To What Extent Can 'Karaoke'-style Subtitles on Digital Video Help Learners of English as a Foreign Language?

David LEES

Current literature on video use for language learning suggests that L2-L2 subtitles (i.e., subtitles displaying the language spoken in the audio) are perhaps more facilitative for learners than L1-L2 subtitles of the more traditional variety. This body of research, however, does not contain much information on the effects of synchronised subtitles upon the perception, noticing and eventual acquisition of L2 vocabulary items. This paper therefore describes an experimental study, investigating the extent to which synchronised 'Karaoke'-style L2 subtitled video can facilitate learners' incidental perception and noticing of L2 vocabulary. A group of 42 intermediate Japanese learners of English were recruited. The study employed a quantitative testing process to collect empirical data, combined with qualitative feedback designed to gauge the participants' first impressions of the 'Karaoke'-style subtitles compared with both standard-subtitled and non-subtitled video.

Data and results gathered from this small-scale study suggest the 'Karaoke'-style subtitled video viewing resulted in a higher degree of L2 vocabulary perception and noticing by students as well as positive opinions of the experience. This paper suggests possibilities for further, more extensive research to expand this field of enquiry.

Foreign language learning is a lengthy and difficult task. Technological advances, including smart-phones, tablets and increasing mobile internet access through such devices might, at surface level, appear facilitative of the acquisition of other languages. As indicated by Salaberry (2001), however, breadth and depth of accessible repositories of digitised L2 information, nor the seemingly 'advanced' technology itself, does not necessarily make students more enthusiastic, nor improve their acquisition. We should not, then, assume that the advance of technology by itself advances pedagogy. Despite this, it could be equally unwise to deny that new computer-assisted, digital technology and its potential scope for pedagogic utility has permitted new possibilities for educators and learners.

Perhaps the most notable technological expansion, all but 'normalised' (Bax, 2006, p. 2) into developed societies, is digital video. Video, though long since considered an "everyday" technology, has experienced a resurgence due mainly to the expansion of digital devices (Broady, 1997). As Broady states, while video as a medium "is no longer a new pedagogical tool" (1997, p. 1), it is often overlooked,

even though it possess several inherent features which support it as a “crucial addition to the teacher’s resources" (Stempleski, 1990, p. 3).

One of these potentially facilitative aspects of digital video is subtitling. The use of subtitles have been a staple of video for many years, though again has experienced a revival of sorts. Youtube, arguably the Western world’s largest digital video sharing site, as well as many online digital video streaming sites support captions in their browser-based software. Many language-learning websites such as English Central, Lingolab and Yolango employ subtitling in their videos. Following on from previous research, in this short study the author intends to further examine the extent to which video and subtitles can influence students’ lexical acquisition. This study will first introduce (or perhaps re-introduce) video as a topic of enquiry, establishing the rationale for researching subtitling and lexical-item noticing. Next, it will posit research questions based on the literature review, followed by the creation of an experimental model to test these questions. The methodology of the experiment will be explained, after which the results will be analysed and discussed in the concluding sections.

AREAS FOR CONSIDERATION

Using “video” for L2 teaching is a fraught field of study. Video faces stern criticism from some quarters, who comment that it is purely an ‘entertainment’ medium (Purushotma, 2005, p. 80). Furthermore, one must consider the pedagogical utility over the technological, not simply focus on the technology itself. Thus, in considering “digital video” and its influence on language acquisition, I will examine the literature from several perspectives.

First, I will briefly cover video as an information-conveying ‘mode’. A ‘mode’, or a ‘medium,’ is a method by which information is carried, expressed and interpreted (Kozma, 1991). Each mode has certain features unique to it, and these features can be a benefit or a hindrance to learning and language acquisition. These features also relate to the cognitive processing of received information. Second, I will look at the technological features of video. Kozma’s (1991) definition of “technology” will be used, with emphasis on the “technological/physical” and “processing capabilities” aspects. Third, I will examine the “subtitling” processing capability. Subtitles have long been a key part of video “technology,” all the more so given the digitisation of the mode. Type, problems and benefits, and cognitive processing will be explored. Fourth, I will consider vocabulary, its types and its acquisition. Schmitt’s Noticing Theory, as well as vocabulary acquisition systems, will be covered in this section to highlight how digital video as a mode might facilitate lexical uptake. Finally, I will endeavour to summarise the main points from these sections in order to inform the research questions for this investigation.

DIGITAL VIDEO AS A 'MODE'

Here, I will examine digital video’s symbol systems to determine its viability as a mode for L2 learning. As we know, people often utilise paralinguistic information (gestures, expressions, contexts and so forth) to help them understand speech-utterances (Hill, 1999, p. 2; Kress, 1996, p. 2). As it combines real-time

visual and aural elements, video is generally believed to be an "obvious", beneficial medium (Willis, 1983, p. 29). Additionally, video's context, realism and motivational characteristics are also frequently said to be of benefit to L2 learners.

Video is considered to 'realistic', in that it displays a combination of real-time audio and visual data simultaneously. As these streams broadly overlap with our day-to-day sensory experience, and is therefore 'familiar', it is likely that this information can be processed relatively smoothly (Alter, 2009). Indeed, although L2 learners may struggle with culturally-specific markers, as Broady (1997) notes; McCloud states that the less abstract (worded, linguistic) and more representational (visual) the data, the easier it is to process (1993, p. 49).

Context is also important; King suggests that "learners' encounters with realistic situations and exposure to living language provide a dimension that is missing in text-book orientated teaching" (2002, p. 510). Further, Willis notes, "[m]ost language students say they find video tape easier to understand than audio tape: Sturtridge (1976) found through experimentation that her students preferred a bad quality video tape to a good quality audio tape" (1983, p. 30). With the exception of overly stylised editing, video is held to be sufficiently contextualised by the situation in which it was captured (Tschirner, 2001, p. 315). Some commentators note that over-edited, over-modified video often creates an 'unnatural' atmosphere; according to King, "purpose-made language video clips quickly lose their appeal" (2002, p. 512).

Complications remain, however. As Broady (1997, p. 4) shows by analysing several French news broadcasts, it is common for an audio-track to convey a culturally exclusive meaning unsupported by the visual data stream. When this occurs, research shows that learners may become confused (Gruba, 2006, p. 86). Video's content and cultural markers should of course be evaluated for learner suitability (Fawkes, 1999, p. 33).

Finally, discussions of video for L2 learning frequently touch on motivation. Motivation is of course a central requirement for L2 acquisition (Lommel, 2006, p. 255), though with regards to the use of video – or indeed "any technology primarily used for entertainment" (Purushotma, 2005, p. 80) – for language learning, some commentators are sceptical of video's utility (Tudor, 1987, p. 203). Despite these limitations, video is frequently noted as "intrinsically motivating" (Bayon, 2004, p. 2; Lonergan, 1984, p. 5), and this motivational power has been commented on by many (Fawkes, 1999; Lonergan, 1984; MacKnight, 1983; Stempleski, 1990). Video essentially 'captures' segments of reality, allowing teachers to provide a motivational glimpse of the target-language environment (Fawkes, 1999, p. 3).

Video as a mode seems promising due to its realistic, contextualising and motivating attributes. Additionally, many significant technological factors, discussed in the next section, support these positive characteristics.

DIGITAL VIDEO: TECHNOLOGY AND PROCESSING CAPABILITIES

Although Salaberry warns that "improved technology does not imply improved pedagogy" (2001, p. 39), teachers should still be wary of the 'detrimental effects' to their students of not keeping up with current hardware, software and

technology, as predicted by Lindenau (1984, p. 123). Also, while many institutions install new CALL labs and computer rooms, many researchers comment that such expansion is frequently under-informed (Levy, 2008, p. 339). Thus, I will examine digital video's "technology" and its "processing capability" (Kozma, 1991, p. 2).

Technological Features

According to Kozma, 'technology' denotes that which digital video is able to accomplish as a result of mechanical and electronic features (1991, p. 3). Compared to "big-TV/VCR-on-wheels-with-blackout-curtains" technology (Sherman, 2008, p. 28), digital video's 'technology' makes the video 'mode':

- *Portable* – video is now portable, stored on Flash memory and other data-storage devices, and can be viewed on a wide range of digital viewers such as music-players, smart-phones, tablet and note-computers;
- *Transferable* – as data, video is now transferable between peers. One video clip can be shared with an unlimited number of people;
- *Editable* – the advances of home computing and free software means that video-editing is widely achievable. A video can be cut down to a target segment, encoded as an audio file, have individual frames captured and used as pictures or have subtitles attached; and
- *Online* – digital video is online, making it widely available, free and easily viewable to users with access to the internet through online digital devices.

Given the increased portability of contemporary digital devices, characteristics such as shape, size, weight and functionality suggest that instead of a traditional textbook, more students might now study using a digital device (Kozma, 1991). Interactivity, an important aspect of L2 acquisition (Haldane, 2007; Sims, 1995; Vygotsky, 1978), has also improved. By interacting with the video and/or discussing it online with their friends, learners can create more cognitive links regarding a word or phrase, potentially increasing retention (Draper, 1996). Additionally, due to this technological portability and interactivity, video has become familiar, having achieved 'normalisation' in daily life (Bax, 2006, p. 2). 'Familiarity', also known as 'cognitive fluency', states that people are more ready and eager to process information in ways that are familiar to them (Alter, 2009; Sherman, 2008).

Processing capabilities

It would seem that digital video has become a highly portable, viewable, interactive and familiar medium due to its technological features. Additionally, however, digital video also possesses important processing capabilities.

When considered from an information-conveying perspective, video as a mode is unique. The symbol systems – both audio and visual streams, running simultaneously in real-time – combine to provide a realistic information stream. As discussed in the previous section, more so than comic books (visual), books (abstract-visual, linguistic) and radio (aural), it is with this mix of both audio and

visual streams that video is considered by many to best capture the contextualised reality experienced by humans (Alter, 2009; Tschirner, 2001). However, while video may possess the same symbol systems as a televised broadcast, video has several processing capabilities which affect the processing of information conveyed to the viewer (Kozma, 1991). As an example of this, note that a TV broadcast runs "front-to-back" (Mackey, 2002), while a digital video clip can easily be paused, searched, rewound, reviewed and viewed with or without subtitles.

The ability to pause, rewind and review information makes the medium 'stable', like a book, instead of 'transitory'. Kozma places great import on this:

> In many situations for fluent readers, reading progresses along the text in a forward direction at a regular rate and the information could just as well be presented in another, more transient medium. But on occasion, processes interact with prior knowledge and skill in a way that relies heavily on the stability of the text to aid comprehension and learning. (Kozma, 1991, p. 5)

Thus, akin to reading a novel, to watch video might occasionally require a 'replay' to reaffirm comprehension (Mackey, 2002). Arguably, compared to previous analogue media, digital video is far more "trackable" (Tschirner, 2001). Pacing – defined by Wright (1984) as the amount of information presented within a set time – can also be an issue with "front-to-back" video. Studies have shown that short-term to long-term information-chunk processing speed is dependent on the length of the chunk and the level of the viewer's background knowledge. It is therefore possible that the pace of information in transient media can cause comprehension loss (Kozma, 1991, p. 16). Students often feel overwhelmed by a new native-speed L2 video, though techniques such as pausing and replaying helps them recover from comprehension failure (Gruba, 2006: 87; Mackey, 2002). Pausing the video and thinking about the visual-audio relationship provides cognitive linking (Mayer, 1994). Gruba's investigation shows this; his research illustrates that learners who engage with L2 video are able to build information signposts based on the audio-visual streams, assisted by paralinguistic information (Gruba, 2006).

Digital video's improved technological features, combined with its processing capabilities and its features as an information-conveying mode, seem to hold potential to facilitate L2 learning. With these considerations, I will now examine subtitling, and consider its benefits and drawbacks.

Digital video and subtitling

Subtitling plays a central role in digital video's processing capabilities. To date, evidence suggests that L2 audio with L1 subtitles generally facilitates acquisition (Broady, 1997). Additionally, researchers consider L2 audio and L2 subtitles to be conducive for "activating language already in the learners heads" (Broady, 1997; Vanderplank, 1990), though "all (variants of subtitles) make different demands on the students' linguistic skills and are equally valuable" (Williams, 2000). A hybrid approach, termed "DualSubs" technology which

enabled learners to watch videos with both L1 and L2 subtitles on the screen, was found to be beneficial, though several viewers experienced confusion (Bayon, 2004). Overall, subtitles are held to:

- enhance the 'stability' of the information in the video, to the degree that several researchers refer to "video" as "videotext"(Gruba, 2004, 2006; Mackey, 2002, p. 22).
- provide an "advanced organiser", allowing students to "pre-load" (i.e., anticipate, activate, or extrapolate) vocabulary visually before they hear it (Gruba, 2004, p. 60).
- increase the saliency of lexical items in the video (Carroll, 2006).

Subtitles are already widely used in digital video. Online websites like EnglishCentral use L2 and dualsubs subtitles. With more 'stable' videotext, authentic videos may become more accessible for L2 learners (Peachey, 2008).

In sum, it would appear that subtitling is believed to aid L2 learner's acquisition, extrapolation, and the ability to 'read' a video's audio track as videotext. The next section will address how video, and in particular its capacity for subtitling, could potentially help these learners to acquire lexical items in their target language.

LEXICAL ITEMS, PROPERTIES AND PROCESSING

Researchers list many issues associated with learning and maintaining an L2 vocabulary, including the required word-knowledge of L2 lexical items (Nation, 1997; Schmitt, 2000), the number of words, as well as interference effects of the learners' L1 knowledge (Ellis, 1997a; Schmitt, 1997). Researchers also remain divided as to how lexical items are acquired, though commentary points chiefly to the 'Weak Interface' system and to the concept of 'Noticing'.

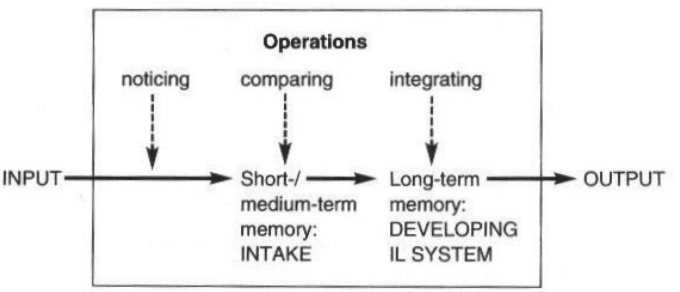
Weak Interface System

The 'Weak Interface' system combines a range of widely accepted views on the cognitive processes undertaken during L2 acquisition (Pica, 2005, p. 276). The Input to which an L2 learner is exposed is not necessarily taken into the learner's short-term memory:

Reflecting on this system, Ellis states that:

> In accordance with the current theories of L2 acquisition, the process by which input becomes implicit knowledge is seen to involve two principal stages: one where input becomes intake, which involves the operation of noticing, and one where intake becomes part of the learner's interlanguage system. (1997b, p. 119)

Figure 1
Diagram of the main SLA processes (Ellis, 1997a)



In addition to the potential problems noted above regarding the nature of L2 vocabulary, acquiring new vocabulary rests first on perceiving and noticing Input (Ellis, 1997b, p. 119). Logically, if a specific lexical item is not perceived by an L2 learner, it cannot be taken into the learner's interlanguage system. This is clearly of central importance, and will be discussed below.

Noticing and Perception

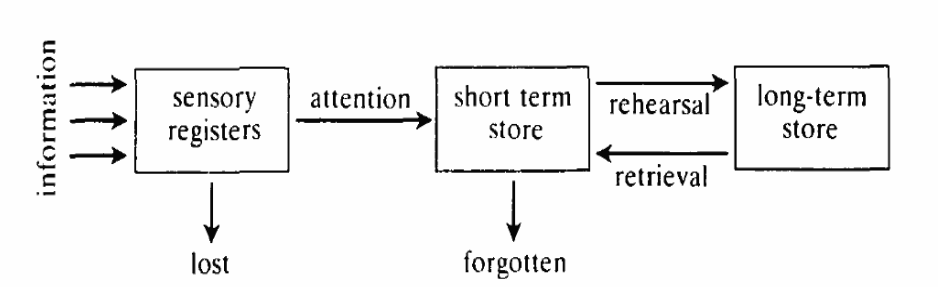
As noted, the 'Weak Interface' position incorporates many of the main theories postulated about L2 acquisition. Schmidt's Noticing theory is one of these. Schmidt states that "conscious" attention is required to gain information from input (Robinson, 1995; Schmidt, 1990), and while "subliminal" language learning is dismissed, incidental L2 acquisition is "clearly possible and effective" (Schmidt, 1990, p. 129). "Consciousness" essentially corresponds to its dictionary definition, though according to Schmidt it covers several different areas, "intention", "awareness", and "attention" (Ellis, 1997a, p. 116).

"Intention" refers to the distinction between explicit and incidental learning. Schmidt notes that this role of consciousness is not particularly a priority, stating that "it makes no difference whether the learner notices a linguistic form in input because he or she was deliberately attending to form, or purely inadvertently" (Schmidt, 1990, p. 139). This, even by itself, could be of importance. As an example, train or bus announcements in a foreign language could be heard as noise at first, then perceived to carry meaning as the learner learns more of the language. "Attention", the second form of consciousness, involves "noticing the properties of the input" (Ellis, 1997a, p. 116) and focuses on identifying specific linguistic information. The meaning understood from train or bus announcements expands gradually from basic information, such as the next stop, to include more complex information (e.g. adverts, driver warnings) as the learner progresses in his or her learning.

"Awareness", the third form of consciousness, states that learners are required to be aware of the knowledge that they learn. This consciousness form deals with three progressive levels of recognition within the input itself, and we should note here that Schmidt defines and treats "noticing" and "perception" in specific ways. The progressive levels of awareness are; *perception* (the information

that is ‘recorded’ by our senses); *noticing* (a focal awareness of a certain part of the information that is perceived by our senses); and *understanding* (perceived and noticed information is compared to currently available and relevant data in order to comprehend its significance) (Schmidt, 1990). The first two stages, *perception* and *noticing*, effectively correspond to the “Noticing” in the first stage of the cognitive processes of second language acquisition (SLA) in Figure 1, while the final *understanding* stage is aligned with the “Comparing” stage. Schmidt’s theory also fits well within most contemporary cognitive memory models such as that proposed by Kihlstrom (1984), as depicted in Figure 2.

Figure 2
Effect of consciousness in a cognitive memory system (Kihlstrom, 1984)



Schmidt's *perception* (the recording of information) takes place at the sensory registers in Kihlstrom's model. Following this, *noticing* and *understanding* provide the 'attention' required to transfer the perceived information to the 'short term store'. Of special importance here are the arrows 'lost' and 'forgotten'. Information not assimilated into the 'long term (memory) store', such as information in the Input not *perceived*, cannot be counted as Intake and is therefore "consigned to oblivion" (Kihlstrom, 1984, p. 165).

Though Schmidt insists that Input can only become Intake through aware *noticing*, his Noticing theory contains notable points which are widely accepted throughout the field of SLA. First, he agrees that *detection / perception* of a language feature is a prerequisite for both implicit and explicit acquisition processes (Ellis, 1997b, p. 116). Indeed, during Schmidt's discussions about his learning Portuguese, he refers to occasions when he began "… hearing things I never heard before" (Schmidt, 1990, p. 140). It is important to note here that it was "heard"; that is, it was isolated and could then be extracted from the audio information stream.

Secondly, we should note that the process of Noticing may be influenced by several factors in the Input itself. Ellis summarises these factors as task demands, frequency, unusual features, salience, interactional modification, and existing linguistic knowledge (Ellis, 1997a, p. 120). Of particular interest here is the *salience* of words. Carroll states that the salience of a word can often be increased by focal accent, volume and extended pauses between words (Carroll, 2006, p. 19). Carroll further considers that "[p]ossibly the first step in learning a language is learning to segment the speech signal so that the continuous sound stream is perceived as a linear sequence of sound forms" (Carroll, 2006, p. 22).

Clearly, though improving the chance that Schmidt's Noticing may occur, neither salience itself nor these other factors guarantee that a certain linguistic feature will be *perceived*, *noticed* and then, as Intake, transferred to the short-term memory store. However, Noticing theory research does suggest that lexical items may be highlighted through several techniques. Next, I shall summarise the findings of the literature review, and posit research questions for this video.

IMPLICATIONS AND RESEARCH QUESTIONS

Studies of video as a technology, medium and pedagogical aid for L2 learning consider video to have a high level of potential due to its realistic combination of visual and audio data (Willis, 1983, p. 29). The visual information stream supports the audio through context and content (Tschirner, 2001, p. 315). Native authenticity both motivates and provides the viewer with a wealth of linguistic and paralinguistic information (King, 2002, p. 510). Furthermore, subtitles grant stability (Gruba, 2004, p. 54; Mackey, 2002, p. 22), facilitate saliency (Gruba, 2004, p. 60), aid organisation and also function as memory "pegs" with which to assist perception, comprehension and retention of form and meaning (Kozma, 1991, p. 181).

Subtitles prove useful in many situations. For lower-level learners they function as a safety-net. L1-L2 subtitles assist with concept-linkage and short chunk translation. L2-L2 subtitles 'pre-load' the vocabulary into the viewer's short-term memory, allowing them to scan through the L2 audio to 'catch' or *perceive* the corresponding words. For higher level learners they can provide a stable transcript allowing multiple "pegs" to remind and re-activate vocabulary, or allow new acquisition. Based on these considerations, as well as previous research (Lees, 2012), I believe that the degree to which further subtitle intervention, such as that provided by synchronised 'Karaoke'-style subtitles, could positively influence L2 learners' ability to perceive and notice individual vocabulary items warrants further investigation. In doing so, I shall seek to answer the following research questions:

- To what degree does synchronised 'Karaoke'-style subtitled digital video facilitate or hinder incidental word-perception when compared with standard-subtitled and non-subtitled digital video?
- What are the general opinions of L2 learners regarding 'Karaoke'-style subtitles compared to standard-subtitled and non-subtitled digital video?

In order to answer these questions, I conducted an experimental investigation. Below, I outline the methodology and the procedure used to administer the experiment and gather the results. After collecting and analysing the data, I will display and discuss the findings in reference to both research questions and the relevant literature.

METHOD

To examine and compare influence on incidental perception, three groups were created. These were a 'Karaoke'-subtitle group, a 'standard'-subtitle group and a no-subtitle group, drawn from 42 pre-intermediate Japanese learners of English, who were taking an elective English class at a Japanese university. Each student received a CD, which contained a copy of the same video – a 5-minute clip from a British TV-show called *Top Gear* – with different types of subtitles as outlined above; 'Karaoke-style subtitles, 'standard' subtitles, and no subtitles. As the TV show is British, the presenters use authentic, native-speed English; additionally, the clip in question was filmed in Japan, in order to ensure some degree of familiarity. The clips were watched simultaneously in a computer lab. Before watching their video, each participant was briefed on the overall context of the video clip and was asked merely to concentrate for the duration. After the first viewing, participants completed a 20-item word-recognition test. At this point, they watched the 'Karaoke'-subtitle video together, and completed a questionnaire comparing their viewing experiences. The surveys and questionnaires were administered online, in English and Japanese, using Google Docs for ease of data collection and collation.

RESULTS

Word Noticing and Perception Tests

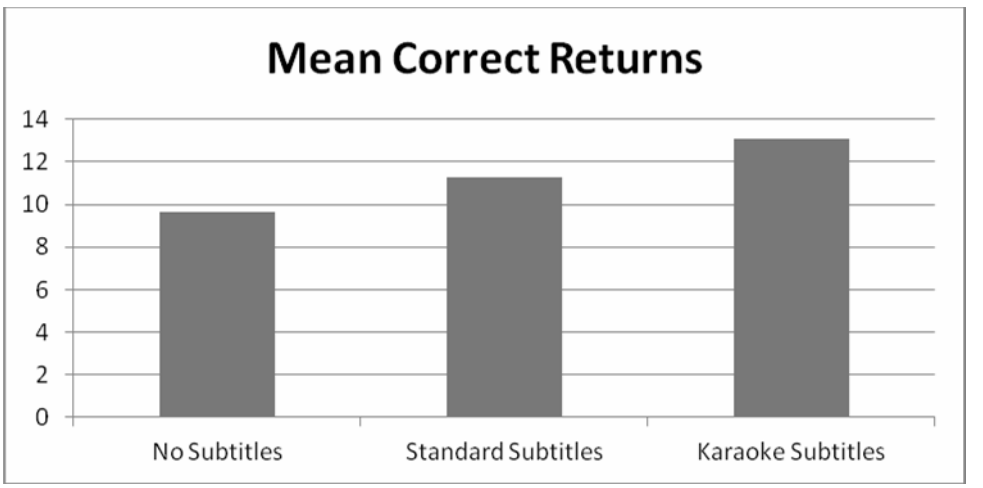
The degree to which individual words were perceived and taken note of by the participants was investigated through the completion of the word-recognition test. The results were identified, tabulated and collected into a chart.

Table 1
Collected Results from the Word Perception and Noticing Test

	Correct	Incorrect
No Subtitles	135	145
n = 280 (mean)	*(9.64)*	*51%*
Standard Subtitles	169	131
n = 300 (mean)	*(11.26)*	*43%*
Karaoke Subtitles	170	90
n = 260 (mean)	*(13.07)*	*35%*

Notably for the correct returns –where the participants were able to correctly identify a word that had appeared in the video, or recognise a word that had not – these results illustrate a trend favouring subtitled videos over non-subtitled videos, and the 'Karaoke'-style-subtitled video in particular.

FIGURE 3
Mean Correct Returns



While this trend alone cannot be considered proof that ‘Karaoke’-style subtitles will always facilitate the perception of individual words to a greater degree than standard subtitles or a non-subtitled digital video, they do suggest that the 'Karaoke'-style subtitles have a facilitative effect on incidental lexical noticing.

Participants’ impressions and comments

Analysis of the quantitative survey of the participants shows that a sizeable majority preferred the 'Karaoke'-style subtitled video over the Standard subtitled and Non-subtitled videos after a second viewing:

Table 2
Collected Results from the Impressions Survey

I thought that the Karaoke-sub video…	***(1)*** *Strongly Disagree*	***(2)*** *Disagree*	***(3)*** *Neutral*	***(4)*** *Agree*	***(5)*** *Strongly Agree*
a) was easier to listen to		5	6	11	20
b) was easier to understand		2	5	11	24
c) was easier to read		2	7	11	22
d) allowed me to better notice words	1	2	5	8	26
e) allowed me to better link the "audio" words to the subtitles		2	4	11	25

It should also be noted that the participants volunteered many comments regarding the experiment and their opinions regarding non-subtitled, standard-subtitled and 'Karaoke'-style subtitled video. A large number of them supported the synchronised style of subtitling, as evidenced by the following remarks:
「字幕はないよりもあるほうが、頭の中にはいってくるので、役に立つと思います。なので、こんなカラオケ字幕のついたビデオも見てみたいです」*"I think with subtitles is better than not having them, as by seeing the words they come into my head more easily, which is useful. I would like to watch more videos with Karaoke subtitles."*

「英語のリスニングをしているときに、一番ききとりにくいのは言葉と言葉をつなげて話されたときなのでカラオケ字幕があると、どの単語を言っているのかすぐわかるのでとてもいいと思う。またこの単語はネイティブの人が話すとこのように聞こえるのかということもわかって、ほかのビデオより学習しやすいと思った」*"When I'm listening to English, the most difficult thing for me is to understand the words when they are all strung together, so for me I found that the Karaoke subtitles made it easy for me to understand what words were being said. I also was able to catch what certain words sound like when a native speaker says them, so I thought that I'd like to use this method to study with other videos too."*

「カラオケ字幕のほうが、理解しやすかったです。どこを話しているのかがわかるからだとおもいます」*"The Karaoke subtitles were easier to understand, because I felt I understood where and when they were saying something."*

「私的にはこのビデオはとても見やすくて分かりやすかったと思う。普通の字幕だと、もしも会話を見失ったときに分からなくなってしまうがカラオケ音声だと言葉を追いやすくていい」*"For me, this (Karaoke subtitle) video*

was easy to watch and understand. With standard subtitles, I sometimes look ahead in the speech and get lost, but Karaoke subtitles are good because they are easy to follow."

「発音や単語と単語のつながりに注意して聞くことができる」*"With (the Karaoke subtitles) I could concentrate and catch the pronunciation, and the different words."*

On the other hand, some participants stated that they preferred standard subtitles:

"I think Standard Subtitles is better, because Karaoke style made me tired."

「ふつう字幕とカラオケ字幕の違いがあまりわからなかった」*"I didn't really feel there was much difference between the Standard and the Karaoke subtitles."*

"I think there are only little bit different, Karaoke and Standard subtitles." (sic)

On balance, therefore, it would seem that the 'Karaoke'-style subtitles, despite their unfamiliar and experimental nature, were relatively well received by the participating English learners. Furthermore, many of the participants themselves reported that the 'Karaoke'-style subtitles enabled them to better 'catch' specific words in the native-speed audio stream, suggesting that the word salience was favourably improved by this type of captioning.

CONCLUSION

Based on the results gathered, analysed and summarised above, it is possible to conclude that the ‘Karaoke’-style L2 subtitled video under investigation in this study can be seen to have positively facilitated learners of English as a second language in incidentally perceiving individual L2 vocabulary items. The experimental trial of ‘Karaoke’-style subtitles in this study revealed that whilst enhancing some of the saliency properties of standard L2 subtitles and maintaining much of their benefits over non-subtitled video, the synchronised subtitles were also, on whole, positively received by the participants. This is not to state that these ‘Karaoke’-style subtitles inherently facilitate language perception by the viewer to a higher degree than standard intralingual L2 (L2-L2) subtitles; the findings of one small-scale study are of course insufficient for such a claim.

However, the results of the investigation are promising. Additional experimental research into this area to confirm the findings of this study would provide a firm foundation for further future inquiries.

REFERENCES

Alter, A. (2009). Suppressing Secrecy Through Metacognitive Ease. *Journal of Psychological Science, 20*(11), 1414-1420.

Bax, S. C., A. (2006). Making CALL Work: Towards Normalisation. *System, 34*, 465-479.

Bayon, V. (2004). Exploiting Films and Multiple Subtitles Interaction for Casual Foreign Language Learning in the Living Room Retrieved 5/1/2009, 2009, from http://www.virart.nott.ac.uk/idue/pdf/icec2004.pdf

Broady, E. (1997). Old Technology New Technology: Video Makes a Come-Back. *New Technologies in Language Learning and Teaching*, 1-14.

Carroll, S. (2006). *Salience, Awareness and SLA*. Paper presented at the 8th Generative Approaches to Second Language Acquisition Conference, Somerville, MA.

Draper, S. (1996). Integrative Evaluation: An Emerging Role for Classroom Studies of CAL. *Computers and Education, 26*(1 & 2), 17-32.

Ellis, R. (1997a). *Second Language Acquisition*. Oxford: Oxford University Press.

Ellis, R. (1997b). *SLA Research and Language Teaching*. Oxford: Oxford University Press.

Fawkes, S. (1999). *Switched On? Video Resources in Modern Language Settings*. Clevedon: Cromwell Press.

Gruba, P. (2004). Understanding Digitised Language Videotext. *Computer Assisted Language Learning, 17*(1), 51-82.

Gruba, P. (2006). Playing the Videotext. *Langauge Learning and Technology: A Media Literacy Perspective on Video-Mediated L2 Listening, 10*(2), 77-92.

Haldane, M. (2007). Interactivity and the Digital Whiteboard: Weaving the Fabric of Learning. *Learning, Media and Technology, 32*(3), 257-270.

Hill, B. (1999). *Video In Language Learning*. London: CiLT.

Kihlstrom, J. (1984). Conscious, Subconscious, Unconscious: A Cognitive Perspective. In K. M. Bowers, D (Ed.), *The Unconscious Reconsidered*. New York: Wiley-Interscience.

King, J. (2002). Using DVD Feature Films in the EFL Classroom. *Computer Assisted Language Learning, 15*(5), 509-523.

Kozma, R. (1991). Learning With Media. *Review of Education Research, 61*(9), 172-212.

Kress, G. V. L., T. (1996). *Reading Images: The Grammar of Visual Design*. New York: Routledge.

Lees, D. (2012). To What Extent can the use of Digital Video Benefit L2 Learners of English? *Himeji Dokkyo Daigaku Foreign Studies Department Kiyou Journal., 21*.
Levy, M. (2008). Language, Literacy and Technology. *TESOL Quarterly, 42*(2), 339-343.
Lindenau, S. (1984). The Teacher and Technology in the Humanities and Arts. *Modern Language Journal, 68*, 119-124.
Lommel, V. (2006). Foreign Grammar Acquisition While Watching Subtitled Television Programmes. *British Journal of Educational Psychology, 76*, 243-258.
Lonergan. (1984). *Video In Language Teaching*. Cambridge: Cambridge University Press.
Mackey, M. (2002). *Literacies Across Media: Playing the Text*. New York: Routledge.
MacKnight, F. (1983). Video and English Language Teaching in Britain. In C. Brumfit (Ed.), *Video Applications in English Language Teaching* (pp. 1-15). Oxford: Pergamon Press Ltd.
Mayer, R. (1994). For Whom is a Picture Worth A Thousand Words? Extensions of a Dual-Coding Theory of Multimedia Learning. *Journal of Educational Psychology, 86*(3), 389-401.
McCloud, S. (1993). *Understanding Comics: The Invisible Art*: Harper Collins Inc.
Nation, P. (1997). Vocabulary Size, Text Coverage and Word Lists. In N. M. Schmitt, M (Ed.), *Vocabulary: Description, Acquisition and Pedagogy* (pp. 6-19). Cambridge: Cambridge University Press.
Peachey, N. (2008). Online Video: What Does it Have to Offer? Retrieved 5/1/2009, 2009, from http://nikpeachey.blogspot.com/2008/12/online-video-what-does-it-have-to-offer.html
Pica, T. (2005). Second Language Acquisition Research and Applied Linguistics. In E. Hinkel (Ed.), *Handbook of Research in Second Language Teaching and Learning* (pp. 263-281). London: Routledge.
Purushotma, R. (2005). You're Not Studying, You're Just... *Language Learning and Technology, 9*(1), 80-96.
Robinson, P. (1995). Aptitude, Awareness and the Fundamental Similarity of Implicit and Explicit Second Language Learning. In R. Schmidt (Ed.), *Attention and Awareness in Foreign Language Learning* (pp. 303-358). Hawaii: University of Hawaii.
Salaberry, R. (2001). The Use of Technology for Second Language Learning and Teaching: A Retrospective. *The Modern Language Journal, 85*(1), 39-56.

Schmidt, R. (1990). Role of Consciousness in Second Language Learning. *Applied Linguistics, 11*(2), 129-158.
Schmitt, N. (1997). *Vocabulary: Description, Acquisition and Pedagogy*. Cambridge: Cambridge University Press.
Schmitt, N. (2000). *Vocabulary in Language Teaching*. Cambridge: Cambridge University Press.
Sherman, J. (2008). Where are we with Media Video?, *IATEFL CALL Review*.
Sims, R. (1995). Interactivity: A Forgotten Art? *Computers in Human Behaviour, 13*(2), 157-180.
Stempleski, S. T., B. (1990). *Video In Action: Recipes For Using Video In Language Teaching*. Cambridge: Cambridge University Press.
Tschirner, E. (2001). Language Acquisition in the Classroom: The Role of Digital Video. *Computer Assisted Language Learning, 13*(3-4), 305-319.
Tudor, I. (1987). Video as a Means of Cultural Familiarization. *System, 15*(2), 203-207.
Vanderplank, R. (1990). Paying Attention to the Words: Practical and Theoretical Problems in Watching Television with Uni-Lingual (CEEFAX) Subtitles. *System, 18*(2), 221-234.
Vygotsky, L. S. (1978). *Mind in Society*. Cambridge, MA: Harvard University Press.
Williams, H. T., D. (2000). The Value of Teletext Subtitling as a Medium for Language Learning. *System, 28*, 217-228.
Willis, J. (1983). The Role of the Visual Element in Spoken Discourse: Implications for the Exploitation of Video in the EFL Classroom. In C. Brumfit (Ed.), *Video Applications in English Language Teaching* (pp. 29-42). Oxford: Pergamon Press Ltd.
Wright, J. C. (1984). Pace and Continuity of Television Programs: Effects on Childrens Attention and Comprehension. *Developmental Psychology, 20*.