 2002; Rubin, 1994; Vandergrift & Goh, 2012) suggest that listening comprehension involves processing and cognition and any definition cannot be limited to results.

Their perspective is that listening comprehension is what occurs in learners from the time of their initial preparation for listening, to the point which spoken text enters the ears, to the point in which the brain processes, and then stores or discards the resulting information.

A universally accepted definition of 'listening comprehension' is not available in the literature nor is that an objective. As mentioned earlier, authors offer slightly different views of the term. This is not a flaw of the studies; rather the nature of listening comprehension makes an all-encompassing definition difficult.

Listening comprehension is a construct, as posited by Vandergrift (2007),

Construct validity in L2 listening assessment is a particular challenge, given the covert nature of listening. Listening processes are difficult to verify empirically, and they interact in complex ways with different types of knowledge. Ultimately, comprehension can only be inferred on the basis of task completion. (p. 203)

Constructs are generally assessed through variables that can be measured. It is helpful to begin with an understanding and identification of the measurable components of listening comprehension addressed by studies. These dependent variables in the studies have become the proxy definition of listening comprehension. This granular view encompasses: learners being able to successfully complete a written test on vocabulary (Markham 1989, 1999; Markham & Peter, 2003; Neuman & Koskinen, 1992); learners being able to comprehend a story and recreate a dialogue (Chiquito, 1995); learners being able to correctly answer reference questions based on the content of material (Garza, 1991); learners being able to correctly identify aural and written words that were used in the material (Markham, 1999; Bird & Williams, 2002); learners being able to measurably improve their ability to understand a second language (L2) (Vandergrift &

Goh, 2012); and learners reflecting their attitudes towards comprehension with captions (Baltova, 1999; Basran & Dilber, 2013; Taylor 2005).

Components

Another means to understand the term 'listening comprehension' is to examine the elements that influence it. Rubin's (1994) review of studies on listening comprehension suggested that it is affected by five major factors including text, interlocutor, task, listener, and process characteristics. These factors are useful because the practitioner can exert some control over them and affect the learners LC (Rost,2013; Vandergrift, 2012). Rubin's (1994) factors are relatively easy for practitioners (myself included) to understand and her research is the most closely aligned with my thought of creating a paper that would be of use to practitioners. In what follows, each variable will be addressed relative to whether it has been studied through video captioning research.

Text characteristics include speech rate, Sandhi (includes assimilation, mutation, contraction, liaison, and elision), stress and rhythm, L1/L2 acoustic differences, syntax, morphology, word order, and discourse markers (Rubin, 1994). These characteristics can affect the level of difficulty of listening material. They should be considered when choosing captioned video because, just as in the selection of any other curricular materials, they should not overwhelm the proficiency level of the students. Some text characteristics can be controlled by the teacher, through limiting the text on-screen to keywords so that students will not be overwhelmed (Guillory, 1999; Hsu, Hwang, Chang, & Chang, 2013; Kikuchi, 1998; Rubin, 1994) . Another text characteristic mentioned in Rubin's (1994) study is text type, for example, news, lecture or dialogue. There are a few studies that suggest that type of video (discussed further in this study) is significant for students' comprehension of video captioned material (Bell,

2003; Bianchi & Ciabattoni, 2008; Markham, 1999). Rubin (1994) also notes that this is a significant variable in listening comprehension overall.

For interlocutor characteristics, Rubin's study cites Markham's (1988) "Gender differences in perceived expertness of the speaker as elements in ESL listening recall" to indicate, that for certain students, the male voice increases listening comprehension. Notably, most of the participants were north Asian where "perhaps female listeners are gradually conditioned to be more attentive to male speakers as a result of gender-related status divisions in the speech community" (Markham,1988, p. 404).

Rubin's 1994 study also brings forward that the type of task has an effect on the language learning outcome. The studies she surveyed noted that multiple choice questions produced "More idea units" (p. 206) than WH questions, that students performed better on tests with "local cues" than "global cues" (p. 206), and that there were "significant main effects for task and semester in the total number of propositions recalled to conclude that tasks affect 1) recall of task-specific propositions, 2) the proportion of macro-propositions to micro-propositions recalled, and 3) the number of distortions" (p. 206). Although it is apparent that tasks can alter listening comprehension, all of the studies surveyed were unable to precisely determine why these task effects occurred.

Listener characteristics are a significant factor impacting listening comprehension, according to Rubin's (1994) review. These include language proficiency, memory, attention, affect, age, gender, learning disability, aptitude, processing skills, background or prior knowledge, biases, motivation, and confidence. Most studies of video captioning take proficiency level into account (Bianchi & Ciabattoni, 2008; Danan, 2004; Markham, 1999; Price, 1983; Vanderplank, 2013). Attention has also been studied extensively through eye movement studies

(d'Ydewalle, 2002) and student interviews (Grgurović & Hegelheimer, 2007). Affect, biases, motivation and confidence have been researched occasionally in video caption studies through posttest interviews (Markham, 1999). Other listener characteristics (i.e., years in the USA, academic major, native language, gender, prior knowledge of the topic, and educational status) have been evaluated and do not seem, as of yet, to have a significant effect on the impact of captions (Markham, 1999).

According to Rubin (1994), process characteristics include cognition, that is, what is happening in the students' heads and strategies, the deliberate process that are utilized by the learners to comprehend. Cognition includes bottom-up processing – looking at the trees in the forest or words and finer segmentations of voice text; top-down processing – seeing the forest or understanding the gist of a voice text; and parallel processing where both top down and bottom up are occurring simultaneously (Rost, 2002; Rubin, 1994; Vandergrift, 2012).

Rubin (1994) further describes that the type of cognitive processing is linked to the proficiency of the LLs. Lower level students tend to use bottom-up processing of captioned video more frequently than higher level students. Bottom-up processing generates a significant short term memory load. Short term memory limitations, also demonstrated to be correlated to proficiency, can create a cognitive overload and can result in comprehension failure for lower level students. In contrast, higher proficiency learners tend to use top-down processing and parallel processing of captions more frequently. The ability to use this processing is directed by students' aptitude, skill in reading captions while listening, prior knowledge of both the language and the topic, schema (the ideas and knowledge of familiar topics) and prior experience with video captions. As stated by Vanderplank (2010), higher proficiency students, in particular those with previous exposure to captions, tend to gain the most listening comprehension from video

captions. Cognitive processing ability must be taken into consideration when determining the

level of captioned video for the LLs in question. Content can also be altered to take into account

student processing speed by using pre-listening activities (e.g., scaffolding, strategy training, and

schema building).

As a final component of listening comprehension, Rubin (1994) describes listening

strategies for improving listening comprehension. In particular, she indicates strategies that are

deliberately taught, learned, and used increase listening comprehension. Two types of strategies

mentioned by Rubin (1994) are: metacognitive: planning, monitoring, and evaluating

comprehension; and cognitive: learning how to store and retrieve information. This study will

discuss developing students' strategy use to aid listening comprehension from captioned video.

Some of Rubin's (1994) factors are addressed within my paper. It is limited to factors in

which pedagogical control has been demonstrated. Before proceeding further with describing

which elements may be controlled, there is value in clarifying that listening comprehension and

reading comprehension are interrelated but not mutually exclusive.

Listening comprehension and reading comprehension

A question arises as to whether listening comprehension improvement from captions is

actually general L2 acquisition improvement from reading comprehension (Rost, 2002). There is

a range of positions in this discussion, with claims that captions both hamper and assist listening

comprehension.

It has been suggested that a gain in comprehension is purely a gain in reading

comprehension and that captions actually interfere with listening comprehension (Danan, 2004).

Taylor (2005) makes this argument in a study of Spanish LLs. He describes students with only

one year of Spanish study did not score as well as participants with three years of Spanish study

when viewing a video with captions. However, Vanderplank (2010) is critical of Taylor (2005) and advances that it:

> …perhaps unwittingly, appears to confirm that unless video material is
>
> carefully selected and graded, together with judicious editing of captions,
>
> attempting to use captioned videos with low-level learners is likely to be
>
> a waste of time and may well produce a negative response from students.
>
> (p. 14)

In a Japanese study by Kamei, Setsuko, Hirose, and Keiko (1994) cited by Kikuchi (1998), students who utilized AV with captions performed no better than students who viewed captioned video without sound. In fact, the students stated that they generally ignored the sound and focused only on reading. A question here arises as to whether strategic training in caption use might have been useful (Danan 2004; Vanderplank, 2010).

Other studies describe how cognitive load factors could create a situation where an additional channel (i.e., captions added to video and sound) could overload cognition and actually hamper language intake and uptake overall. This is the position taken in studies where lower proficiency students perform poorly when captions are added to AV input. (Bianchi & Ciabattoni, 2008). These authors and others (Vanderplank, 2010) acknowledge the small sample size of this study. The small sample reduces the statistical reliability of measurable increases. Larger samples have shown that there is a gain for lower proficiency students (Markham, 1999; Price, 1983).

There is a significant body of research on 'Dual Coding Theory' (Chang, Tseng, & Tseng, 2011; Danan, 1992, 2004; Mayer, 2005) that finds that reading comprehension synergistically aids listening comprehension. Listening comprehension is considered the most difficult of the

skills to master (Vandergrift, 2012). Comparatively, reading comprehension is often mastered first (Rost, 2002). Using reading comprehension through captions provides a means to improve listening comprehension. As Rost (2002) points out "Because there are few reliable markers in the speech code for word boundaries even a fluent listener may require one or two seconds to recognize words in the speech stream" (p. 96). He also suggests that reading helps word boundary segmentation so that bottom-up processing of individual words in a sound stream can occur. This is often the experience expressed by students who are familiar with captions in English television and movies (VanderPlank, 2010).

Perhaps the most significant research in this arena is the Bird and Williams (2002) study of listening comprehension, which measured sound recognition from captioned video. Participants viewed either captioned or non-captioned material. Their study quite clearly indicated that listening comprehension improves from captioned material.

Additionally, large studies, with a hundred or more participants sample sizes (Kikuchi 2006, Price 1983; Mitterer & McQueen, 2009) indicate positive listening comprehension results from captioned material, regardless of proficiency. This may in part be the positive effect of larger sample sizes enhancing the statistical significance of small improvements found with lower proficiency students.

Finally, Danan (2004) and Vanderplank (1988) indicate that listening comprehension increases over time with continuous exposure to captioned material. The advantage of exposure points out that Taylor's (2005) study was correct in arguing that caption use should be taught to even the lowest proficiency students who generally do not initially show listening improvement.

What is the conclusion then? It is safe to say that research indicates reading captions can effectively improve listening comprehension. More specifically, captions assist with word

boundary recognition, bottom up processing, and synergistically aid listening. If it is done frequently, using captions is a skill that improves over time. Now that a definition of listening comprehension has been addressed, it is time to define what captions are.

Captions

Definition

The terms 'captions' and 'subtitles' are used somewhat interchangeably. To avoid confusion, however, the terms need definition and the distinction needs clarification. 'Captioning' is the use of same language text on a video screen concurrent with the audio of spoken text. It was initially developed as a means to provide input to the hearing impaired for television (Price, 1983). Generally, captioning includes text of the spoken audio and descriptive text describing other sounds heard concurrently (for example, "audience clapping", "phone ringing") (Livingston, 2013). Captions are commonly found in television and web video. Captions are regulated in Canada (Canadian Association of Broadcasters, 2012) and the USA, and "must match the spoken words in the dialogue and convey background noises and other sounds to the fullest extent possible" (Federal Communications Commission, 2014).

It can be observed that, even though captions are regulated for television and some media, there seems to be some free license with the term. A good example of this is YouTube[©], where, although the CC symbol (closed caption) is used, it appears that inadequate voice recognition software is being used to convert the spoken text to written text: Here is an example of poor quality captioning:

The spoken text was: "it could be any dog and in fact many dogs play this role"

https://www.youtube.com/watch?v=UxqRw-bo84I

Another consideration with captions is whether they are have been prepared for live video (sports or new broadcasts are good examples) versus prerecorded television or video. Captions for live video generally lag the spoken text because time is required to transfer the spoken text to captions on the screen. Pre-recorded video generally has no time lag between the spoken text and the captions.

'Subtitles' are usually in the viewer's L1 or other language text on screen concurrent with the L2 spoken audio. The text may be paraphrased so that the text on the screen matches the scene (Livingston, 2013). Subtitles are not regulated and are most commonly used in movies.

A somewhat unfortunate side effect of subtitling's lack of regulation (Livingston, 2013; Cintas, 2005) is the uncertain quality of subtitles (Cintas, 2005). This is compounded by low cost subtitling software that has created significant competition in the subtitling market and driven down quality. One of the unfortunate side effects of some of the subtitling available today is incorrect and sometimes humorous translations, such as, "Mr. Bush becoming literally el señor Arbusto in Spanish" (Cintas, 2005, p. 2).

For the purpose of this study, the words 'subtitling' and 'captioning' will be used. Unless noted otherwise, 'Subtitling' will refer to use of the students' L1 for the written text. It will be assumed to be on-screen written text in the students' language while the spoken text is in the L2. 'Captioning' will be used where the language of the written text is in the students' L2, the term.

With captions and listening comprehension defined, the next section will address some of the control variables of LC. These include: learner characteristics including, age, language proficiency, motivation, and process characteristics; the type of task; and the nature of the captioned video (captions and type). Teachers will need to be aware of and need to manage these when using captioned video.

Variables affecting listening comprehension and the use of AV with captions

Listener

The listener is perhaps the most complex variable in the studies on listening comprehension. One study (Markham 1999) lists nine characteristics including: years in the USA, age, academic major, gender, education level, native language, TOEFL score, university placement score and prior knowledge. Rost's (2002) text on listening identifies ten listener variables: attention, affect, age, gender, learning disability, aptitude, processing skills, background, biases, motivation, and confidence. The small overlap between these lists is noteworthy. Many of the attributes for example, years in the USA, academic major, native language, gender, prior knowledge of the topic, and educational status seem to have little effect on listening comprehension (Markham, 1999). Among the attributes thought to have an effect are age, proficiency, motivation, and process. Each is discussed below.

Age

My study focuses on adult learners so clarification on the effect of age is necessary. Early studies in the use of captioning indicate its value for adults:

> …CCTV[1] has great potential for teaching adult ESL students. Captioning transforms the seductive medium of television into a literacy and language learning tool and helps introduce newcomers to an important conveyor of culture and information. Smith warns, however, that care must be taken in identifying suitable programming. Students approach programs with varying degrees of linguistic proficiency and familiarity with the cultural contexts involved. (Spanos & Smith, 1990, p. 5)

Spanos and Smith indicated that adult learners and younger learners develop similar behaviors and strategies in caption use. These include writing down new vocabulary, oral repetition, high levels of attention and post treatment use of vocabulary. In contrast, in a study of the Dutch language (d'Ydewalle, 2002) found that adults treat subtitles differently than younger people:

> Elderly people complain more about subtitles than other age groups. d'Ydewalle, Warlop, and Van Rensbergen (1989), using again eye-movement recordings, found that with longer subtitles, younger people looked longer at the subtitle than the older people. As younger people read faster than older people and therefore finish reading earlier, younger people start re-reading the subtitles and therefore, linger longer in the subtitles. The age-related difference does not occur with shorter subtitles because in that case nobody has extra time available due to their shorter presentations. Watching subtitled programs requires to integrate

---

[1] CCTV – Closed Caption Television

continuously the information from the image, subtitles, and sound, and older

subjects are considerably slower to integrate; therefore, older subjects return to

the image as quickly as possible after a first reading of the subtitles. A number of

detailed analyses of the data as well as the screening of the video recordings

confirmed the above explanation. (p. 61)

This difference in reading speed related to age indicates the need to limit the amount of

information in subtitles for adults, but clearly, adults can acquire language from subtitles through

processes similar to those used by younger people (d'Ydewalle, 2002). Older people just require

more time.

Proficiency

Language proficiency is a consistent control variable in research on caption use. Some

studies indicate that lower proficiency students do not gain a significant advantage from

captioned material (Bianchi & Ciabattoni, 2008; Taylor, 2005). These studies are heavily

counterweighted by studies with much larger sample sizes that indicate captioning may benefit

even the lowest level students (Kothari, Takeda, Joshi, & Pandey, 2002; Kothari, Pandey, &

Chudgar, 2004) and benefit students at all levels (Price, 1983; Vanderplank, 2013).

Many studies indicate that captioning can be beneficial for lower proficiency students

when captions and the material are suitable to the LLs' level. Kothari et al. (2002) use the term

"neo-literate" to describe students who are just at the beginning of literacy development. Kothari

et al. demonstrated that captioning popular Indian songs improved these students listening

comprehension. Guillory's (1999) and Kikuchi's (2003) research focused on keyword captions

for lower proficiency learners. Spanos and Smith (1990) noted that videos must be appropriate to

the age and proficiency of the students.

Motivation

Motivation has been shown to have an influence on listening comprehension. As is described below interrelated criteria are identified in the literature as positively motivating LC.

First, teachers can increase motivation through the use of pedagogical techniques and interesting teaching and learning materials (Dörnyei, 2001). Teachers can enhance L2 student motivation: "Without sufficient motivation, however, even the brightest learners are unlikely to persist long enough to attain any really useful language" (Dörnyei, 2001, p. 5). "So, whose responsibility is it to motivate learners? My guess is that it is every teacher who thinks of the long-term development of his/her students" (Dörnyei, 2001, p. 27). Dörnyei (2001) also points out that there are "*course specific motivational components* (related to the syllabus, the teaching materials, the teaching method and the learning tasks)" (p. 19). Film is teaching material that Baltova (1991) recommends is "an entertaining authentic source" that "can have considerable educational value" and "is consistent with a growing consensus among educators today that the combination of text, picture, and sound (especially when an interactive element is added) can promote learning" (p. 151).

Secondly, Baltova (1999) suggests that film can be made more useful by the addition of captions. Baltova's study points out the positive potential that video has as a motivational activity, but there is a risk of video being perceived as entertainment and not educational. Her paper advises that the addition of captions helps transition video from being purely entertainment to an educational tool.

Thirdly, captions motivate students primarily because they increase listening comprehension (Garza, 1998; Basaran & Dilber, 2013; Danan, 2004; Kuppens, 2010). . Research that measured perception indicated that the students preferred captions to no captions. They

believed it enhanced their learning (Baltova, 1999; Danan, 2004; Kikuchi, 2006; Stewart & Pertusa, 2004; Vanderplank, 2013).

Finally, the research into the continual use of captions demonstrated LC increased in conjunction with learners' attitudes towards captioning. (Basaran & Dilber, 2013; Garza, 1991; Danan, 2004; Kuppens, 2010) Two studies indicated that familiarity with captions seemed to impact their effectiveness. In Garza's (1991) research, participants studying Russian did not perform as well with captions as participants studying English with captions. This was due in part to the increased availability of and familiarity with English captioned films and resulted in higher listening comprehension scores. In another study Turkish students gained significantly from captioned film (Basaran & Dilber, 2013). Students stated that they frequently used captioned television and felt quite comfortable watching a film using English captions. Danan (2004) and Kuppens (2010) note that Europeans, familiar with English film and television, use captions extensively to learn language.

Motivation, although it can be difficult to quantify because it is a construct, was assessed in the research presented above through participant interviews. In summary, the research indicates, students liked captioned material, they believed it helped their listening comprehension, they perceived it made video less entertainment focused and more education focused, and they thought that ongoing exposure to captions helped them more than single or infrequent exposure.

Motivation and processing have some similar characteristics. They are both internal to the student, have an impact on listening comprehension, are affected by captions and are elements of LC over which teachers have some control.

Process

Exploration of cognitive processes aids in understanding why captions are useful. Danan (2004) eloquently discusses why this occurs using studies she reviewed or authored. Specifically she presents the idea that listening comprehension improves if material is presented both visually (images and text) and aurally. The brain accepts the information through auditory and visual channels separately but then synergistically combines them to achieve full meaning comprehension. This finding is supported by studies from Chang, Tseng, and Tseng (2011); Danan (1992); and Mayer (2005).

Cognitive processing can be enhanced by using specific techniques. For example, as was mentioned earlier, repeated use of captions augments the skill of students in utilizing captions for LC (Vanderplank, 2013). Additionally, research by Guillory (1999), Kikuchi (1998), and Kothari et al. (2002, 2004) specifically comment that pedagogically reducing the text on screen to simpler or keyword captions reduces the risk of cognitive overload and develop LC.

Vandergrift and Goh (2012) devote a significant portion of their text on metacognitive processes to 'learning to learn to listen'. They describe the value of acquiring listening strategies (see Appendix 1) including the metacognitive processes of planning, monitoring, evaluating, and problem solving to enhance listening comprehension. This is very similar to Rost's (2002) focus in his listening text where he recommends: "Listening strategies include predicting, inferring, listening for the main idea, and listening for details" (p. 111)

In conjunction with teaching strategies, Vandergrift and Goh (2012) point out specific teaching instructions that can guide caption use in the classroom:

> • Use materials with captions, to reinforce and confirm understanding of
> an aural text: for example, with a repeat listen. Captions can draw

attention to the difference between what they hear and the written form of the same message. This helps direct attention to gaps in understanding during repeat listens.

• Captions can also help learners develop word segmentation skills and gain insight into their comprehension errors. (p. 235)

Pedagogically, teachers have some control over motivation and can enhance student processing by teaching and utilizing strategies. Perhaps the most significant aspect of what a teacher can do in the classroom is to provide the most appropriate tasks for the specific outcome desired. I will now focus attention on tasks.

Task

Tasks, the actual activities that a teacher uses for a class, are generally well within the control of the teacher. Whether the tasks consist of actual activities, evaluations, student presentations, assigned group work, Q&A sessions, or other exercises the teacher typically manages these through the lesson plan prepared for the class. In the research, covered by my paper, there are examples of task management using captioned video.

Although not directly related to captioned video, one interesting example of task sequence to promote LC was a study conducted with students of Russian at the American Defense Language Institute. The students were deliberately kept from oral language production for a four week period (Rubin, 1994): They were allowed written language production only. The result was a significant improvement in listening comprehension when compared to a control group (Rubin, 1994). There is some correlation between tasks and LC, but more research needs

to be done in order to successfully apply pedagogy (Rubin, 1994). In a similar manner, Danan's (2004) study indicated that there was more LC benefit to presenting captioned video after initially presenting the same video uncaptioned. In both of these studies, the task sequence made a difference in LC.

A second example regarding task selection is that the use of authentic material for classroom tasks has become part of current pedagogy (Nunan, 2004). Nunan also states that authentic material is one of the cornerstones of Task Based Learning and Teaching (TBLT). Students viewing and listening to captioned video is an authentic task (Baltova, 1999; Kikuchi, 2002) that can help improve LC.

A third example, as was discussed earlier with motivation and process, is repeated use of captioned video enhances LC more than single use, "Exposing students only to one movie does not seem to be sufficient to see the effects of a captioned movie on listening comprehension. It can be said that exposing students to at least more than one movie is necessary to confirm the positive effects." (Kikuchi, 2002, p. 177). The ongoing use of captioned video as a task is a useful means to enhance listening comprehension.

In summary, tasks are an element of video with captions that can be managed by teachers. Some guidelines, indicated by the research above, for improving LC include, using task sequencing as a means to scaffold and build listening comprehension, using captioned video as an authentic task, and using repetition of captioned video as a means to increase captioning effectiveness. Task control is one way that teachers can use captioned video effectively, another is management of the captions themselves. In what follows management of captioned video will be discussed.

Text

Captions

From a pedagogical perspective, controlling the captions in video is the area in which teachers have the greatest ability to manipulate the input for LC when compared to other features covered thus far. There are three means to control input described in the research included in my paper. The first is to manipulate the captions, the second is to provide different formats of the onscreen text to match the proficiency of the learners and the third is to select, based on student proficiency, previously captioned or subtitled material.

For the first form of input control, Markham (1999) demonstrated creating precise captions that reflect the aural spoken word has been shown to improve listening comprehension for intermediate and higher proficiency LL's. In his study Markham described that creating captions that were an exact match to the spoken text was time consuming and difficult. Although this represented a significant effort on the part of the researchers (and teachers that might follow this example), Markham found that "the availability of captions did significantly improved the ESL students' listening ability to recognize words on the video-tapes that also appeared on the subsequent listening-only multiple-choice tests" (p. 324).

Another example of text manipulation is Guillory's (1999) study which compared proficiency level to the use of limited or keyword captions. Citing processing theories as a basis for the result, she discovered that lower proficiency students were better able to attend to the auditory channel if the text on screen was limited to keywords. Two hundred American participants in French classes participated in the experiment. She found that the listeners consistently attempted to read every word on screen. If the captioned video is too advanced, that is, beyond their proficiency level, learners may have difficulty catching up with fast dialogues

and comprehending vocabulary. Selected keyword captions let the learners focus more on listening than reading, which allowed them to comprehend and access meaning without having to read every word.

Results similar to Guillory (1999) were demonstrated by Hsu, Hwang, Chang, and Chang's (2013) research showing keyword L1 and L2 captions on mobile devices were just as effective as full text L1 and L2 captions for elementary school children in Taiwan. They assessed both vocabulary acquisition and listening comprehension over a four week period. The authors found that there were statistically significant increases in vocabulary acquisition between the control group (non-captioned) and the experimental groups (captioned). Listening comprehension gains were smaller between the groups and did not achieve statistical significance. The authors attribute that to the elementary students lack of English vocabulary.

Another means to control input is to alter the format of the text presented to the learner. In multimedia settings, captions and transcripts show similar advantages. In their research, Grgurović and Hegelheimer (2007) compared the effectiveness of captions, transcripts, and other help supports for listening comprehension. They also measured frequency of use and student perceptions of each of these supports. While students were viewing the material, they were provided with three onscreen options: captions, transcripts, and a dictionary. Students at the intermediate level were most likely to use no help, followed by captions and then transcripts. Other supports (e.g., dictionaries) were not used at all. Students used and preferred captions over transcripts. The authors suggest that this may in part be due to the prevalence of and familiarity with captions in all video and multimedia environments. Although Grgurović and Hegelheimer's (2007) study revealed many interesting effects related to visual and aural text, the most

significant, for my paper's purposes, is the conclusion that transcripts and subtitling demonstrate similar gains in LC.

Two conclusions can be drawn from controlling captions to benefit listening comprehension. First, as was discovered by Guillory (1999); Markham (1999); and Hsu, et al. (2013), captions should be manipulated to match the proficiency of the LL. Higher proficiency learners gain the most from exact captions and low through high proficiency learners benefit from key word captions. Secondly, as was noted by Grgurović and Hegelheimer (2007), transcripts and captions show equal benefit for LC.

The advantage of keyword captions is that it takes less time than creating onscreen full text captioning. For videos that have actual transcripts readily available, the process is further simplified because it becomes a matter of identifying keywords and providing them as transripts or as is discussed later in this section, applying the keyword captions to the video. Alternatively, another option is to consider videos that are already prepared with captions or subtitling. It is to this topic I will now turn.

Subtitling (L1 and L2) has been a standard feature of movies since the early 1900s (Ivarsson, 2004). Subtitles are currently available in non-L1 theatrical movies and alternate medium (TV, DVD, computer, streamed, etc.) movies as well as TV shows. Subtitles are readily available for teachers to use in these medium and have been shown to generate LC gains (Vanderplank, 2013). A question arises as to the advantage of using subtitles in the L1 with L2 spoken text versus using subtitles in the L2 with L2 spoken text. Several interesting studies have indicated that subtitling in either the L1 or L2 generate similar increases in comprehension (Obari & Hiroyuki, 1996; Mitterer & McQueen, 2009; Setsuko, 1991; Vanderplank, 2013).

In contrast, but in a very specific situation, Mitterer and McQueen's (2009) study of 160 Dutch participants, fluent in English, demonstrated that even though the videos contained unfamiliar regional Australian and Scottish dialects, students were better able to recognize and utilize the language when English subtitles were used over Dutch subtitles. Dutch subtitles were more effective than no subtitles.

In general it is clear that subtitling available in videos has a LC benefit that teachers can utilize for their students (Vanderplank, 2013). As described in studies cited by Kikuchi (1998), the gains from either L1 or L2 subtitling are similar, and that English subtitles are superior for listening comprehension for very high proficiency learners (Mitterer & McQueen, 2009).

Even though L1 and L2 subtitling may show similar benefit for most proficiency levels, from a practical perspective L1 subtitling may not be useful in an ESL classroom where the teacher does not know the students' L1s. Given the subtitling quality issues discussed earlier, translation may be incorrect and the teacher may unwittingly expose the student to errors. Secondly, in classes where there is more than one L1, use of L1 subtitling maybe logistically difficult (impossible?) and will only benefit those able to read the L1 captions.

Managing caption and subtitle quality

Although it has been demonstrated that subtitles and captions do have an LC benefit, caution must be used in the selection of material. As described earlier (see Captions: Definition) captioning is regulated in Canada and the USA. Captions should be an accurate reflection of spoken text, but as was noted with the YouTube © example, they can also vary in quality and they can sometimes lag the actual spoken text. Subtitles, on the other hand, are not regulated. They may include paraphrase and may not accurately reflect the spoken text. Cintas (2005) describes a deterioration in subtitle quality because the cost of subtitling software has declined

and competition in an unregulated atmosphere has reduced quality. So, aside from carefully reviewing caption and subtitle quality prior to classroom use (see Appendix 4), what other options do teachers have for managing this quality issue?

With the advent of inexpensive subtitling tools, teachers can now create quality captions. Freeware such as Camstudio (CamStudio Group, 1998) and Screencast-O-Matic (FFmpeg, 2014) for onscreen recording of video and Windows Movie Maker (Microsoft Corporation, 2011) or iMovie (Apple Inc., 2010) for captioning are simple to use and make captioning capability universally available (see Appendix 3). Krajka (2013) presents simple online tools for captioning for example, Amara (Participatory Culture Foundation, 2014) and Dotsub (Dotsub.com, 2014) are also simple to use and can use online video (YouTube ©, Twitter © and Facebook ©) (Appendix 4). He also specifies two online tools that can capture and download online video: YouTube downloader and Keepvid.com.

Having the ability to control captions and subtitling is a significant advantage for teachers attempting to improve students' LC. Likewise, control over the type of video has particular benefits to LC.

Type

Some types of captioned video are more beneficial than others for LC. Type was described earlier in the section on LC Components and includes genre (Garza, 1991) such as dramatic feature film, light comedy feature film, news/documentary, animated feature and music video. Type also encompasses consideration of other factors including proficiency level (Spanos & Smith, 1990), whether the video is abstract and/or has low semantic correlation (correlation between visual imagery and the text) (Bianchi & Ciabattoni, 2008) and authenticity (Basaran & Dilber, 2013).

For the first component of type, genre, Garza (1991) and Bell (2003) describe its effects on listening comprehension. Garza (1991) compared dramatic feature films, light comedy feature films, news/documentary, animated features and music videos. In his study all of the captioned video genres showed a significant gain in LC over non-captioned video of the same genre. Significantly, among genres, the largest improvement in comprehension was achieved with the captioned music video over the uncaptioned music video, likely because of the confusion created by the music as extra aural material. The captions clarified the aural text.

Bell (2003), in his study of the use of TV news, although not specifically focused on captions, made some interesting comments on the use of this genre for listening comprehension. The author recommends that TV news can be used as a learning tool if it is appropriate:

> …, particular formal features of TV news can have important effects on student comprehension. The trends in American network news towards dramatic framing and the news as infotainment make network news more accessible to second LLs. The more visual and varied style of American TV news, the tendency to more narrative texts, and the tendency towards more perspectival reporting aids student listening comprehension. At the same time, raising students' consciousness of these formal features of TV news in both the target and student culture could have beneficial results for both listening comprehension in particular and in media literacy in general. (p. 15)

Bell also observed that to be effective, TV news requires appropriate scaffolding to build schemata prior to presentation and must have authentic language.

Authenticity is another aspect of captioned video type. A common definition of authentic material is "any material which has not been specifically produced for the purposes of language teaching" (Nunan, 1989, p. 54).

In terms of authenticity, Basaran and Dilber (2013) in an EFL context recommend:

Consequently, captions can be regarded as language learning sources that are full of authentic language materials for learners, making them a plausible partial solution to the lack of opportunities to communicate with native speakers. In ESL contexts this capability is available to those students that do not have the opportunity to connect with native speakers. (p. 93)

A third aspect of video type is level. It is a criterion that should be taken into account when captioned video is selected. This aspect has many sub-elements including grammar, vocabulary, types of discourse, but simply put, Krashen's (1981) 'Comprehensible Input Hypothesis' recommends that the proficiency of the student must be considered when selecting input. In this case, the captioned video is the input. Additionally, the need to consider appropriateness comes into play when using captioned video. Spanos and Smith (1990) indicated that care must be taken in selecting the type of material presented to adults stating, "Sesame Street could be insulting to adult learners" (p. 3).

Semantic correlation is an element related to level. Semantic correlation is a measure of how well the text (spoken and captioned) relates directly to the onscreen imagery. Bianchi and Ciabattoni's (2008) study used an interesting technique to determine how semantic correlation was significant when captioning was used. In this study the authors compared results between viewing captioned versions of *Fantasia* and a *Harry Potter* movie. Participants had not seen the movies prior to the study. *Fantasia* is an animated video dominated by fantasy imagery accompanied by classical music and overlaid with a narrated story. *Harry Potter* is a story of the interpersonal interactions among a group of young students and their professors at a witchcraft academy. Although neither reflects real life, *Harry Potter* has a high level of semantic correlation

between the onscreen imagery and the spoken text, whereas *Fantasia* does not. The authors found consistently higher listening comprehension in L2 students for the *Harry Potter* movie (Bianchi & Ciabattoni, 2008).

In summary, many variables affecting listening comprehension and captioned video can be manipulated by the teacher. Listener characteristics, although they cannot be directly altered, can be managed appropriately. For example, age, proficiency and motivation factors can be taken into account by appropriate captioned material selection; and motivation can be increased by using captions judiciously and repeatedly. Student processes can be enhanced through teaching of cognitive and listening strategies. Using unaltered captioned video is an authentic task, that if used repeatedly, can enhance LC and, where students need to receive input to match their needs, captioned video it can be pedagogically altered.

Discussion

Language and the adult ESL learner

The emphasis of this study is on adult ESL learners, so the question arises as to the applicability of the research cited. In brief, there are few studies that actually address this learner type, but those that do (e.g., Spanos & Smith, 1990; Park, 1994) generally cite the studies reviewed here as research evidence for caption use. Additionally, student age was a feature in the research. The research demonstrated that all LLs, regardless of age, showed similar patterns of behavior and similar results, but the older adult learners required more time to complete reading the captions.

The question of language as a factor is also considered. In the studies surveyed here, English in both ESL and EFL contexts dominated followed by Spanish, French, and Russian contexts. Using results as a measure of generalizability, most studies indicated an overall

improvement in listening comprehension as a result of the use of captions regardless of language. Those that mentioned language at all mentioned it in light of the relative availability of captioned or subtitled material in that language or mentioned orthographic differences in written text [English and Russian] (Garza, 1991).

Understanding listening comprehension

Understanding listening comprehension is a major focus of this paper. It is a multifaceted construct that varies in definition from one study to the next. The construct is comprised of measures used to assess it, processes that are engaged in achieving it, and features that influence it.

The influence combined modalities of audio/video (AV) and captioning have on listening comprehension.

The first research question was addressed by bringing forward research that indicates some specific influences that captions have on listening comprehension. In general, most studies indicate that captioned video has a significant positive influence on listening comprehension. First captions assist with word boundary segmentation so that the listener can comprehend words in a voice text. Secondly there seems to be a synergistic effect from combining listening and reading that aids the learners in developing meaning from a voice text. Thirdly, caption use is a learned behavior that can be enhanced by teaching listening and cognition strategies and finally there seems to be a positive cycle in use of captioned material in which the beneficial effects on listening comprehension increase with more continual use.

Ideal conditions for improving listening comprehension through captioned audio-visual

The second research question regarding how video captions improving listening comprehension was addressed comprehensively in Rubin (1994); Hoven (1999); Rost (2002);

and Vandergrift and Goh (2012). My study identified research covering primary variables that influence listening comprehension. These include include captions, proficiency, motivation, strategies, perception and type of video. Other factors including age, education type, language studied, learner type, intelligence, cultural background, and disability have either not been studied in depth or appear to have no influence on listening comprehension. Ideally, then, in addressing the second research question, the features that are known to have influence could be effectively managed to provide the greatest gains possible in listening comprehension for the LL. They are summarized in what follows.

In a learner-centric model, proficiency is an important consideration in the use of captioned video as a pedagogical tool. The studies have indicated that intermediate and higher level proficiency students seem to learn the most from captioned video. Lower proficiency students can also learn from captioned video, but require pedagogically controlled material to make gains. Material too far above their level or not semantically correlated to the visuals is detrimental.

Pedagogical control, by creating keyword specific captions, has been shown to be beneficial for all levels of LLs and for lower proficiency learners in particular. Precise captions in the language of the audio have been shown to generate the greatest increase for intermediate and higher level students. Care must be taken when using authentic subtitled (same language or otherwise) material because subtitles do not have the same regulated rigour applied to most captions.

Listening strategies should be explicitly taught to achieve the greatest increase in LC. If the general northern European (Dutch, and Scandinavian in particular) experience is an indication, learners gain more from repeated use of captioned material reference.

The final major feature in improving listening comprehension through captioned video is video type. Although this is somewhat less quantifiable because it should be driven by student characteristics and interest, there are some pedagogical considerations. Semantic correlation is a major issue here, the voice (aural and written) should correlate with the visuals on screen to truly enhance listening.

Conclusion

This study addressed the use of captioned videos for improving listening comprehension. Listening comprehension improvement in L2 learners has been demonstrated through the numerous studies of the effects of captioned video. Listening comprehension through captions is not just improved reading comprehension. Use of captioned video in the adult ESL classroom can be successful if major factors influencing listening comprehension are controlled such as matching the captions to the proficiency level of the students, ensuring the appropriate type of video is chosen, managing the match between the captions and the spoken text, teaching students listening strategies.

Even though there are many studies that demonstrate the effectiveness of caption use for listening comprehension, more research could be carried out. Specifically, factors suggested by Rubin (1994) including, text, interlocutor, task, listener, and process; suggested by Rost (2002) including processes and cognition; and suggested by Vandergrift (2012) including process and metacognition, have not all been assessed in depth for their impact and whether or not caption use is influenced by these elements. Going forward it seems that more studies, preferably with larger sample sizes to increase statistical reliability, might be beneficial.

A few examples are expansion of studies on semantic correlation related to video type. Although it seems that there is some measure of improved listening comprehension with

captioned video that has higher semantic correlation, the samples for these studies and their breadth seem small. This study would aid practitioners in having another tool in material selection.

More studies into the effect of video type and authenticity would also be beneficial. For example does captioned authentic or does captioned pedagogical material aid more with listening comprehension? This knowledge could aid practitioners in material selection and in the effort made to create material.

Finally, studies that are focused specifically on the adult ESL classroom would be useful to determine if the results of this literature review are indeed generalizable to that context.

References

Baltova, I. (1999). *The effect of subtitled and staged video input on the learning and retention of content and vocabulary in a second language* (Unpublished doctoral dissertation). Toronto: Graduate Department of Curriculum, Teaching and Learning Ontario Institute for Studies in Education of the University of Toronto. Retrieved from http://www.collectionscanada.gc.ca/obj/s4/f2/dsk3/ftp04/nq41096.pdf

Basaran, S., & Dilber, E. (2013). Effects of captioned TV shows on language learner motivation and perception. *Respectus Philologicus, 23*(28), 83-96.

Bell, D. M. (2003). TV news in the EFL/ESL classroom: Criteria for selection. *The Electronic Journal for English as a Second Language, 7*(3). Retrieved from http://tesl-ej.org/ej27/a2.html

Bianchi, F., & Ciabattoni, T. (2008). Captions and subtitles in EFL learning: an investigative study in a comprehensive computer environment. In M. P. A. Baldry (Ed.), *From Didactas to Ecolingua* (pp. 69-80). Trieste, Italy:EUT. Retrieved from http://www.openstarts.units.it/dspace/bitstream/10077/2848/1/bianchi_ciabattoni.pdf

Bird, S. A., & Williams, J. N. (2002). The effect of bimodal input on implicit and explicit memory: An investigation into the benefits of within-language subtitling. *Applied Psycholinguistics, 23*, 509-533.

Camstudio Recorder (Version 2.7.2) [Computer Software]. CamStudio.org. Retrieved from http://camstudio.org/

Canadian Association of Broadcasters. (2012, August). *Closed captioning standards and protocol for Canadian English language television programming services.* Retrieved from http://www.cab-acr.ca/english/social/captioning/cc_standards.pdf

Chang, C., Tseng, K., & Tseng, J. (2011). Is single or dual channel with different English proficiencies better for English listening comprehension, cognitive load and attitude in ubiquitous learning environment? *Computers and Education, 57*, 2313-2321.

Chiquito, A. B. (1995). Metacognitive learning techniques in the user interface: Advance organizers and captioning. *Computer and the Humanities, 28*, 211-223.

Cintas, J. (2005). Back to the future in subtitling. *MuTra 2005 - Challenges in Multidimensional Subtitling* (pp. 1-17). London: EU-High-Level Scientific Conference Series.

d'Ydewalle, G. (2002). Foreign-language acquisition by watching subtitled television programs. *Journal of Foreign Language Education and Research, 12*, 59-77.

Danan, M. (1992) Reversed subtitling and dual coding theory: New directions for foreign language instruction *Language Learning, 42*(4), 497-527.

Danan, M. (2004). Captioning and subtitling: Undervalued language learning strategies. *Meta: Translators Journal, 49*(1), 67-77.

Dornyei, Z. (2001). *Motivational strategies in the language classroom.* Cambridge, U.K.: Cambridge University Press.

Dotsub (2014) [Computer software]. New York, NY:Dotsub.com

Federal Communications Commission. (2014). *Closed captioning on television*. Retrieved from http://www.fcc.gov/guides/closed-captioning

Garza, T. J. (1991). Evaluating the use of captioned video materials in advanced foreign language learning. *Foreign Language Annals, 24*(3), 239-258.

Grgurović, M., & Hegelheimer, V. (2007). Help options and multimedia listening: The students' use subtitles and transcript. *Language Learning and Technology, 11*(1), 45-66.

Guillory, H. (1999). The effects of keyword captions to authentic French video on learner comprehension. *Calico Journal, 15*(1-3), 89-108.

Hoven, D. (1999). A model for listening and viewing comprehension in multimedia environments. *Language Learning and Technology, 3*(1), 88-103.

Hsu, C., Hwang, G., Chang, Y., & Chang, C. (2013). Effects of video caption modes on English listening comprehension and vocabulary acquisition using handheld devices. *Educational Technology & Society, 16*(1), 403-414.

iMovie (Version 2010) [Computer software]. Santa Clara, CA: Apple Inc.

Ivarsson, J. (2004). A short technical history of subtitles in Europe. *Subtitling and Translation*. Retrieved from: http://www.transedit.se/history.htm

Kikuchi, T. (1998). A review of research on the educational use of English captioned materials in Japan. *Memoirs of Numazu College of Technology, 32*, 147-160.

Kikuchi, T. (2002). An experiment on the long-term effects of keyword captions. *Memoirs of Numazu College of Technology, 36*, 171-180.

Kikuchi, T. (2006). Listening to English movies: effects on junior high school students' intrinsic interest in learning English. *Japanese Journal of Educational Psychology*, 254-264.

Kothari, B., Pandey, A., & Chudgar, A. (2004). Reading out of the "idiot box": Same language subtitling in India. *The Massachusetts Institute of Technology Information Technologies and International Development, 2*(1), 23-44.

Kothari, B., Takeda, J., Joshi, A., & Pandey, A. (2002). Same language subtitling: A butterfly for literacy? *International Journal of Lifelong Education, 21*(1), 55-66.

Krajka, J. (2013). Audiovisual translation in LSP - A case for using captioning in teaching languages for specific purposes. *Scripta Manent, 8*(1), 2-14.

Krashen, S. (1981). *Second language acquisition and second language learning.* University of

    Southern California: First internet edition December 2002. Retrieved from

    http://sdkrashen.com/content/books/sl_acquisition_and_learning.pdf.

Kuppens, H. (2010). Incidental foreign language acquisition from media exposure. *Learning,*

    Media and Technology, 35(1), 65-85.

Livingston, J. (2013). Closed captions vs. subtitles. *Closed Captioning & Subtitling Software*.

    Retrieved from: https://cpcweb.zendesk.com/entries/25142571-Closed-Captions-vs-

    Subtitles

Markham, P. L. (1988). Gender differences and the perceived expertness of the speaker as factors

    in ESL listening recall. *TESOL Quarterly 22*(2), 397-406.

Markham, P. L. (1999). Captioned videotapes and second-language word recognition. *Foreign*

    Language Annals, 32(3), 321-328.

Markham, P., & Peter, L. (2002-2003). The influence of English language and Spanish language

    captions on foreign language listening/reading comprehension. *Journal of Educational*

    Technology Systems, *31*(3), 331-341.

Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The*

    Cambridge Handbook of Multimedia Learning (pp. 31-47). New York, NY: Cambridge

    University Press.

Mitterer, H., & McQueen, J. (2009). Foreign subtitles help but native-language subtitles harm

    foreign speech perception. *PLoS ONE, 4*(11).

Neuman, S., & Koskinen, P. (1992). Captioned television as comprehensible input: effects of

    incidental word learning from context tor language minority students. *Reading Research*

    Quarterly, 27(1), 94-106.

Nunan, D. (1989). *Designing tasks for the communicative classroom*. Cambridge, U.K.: Cambridge University Press.

Nunan, D. (2004) *Task-based language learning and teaching*. Cambridge, U.K.: Cambridge University Press.

Park, C. (1994). Closed captioned TV: A resource for ESL literacy education. *ERIC Digest* ERIC Identifier: ED372662.

Price, K. (1983). Closed-captioned TV: An untapped resource. *MATESOL Newsletter, 12*, 1-8.

Robin, R. (2011). Listening comprehension in the age of Web 2.0. In N. Arnold, & L. Ducate, *Present and future promises of CALL: From theory and research to new directions in language teaching* (pp. 93-130). San Marcos, Texas: Computer Assisted Language Instruction Consortium (CALICO) Texas State University.

Rost, M. (2002). *Teaching and researching listening*. Essex. UK: Pearson Education Limited.

Rubin, J. (1994). A review of second language listening comprehension research. *The Modern Language Journal, 78*(2), 199-221.

Rubin, J. (1995). The contribution of video to the development of competence in listening. In D. Mendelson, & J. Rubin, *A guide for teaching second language listening* (pp. 151-165). San Diego: Dominie Press.

Screencast-O-Matic (2014) [Computer Software]. FFmpeg. Retrieved from http://www.screencast-o-matic.com/.

Spanos, G., & Smith, J. (1990). Closed captioned television for adult LEP literacy learners. *ERIC Digest* ERIC Identifier: ED321623

Stewart, A., & Pertusa, I. (2004). Gains to language learners from viewing target language closed-captioned films. *Foreign Language Annals, 37*(3), 438-442.

Taylor, G. (2005). Perceived processing strategies of students watching captioned video. *Foreign Language Annals, 38*(3), 422-427.

Vandergrift, L. (2007). Recent developments in second and foreign language listening comprehension research. *Language Teaching, 40*(03), 191 - 210.

Vandergrift, L. (2010). Researching listening. In B. Paltridge, & A. Phatki, (Eds.) *Continuum Companion to Research Methods in Applied Linguistics* (pp. 160-173). New York, NY: Continuum International.

Vandergrift, L., & Goh, C. C. (2012). *Teaching and learning second language listening: Metacognition in action*. NY: Routledge.

Vanderplank, R. (1988). The value of teletext sub-titles in language learning. *ELT Journal, 42*(4), 272-281.

Vanderplank, R. (2010). Deja vu? A decade of research on language laboratories, television and video in language teaching. *Language Teaching, 43*(1), 1-37. doi:10.1017/S0261444809990267

Vanderplank, R. (2013). 'Effects of' and 'effects with' captions: How exactly does watching a TV programme with same-language subtitles make a difference to language learners? *Language Teaching*, 1-16.

Windows Live Movie Maker (Version 2011) [Computer Software]. Redmond, WA: Microsoft Corporation.

Winke, P., Gass, S., & Sydorenko, T. (2010). The effects of captioning videos used for foreign language listening activities. *Language Learning and Technology, 14*(1), 65-86.

Appendix 1

Handout on Listening Strategies:

Stage of Listening Instruction (Vandergrift & Goh, 2012, p. 150)

　　1. Pre-listening—Planning/predicting stage

After learners have been informed of the topic and text type, they predict the types of information and possible words they may hear.

　　2. First listen—First verification stage

a. Learners verify their initial hypotheses, correct as required, and note additional information understood

b. Learners compare what they have understood/written with a partner, modify as required, establish what still needs resolution, and decide on the important details that still require special attention.

　　3. Second listen—Second verification stage

a. Learners verify points of earlier disagreement, make corrections and write down additional details understood.

b. Class discussion in which all class members contribute to the reconstruction of the text's main points and most pertinent details. Interspersed with reflections on how learners arrived at the meaning of certain words or parts of the text.

　　4. Third listen—Final verification stage

Learners listen specifically for the information revealed in the class discussion which they were not able to make out earlier. This listen may also be accompanied by the transcript of all or part of the text.

　　5. Reflection and goal-setting stage

Based on the earlier discussion of strategies used to compensate for what was not understood, learners write goals for the next listening activity.

Appendix 2 Sample Lesson

Objective – Enhanced listening comprehension

Level CLB 6

Material: Teacher adapted captions added to a YouTube. Example: a job interview

analysis. https://www.youtube.com/watch?v=VFTNOF77bMs

1) Review Vandergrift handout on listening strategies

2) Discuss potential vocabulary

3) Handout - Overview Comprehension questions (create on the basis of the video)

   For example,

   a. What are the main points in the video?

   b. Was there vocabulary that you didn't understand?

   c. What was the genre of this video?

   d. What did you like or not like?

4) Listen to the first sequence

5) Discuss the video in a group using the questions as a guide

6) Handout -detail comprehension questions, for example

   a. What are the details of the video?

   b. What message was the video trying to convey?

   c. Are there other vocabulary/grammar items to note?

   d. Do you notice specific ways (pronunciations, accents, etc.) of speaking?

   e. What are the speech acts?

7) Listen again

8) Discuss in groups of two and answer the questions

9) 3rd Listen to confirm the details of the video

10) Discuss and complete the answers

    a. Record new vocabulary and structures

    b. Discuss the pragmatics

    c. Note specific pronunciation elements (prosody, speech rate, Sandhi (includes assimilation, mutation, contraction, liaison, and elision), stress and rhythm, L1/L2 acoustic differences, syntax, morphology, word order, and discourse markers) that were significant.

11) Review listening strategies. What actions would you take the next time you listen to improve your understanding? How applicable was this listening to day to day conversational listening? How can you apply it?

Appendix 3

Preparing a Captioned Video

1) If it is not installed, Install Windows Movie Maker on your PC (free from Microsoft)

2) Use Keepvid.com to download the video you choose from YouTube

    a. Paste the YouTube URL to Keepvid.com

    b. Download a copy of the video – use a compatible, fairly high quality format

      (3GP, 244p is Good)

3) Open Movie Maker

a. drag the video file into the open space.
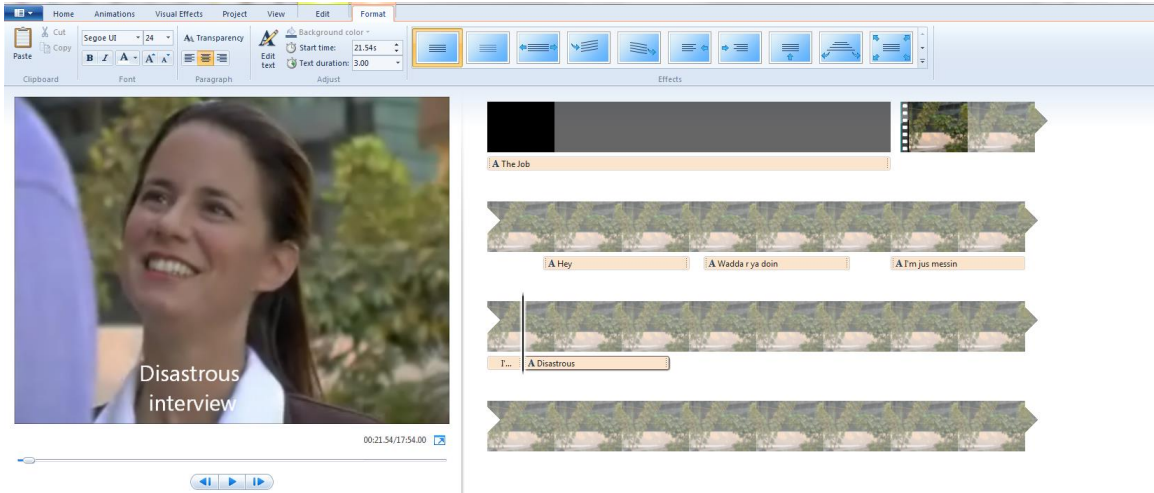
b. Play the video to the point at which a caption is needed. Click Caption on the Home ribbon and enter a caption in the text box on the left screen.

c. Listen to the video again and observe the location of the caption, drag it to where it is needed.

d. Continue until complete

e. Save the movie with one of the settings – the higher the setting the larger the file.

Appendix 4

Sample Captioned Videos from YouTube

1) **Tess of the D'urbervilles** : ENG SUB (full movie) with captions

   Authentic video, natural English, good caption quality. CLB 5 plus

   https://www.youtube.com/watch?v=NwEjOOdqGH0

2) **A Life in Japan - Documentary (English with English subtitles)**

   Documentary – authentic story of life in Japan for English speaking immigrants. Professionally

   done. Excellent material for classroom discussion. Nineteen different voices for students to hear.

   CLB 5 plus.

   https://www.youtube.com/watch?v=prNYOW0_kms

   Description from the YouTube Site

   PUBLISHED ON **Jan 2, 2013**
   Ever wondered how a life in Japan might be? 19 foreigners share their personal experiences, likes and
   dislikes. See the country through the eyes of, mostly western, foreign residents, who have lived in Japan
   between a few months and several decades.

   Without subtitles: https://www.youtube.com/watch?v=WIkQA...
   With Japanese subtitles: https://www.youtube.com/watch?v=trYqp...
   With Swedish subtitles: https://www.youtube.com/watch?v=DJeEE...

   A special edit of the final music: https://www.youtube.com/watch?v=CTiWd...

   Home page: http://www.alifeinjapan.com
   Facebook: https://www.facebook.com/pages/A-Life...
   Mimizukuzu: http://www.mimizukuzu.com

   ---

3) **English Conversations** – Medium (CLB 4-6)

   Good short authentic clips with captions. Lessons that follow each clip are primarily grammar

   related. Useful for independent learning

   https://www.youtube.com/user/kenhhoctienganh/videos

a)   Lesson 1

https://www.youtube.com/watch?v=loNGxjKkMgE&list=TLDxJhradIs8xG5uXwQyEWhgWDywJ21vIk

b)   Lesson 2

https://www.youtube.com/watch?v=4wE9LOKTT-E&list=TLyB_PY9SeW3fevfRCWoqL6GrtHXM227W0

c)   Lesson 3

https://www.youtube.com/watch?v=de_nb969txM

d)   Lesson 4

https://www.youtube.com/watch?v=00oRxjpPgGY

e)   Lesson 5

https://www.youtube.com/watch?v=3-7d6CMJz6M

f)   Lesson 6

https://www.youtube.com/watch?v=lbqOnmLgRCw

4) **Trung Mai** – as listed below this is a compilation of lessons from various sources

https://www.youtube.com/user/vantrungmai?feature=watch

a) **English Conversation** — very slow and clearly spoken English – Somewhat unnatural but useful for Lower levels

Two passes through the video – one listening only, Second listening with captions and pauses for student repetition. https://www.youtube.com/playlist?list=PL6E41B95CF7A7C1A8

b) **New Headway** – Interesting, and quite authentic stories about roommates in an apartment. All of the video is captioned, but with numerous spelling errors. Intermediate and above (CLB 5 plus). Could be used in the classroom.

https://www.youtube.com/playlist?list=PL23BAE1107EEE47A4

c) **Living English** – Australian videos of a variety of realistic life situations. Each short clip is interspersed with teaching on a specific grammar topic. Likely very good for independent

learning. Low intermediate (Less than CLB 5)

https://www.youtube.com/playlist?list=PLB1FB62B1C35C0085

d) **Study English** – Several Australian videos designed for IELTs preparation. Two steps: 1st short authentic video, 2nd extensive instruction with captioned video portions that highlight the form. A bit boring. Intermediate to advanced (CLB 6 and above) Independent learning. https://www.youtube.com/playlist?list=PL442160712B47E422

5) **English with Mr. Duncan** –Over 80 captioned British English vocabulary videos of "Mr Duncan" lecturing and sometime embodying vocabulary. Very clear British accented speech, somewhat slow speech rate, but quite entertaining and engaging. Many lectures are at a high vocabulary level and will require scaffolding. (CLB 6 and above) Excellent for either independent study or classroom use. http://www.youtube.com/user/duncaninchina?feature=watch

6) **Rachel's English** – Hundreds of free videos available on YouTube and a website. Clear pronunciation and skills development for American English. Should work for any level but some of the vocabulary is at least CLB 3. Independent study or classroom use. There is additional commercial content.

      a. **YouTube**:   https://www.youtube.com/user/rachelsenglish/videos

      b. **Web**: http://www.rachelsenglish.com/