Revisiting 'Karaoke'-style Subtitles on Digital Video: To What Extent Do They Help or Hinder Recognition?

David LEES

The current body of research concerning video and subtitling suggests that while traditional interlanguage (L1-L2) subtitles may help on-the-spot understanding of a given scene or utterance, intralanguage (L2-L2) subtitles help learners' L2 recognition, noticing and general acquisition of vocabulary to a higher degree. This research, however, is not concerned with whether or not synchronized subtitles facilitate or hinder the recognition, noticing and eventual acquisition of L2 vocabulary items.

A pool of 50 intermediate-level Japanese learners of English were recruited, and split into three groups. An authentic L2 video-clip was shown to these three groups; each group watched a different version of the video-clip – a non-subtitled, a subtitled, and a synchronised 'Karaoke'-style L2 subtitled version. An experimental study then tested the participants' incidental perception and noticing of certain L2 vocabulary. Data and results gathered from this repeated small-scale study align with the previous study; chiefly, it reports a higher degree of L2 vocabulary perception and noticing by students viewing the 'Karaoke'-style subtitled video, as well as receiving positive opinions from the participants concerning synchronized subtitling. This paper finally posits further possibilities for future research to expand our body of knowledge on synchronized subtitles.

INTRODUCTION

Learning a foreign language (L2) remains a difficult task, despite the advances in technology that modern societies have long since adopted and currently take for granted. This should, of course, be expected as Salaberry (2001) points out, neither the accessible repositories of digitised L2 information, nor the modern and advanced technology itself, necessarily makes the student more enthusiastic, more aware, more intelligent, more diligent, or ultimately improve their acquisition. One would be unwise to assume, then, that the general advance of technology alone thusly advances pedagogy. Despite noting this, one would be equally unwise to deny that new computer-assisted, digital technology has scope for pedagogic utility. Concerning scope, then, the most notable technological advancement all but completely adopted and 'normalised' (Bax, 2006) into modern societies is digital video. Video, though long since considered an "everyday" technology, has since experienced a resurgence due mainly to the expansion of digital devices (Broady,

1997). Although video as a medium is no longer a new pedagogical tool (Broady, 1997), according to Stempleski it is often passed over as a language medium or delivery-method, though it does possess several innate features which support it as a valuable addition to the teacher's resources (1990).

One of these potentially facilitative aspects of digital video is subtitling, otherwise known as captioning. Subtitles have been a particular staple of the video medium for many years, chiefly as intralanguage aides for the hearing-impaired or as interlanguage aides on L2 movies and recorded broadcasts, though they have recently come to language learning amid a digital revival. *Youtube, YouKu, NicoNicoDouga*, as well as many online digital video streaming sites support captions in their browser-based software. Many language-learning services such as *English Central, Yolango* and *Lingolab*, employ subtitling in their videos. Following previous research conducted by the author of this paper (Lees, 2012, 2014), this study revisits the experiment in order to test if its results are replicable, and to ascertain once again whether or not video and subtitles can potentially influence students' noticing and recognition of lexical items.

First, I will introduce video as a topic of enquiry, conduct a literature review and establish the rationale for researching subtitling and lexical-item noticing. Next, I will summarize both the main points of the literature review and the findings of the previous experiment, to inform the research objectives for this investigation. Third, I will outline the methodology, which will remain the same as the former. Finally, after gathering the results, analysis will be conducted on the data from this current experiment, which will be compared to the previous data, and its implications discussed.

LITERATURE REVIEW

Areas For Consideration

Video, though frequently thought to be purely an 'entertainment' medium (Purushotma, 2005, p. 80), frequently gets short shrift in language education; it is often under-utilised, or over-utilised while uninformed, and as a result has not been employed as widely or as successfully in L2 education as it could have (Fawkes, 1999; MacKnight, 1983). While this is currently changing in this field, as evidenced by both the online digital-videos of *English Central* and the recent introduction of "6 Skills" textbooks by Oxford University Press (which include "Viewing" and "Presenting" alongside the more traditional 4 Skills) (OUP, 2015), understanding video requires understanding multiple perspectives; chiefly, an understanding of the medium itself, the technology that supports it, and the cognitive processes associated with it, in addition to employing it with a pedagogically sound approach. Regarding digital video's influence on language acquisition, this paper will examine literature from several perspectives.

First, I will briefly cover video as an information-conveying 'medium'. A 'medium', or a 'mode', is method by which information is carried, expressed and interpreted (Kozma, 1991). Each 'medium' has certain features unique to it, and these features can act as an aide or as a hindrance to L2 acquisition and learning. These features also touch on 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 medium. Type, problems and benefits, and cognitive processing will be explored. Fourth, I will consider vocabulary, its types and its acquisition. Schmidt's 'Noticing Theory' will be applied in this section to highlight how digital video could potentially facilitate lexical uptake. Last, I will summarise the main points from these sections, so as to condense the information for ease of reference when reviewing the previous study's findings.

Digital Video As A Medium

This section will examine digital video's symbol systems – i.e., how video as a medium conveys its meaning – with reference to it being a medium for L2 learning.

Humans frequently utilise paralinguistic information (gestures, expressions, and contexts) to help them understand speech-utterances (Kress, 1996). As it combines real-time visual and aural elements – as opposed to, say, a book, which only contains visual, word-based meaning – video is generally held to be an "obviously 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, though admittedly not always (Bayon, 2004; Fawkes, 1999).

Video is considered 'realistic'; it displays audio and visual data simultaneously. This essentially overlaps with humans' day-to-day sensory experience, making it all the more 'familiar' in both definition and concept, which makes it highly likely that this information can be processed 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).

Also of note is 'context.' 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). Along these lines, Willis (1983) notes:

> Most language students say they find video easier to understand than audio: Sturtridge (1976) found through experimentation that her students preferred a bad quality video tape to a good quality audio tape. (1983, p. 30)

As such, many hold video to be sufficiently contextualised by the situation in which it was captured (Tschirner, 2001), baring overly stylised edits and creative license. Frequently, over-edited, over-modified video creates an 'unnatural' atmosphere, which, according to King, causes purpose-made language video clips to "quickly lose their appeal" (2002, p. 512).

Complications remain, however. As Broady (1997) demonstrated 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. Learners can easily become confused in such situations (Gruba, 2006), which highlights the need for educators to evaluate videos' content and cultural-markers for suitability (Fawkes, 1999), as is frequently done with print media such as textbooks and graded readers.

Finally, discussions of video for L2 learning frequently touch on 'motivation.' Motivation is of course a central requirement for L2 acquisition (Lommel, 2006), though with regards to the use of video or indeed "any technology primarily used for entertainment" (Purushotma, 2005, p. 80), for language learning, some remain sceptical of video's utility. However, video is frequently noted as "intrinsically motivating" (Bayon, 2004, p. 2; Lonergan, 1984, p. 5), and this motivational power is 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), as well as use these glimpses for L2 learning.

Video as a medium offers realistic, familiar, contextualising, and motivating attributes through its concurrent mix of aural and visual symbol systems, as well as the cultural markers it may display. Additionally, many significant technological factors, discussed in the next section, support these characteristics.

DIGITAL VIDEO - TECHNOLOGY AND PROCESSING CAPABILITIES

As previously stated, Salaberry cautions that "improved technology does not imply improved pedagogy" (2001, p. 39). However, Lindenau also warns that teachers should still be wary of the potentially detrimental effects to their students of not keeping current, with their teaching methods as well as with hardware, software and technology (1984), while at the same time keeping sufficiently informed of their efficient and suitable utilisation (Levy, 2008). Thus, I will now 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), which merely permits a video to be shown to a class, usually only once, front-to-back (Mackey, 2002, p. 174), with little option for replay or other comprehension-facilitating interaction with the content (Gruba, 2006), digital video's 'technology' allows the video 'medium' to become:

- 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 digital video clip can be shared with an unlimited number of people, enhancing potential for individual use in an L2 learning classroom or at home.
- 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.

- Online – digital video is online, making it widely available, free and easily viewable to users with access to the internet through online digital devices.
- A more stable medium – digital video is a more 'stable' medium than previous video formats, allowing for easier replay, tracking and reviewing.

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), has also improved. By interacting with the video, discussing it online with their friends, learners can create more cognitive links regarding a word or phrase, potentially increasing its retention (Draper, 1996). Additionally, due to this technological portability and interactivity, video has become familiar, having achieved 'normalisation' (Bax, 2006, p. 2). Furthermore, 'familiarity', or 'cognitive fluency', states that people are more ready and eager to process information in ways that are familiar to them (Alter, 2009, p. 2; Sherman, 2008, p. 28). As many students in modern societies view videos at home and on the way to school, video viewing literacy may be positively influenced in this regard.

Processing Capabilities

Digital video, therefore, has become a highly portable, viewable, interactive and familiar medium due to its technological features. Additionally, digital video also possess important processing capabilities – i.e., the ways that the medium (video) can be manipulated through the technological processes (digital).

As discussed in the previous section, more so than comic books (visual), books (abstract-visual, linguistic) and radio (aural), it is due primarily to video's unique blending of both audio and visual streams that it is considered by many media researchers to best capture the contextualised reality experienced by humans on a day-to-day basis (Alter, 2009; Tschirner, 2001). However, while video may possess the same symbol systems as a "televised broadcast", video has several processing capabilities which effect 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 be paused, searched, rewound, and reviewed with or without subtitles.

This ability to pause, rewind and review information makes the medium 'stable', like a book, instead of 'transitory', like a TV broadcast. Kozma (1991) 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. (1991, p. 5)

Thus, akin to reading a novel, to watch video might occasionally require a 'replay' to affirm understanding (Mackey, 2002). Arguably, compared to previous analogue media, digital video is far more "trackable" (Tschirner, 2001). Pacing, which Wright defines as (1984) the amount of information presented within a set time, can also be an issue with viewing broadcasts. Studies have shown that short-term to long-term information-chunk processing speed is dependent on the length of the chunk and background knowledge. It is therefore possible that the pace of information in transient media can cause comprehension loss (Kozma, 1991). 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 supports cognitive linking (Mayer, 1994). Gruba's investigation shows this; his research illustrates that learners who engage with a selected L2 video are able to build information signposts based on the audio-visual streams, assisted by paralinguistic information (Gruba, 2006).

The enhanced technological features of digital video, combined with its processing capabilities and its features as an information-conveying medium are of potential help to L2 learning. Following on from these points, I will now briefly examine subtitling.

DIGITAL VIDEO AND SUBITLING

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, p. 7; Vanderplank, 1990, p. 222), though "all (variants of subtitles) make different demands on the students linguistic skills and are equally valuable" (Williams, 2000, p. 19). 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" (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 *English Central* use both L2 and L1 subtitles. With more 'stable' videotext, it is thought that many more authentic videos may become more accessible for L2 learners (Peachey, 2008).

Thus, it would appear that subtitling is held to aid L2 learners' acquisition, extrapolation, and the ability to "read" a video's audio track as a "videotext." The next section will consider the cognitive perspective, and how video, and in particular its capacity for subtitling, could potentially help these learners to acquire new L2 lexicon.

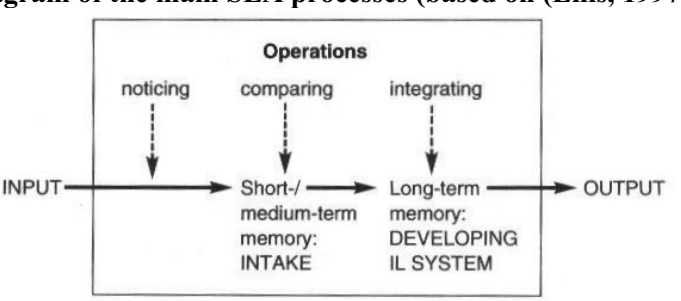
LEXICAL ITEMS, PROPERTIES AND PROCESSING

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 process of 'Noticing.'

Weak Interface System

The 'Weak Interface' system adopts a broad mixture of generally accepted perspectives on the cognitive processes undertaken during L2 acquisition (Pica, 2005, p. 276). First among these is that the Input that an L2 learner is exposed is not necessarily taken into the learner's short-term memory:

FIGURE 1
Diagram of the main SLA processes (based on (Ellis, 1997b))



Reflecting on this system, Ellis (1997b) 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)

As Ellis states, the acquisition of new vocabulary thus begins with the perceiving and noticing of Input; as it logical and reasonable to consider that if a specific lexical item is not perceived by an L2 learner, it will not be taken into the learner's interlanguage system. The importance of this will be discussed below.

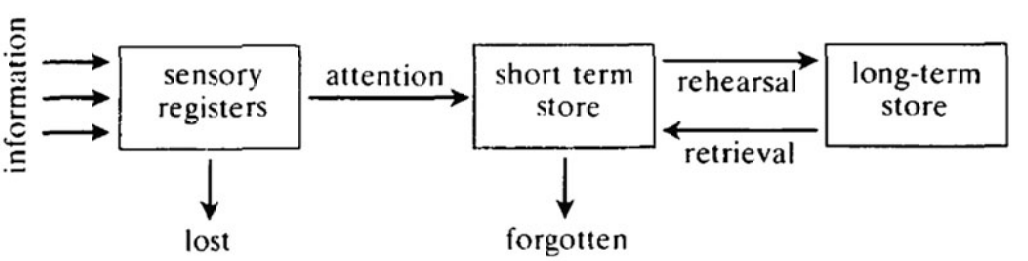
Noticing and Perception

Schmidt's 'Noticing Theory' is another central tenet of the 'Weak Interface' system. It maintains 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).

Referring to 'intention,' and also the distinction between explicit and incidental learning, Schmidt notes that intention is not 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" (1990, p. 139). This, even by itself, could be of importance while 'attention,' the second form of consciousness, involves "noticing the properties of the input" and focuses on "identifying linguistic information (Ellis, 1997a, p. 116)."

The third theory of consciousness, 'awareness,' 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 it is worth noting here that Schmidt defines and treats 'noticing' and 'perception' differently, despite their similarity in meaning. 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, p. 132). The first two stages, *perception* and *noticing*, effectively correspond to the 'Noticing' in the first stage of the cognitive processes of second language acquisition (Fig. 1), while the final *understanding* stage is aligned with the 'Comparing' stage. Schmidt's theory also fits well within most contemporary cognitive memory models:

FIGURE 2
Effect of consciousness in a cognitive memory system model
(based on (Kihlstrom, 1984))



Perception (the recording of information) takes place at the sensory registers. 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 the “detection (*perception)* of a language feature is 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-stream information as a discrete word instead of a series of sounds and noises.

Second, it would be incorrect to overlook the fact that the cognitive processes in the Noticing Theory process can be influenced by several factors in the Input itself. These factors can are summarised as ‘task demands,’ ‘frequency,’ ‘unusual features,’ ‘salience,’ ‘interactional modification,’ ‘existing linguistic knowledge’ (Ellis, 1997a, p. 120). Of particular interest here is the ‘salience’ of words, a concept briefly noted in the previous section on subtitling. Word ‘salience’ can often be increased by focal accent, volume and extended pauses between words (Carroll, 2006, p. 19). Carroll further adds that:

> Possibly 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)

While potentially improving the chance that a word-sound may be heard, neither ‘salience’ itself nor these other factors guarantee that a certain linguistic feature will be *perceived*, *noticed* and then transferred to the learner’s short-term memory store as Intake. However, research conducted into Noticing Theory suggests that lexical items may be highlighted through several techniques (Carroll, 2006; Ellis, 1997a; Pica, 2005).

IMPLICATIONS AND RESEARCH QUESTIONS

The theories and literature reviewed in the previous sections are summarised, for the purposes of brevity, as follows. Investigations into video as an information-conveying medium, with specific reference to foreign language learning and language acquisition, consider video to have a high level of potential due to its realistic, familiar blend of simultaneous audio and visual data (Sherman, 2003; Willis, 1983). The visual information stream supports the audio through context and content (Tschirner, 2001), while native authenticity both motivates and provides the viewer with both linguistic and paralinguistic information (King, 2002).

Furthermore, subtitles grant stability (Gruba, 2004; Mackey, 2002), facilitate saliency (Gruba, 2004), aid organisation and also function as memory “pegs” with which to assist perception, comprehension and retention of form and meaning (Kozma, 1991). Subtitles are held to be mainly facilitative 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" to remind and re-activate vocabulary, or allow new acquisition. Schmidt's theories on noticing remind us that if a word is not first perceived, it not become Input (Robinson, 1995; Schmidt, 1990), and that word saliency, which can be enhanced by subtitling, positively facilitates the noticing of vocabulary in input (Carroll, 2006).

In addition to these findings, the results from the previous research on this topic also need to be taken into consideration (Lees, 2014). The results from this investigation are displayed in Table 1.

The results show that 'Karaoke'-style subtitles appear to facilitate incidental vocabulary recognition to a higher degree than both standard-subtitles and no subtitles, although analysis demonstrated there was not a substantial statistical difference. This research received feedback from several teachers and researchers at seminars and tech-focused workshops; aside from the need, frequently voiced, for both a larger sample-size and the need to replicate the experiment with a different video, much of the feedback questioned whether or not unknown or unfamiliar target vocabulary could be focused on through dual-subtitle intervention, akin to Bayon's (2004) research.

TABLE 1
Collected Results from the Previous Word Perception Test

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

In attempt to answer both of these points, a different video-clip using low-frequency vocabulary items was employed, and a further stage of subtitle intervention was added. I continue to hold that the degree to which further subtitle intervention, such as that provided by synchronised 'karaoke'-style subtitles (and potentially by restrained use of dual-subs to focus on target vocabulary), could positively influence L2 learners' ability to perceive and notice individual vocabulary items warrants further investigation. In doing so, this study seeks to answer the following research questions:

1. 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?
2. To what degree do synchronised 'Karaoke'-style L2 subtitles (with L1 translations on the target vocabulary) on a digital video-clip facilitate or hinder incidental word-perception?
3. 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. I will 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

The participants for this small-scale investigation were drawn from 52 intermediate-level learners of English, currently studying in an elective general-English course at a Japanese university. The participants' agreed to watch a 15-minute clip from the British TV show "Doctor Who." As discussed in the literature review this clip was chosen due to its authenticity (i.e., it is a native-speed spoken-English TV programme, aimed primarily at teenage native-speakers) as well as its balance of vocabulary. 94% of the words in the script are in the first 2000 words of the General Service List, with only a cumulative 6% from the Academic Word List and other off-list, low frequency vocabulary items. In an effort to follow the advice received from previous feedback, the words to be tested were chosen primarily from amongst these lower-frequency words whilst also being less salient. These words are listed in Table 2 below.

TABLE 2
Target Vocabulary and Word-Frequency

Target Word	Frequency	Target Word	Frequency
1. emergency	1734	11. ***fairly***	*2170*
2. ordinary	2282	12. ***same***	*161*
3. craving	12711	13. solid	2021
4. ***lobby***	*3764*	14. draft	2640
5. soaking	29325	15. ***backwards***	*6816*
6. perfectly	2451	16. split	6726
7. ***scanner***	*8335*	17. breach	8001
8. disgusting	9122	18. ***speak***	*335*
9. rubbish	15524	19. snap	2613
10. poison	7472	20. brand new	2586

To investigate the influence on the incidental word-perception of these target words, the participants were divided into four groups: a) a no-subtitle group, b) a standard L2-L2 subtitle group, c) a 'karaoke'-style L2-L2 subtitle group, and d) a 'karaoke'-style L2-L2 subtitle group with the target words translated into the participants' L1, Japanese.

The participants were introduced to the "Doctor Who" TV show, and were instructed to concentrate while they watch the video-clip back-to-back. After the first viewing, the participants undertook a 20-word recognition test to determine which words they "noticed" in the video clip, with 14 of the words present, and 6 of the words absent, in order to test their incidental perception and noticing of these target words (Appendix 1: Word Perception Test). The participants then answered a short questionnaire, which sought their opinions of the video-clip that they watched and the degree of usefulness of the subtitle intervention (Appendix 2: Survey A). Finally, they watched a shortened clip of the 'Karaoke'-style subtitled video to compare to the version that they previously watched, and their opinions were collected in a second survey (Appendix 3: Survey B). Unfortunately, three of the participants slept through the experiment, so in accordance with Schmidt's attention theory their data was discarded.

RESULTS

Word Noticing and Perception Tests

The data gathered from the Word Noticing and Perception Tests of this study and from previous research was collated and displayed in Table 3 below for comparison purposes.

TABLE 3
Collected Results from the Word Perception and Noticing Tests

Video-Type	**Previous Research** *Correct Answers*
No Subtitles (n=14)	9.74
Standard Subtitles (n=15)	11.16
Karaoke Subtitles (n=13)	13.07
	Current Research *Correct Answers*
No Subtitles (n=12)	10.18
Standard Subtitles (n=13)	12.25
Karaoke Subtitles (n=12)	13.27
Karaoke+Dual Subtitles (n=12)	14.27

As the results show, the general trends from both investigations suggest that, grading from non-subtitled video to synchronised-subtitled video, a higher degree of subtitle intervention leads to a higher degree of incidental word perception and noticing. 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, especially given that the previously collected data did not show clear statistical difference, they do suggest that the 'Karaoke'-style subtitles have a facilitative effect on incidental lexical noticing.

The data from this current investigation put into an array for comparison and were run through independent-samples T-tests. Due to the number of tests to run, the results are displayed in Table 4 below, and the significantly different results shaded for emphasis.

The results below illustrate that much of the data shows a statistically significant degree of difference. For example, when comparing the No Subtitles group with all of the subtitled groups (Standard Subtitles, Karaoke Subtitles and Karaoke+Dual Subtitles), the independent T-Test results return values of under p=0.05 for all groups; p=0.047, p=0.001 and p=0.000 respectively. This strongly suggests that there is little chance that the differences between the No Subtitles results and the subtitled groups are due to coincidence.

Between the Standard Subtitles group (M=12.25, SD=2.70) and the Karaoke Subtitles group (M=13.270, SD=2.37), however, despite the higher average of correct answers in the Karaoke Subtitles group the t-test results show no significant difference; t(21)= 1.22, p=0.235. The data does suggest a statistical difference between the Standard Subtitles group (M=12.25, SD=2.70) and the Karaoke+Dual Subtitles group (M=14.27, SD=2.95), with a t-test result of t(21)= 2.21, p=0.031. In this case, perhaps the increased subtitle intervention, with the target words translated into the participants’ L1 during the speech utterance, proved to be an important factor in assisting incidental perception.

TABLE 4
Independent T-Test Results for the Word Noticing and Perception Tests

	No Subtitles (M=10.18, SD=2.29)	**Standard Subtitles (M=12.25, SD=2.70)**	**Karaoke Subtitles (M=13.27, SD=2.37)**	**Karaoke+Dual Subtitles (M=14.27, SD=2.95)**
No Subtitles (M=10.18, SD=2.29)		t(21)= 2.11, p=0.047	t(20)= 3.72, p=0.001	t(20)= 4.68, p=0.000
Standard Subtitles (M=12.25, SD=2.70)	t(21)= 2.11, p=0.047		t(21)= 1.22, p=0.235	t(21)= 2.21, p=0.031
Karaoke Subtitles (M=13.27, SD=2.37)	t(20)= 3.72, p=0.001	t(21)= 1.22, p=0.235		t(20)= 1.45, p=0.162
Karaoke+Dual Subtitles (M=14.27, SD=2.95)	t(20)= 4.68, p=0.000	t(21)= 2.21, p=0.031	t(20)= 1.45, p=0.162	

The data gathered from the current investigation fits the trend, if not exactly to the letter, of the findings of the previous research; mainly, that the groups who watched video-clips with synchronised subtitles, in this case both the Karaoke and the Karaoke+Dual subtitles, are shown to have empirically outperformed both the No Subtitles and Standard Subtitles groups.

Impressions Survey and Comments

Analysis of the qualitative surveys also revealed some interesting information, which is displayed in the tables below. The first table shows the results from Survey A (Table 5), which was administered directly after the Word Perception and Noticing Test. The responses were separated based on the version of the video the participant watched. The highest results have been shaded in order to highlight the general trends.

TABLE 5
Collected Results from the Initial Impressions Survey, Survey A

NO SUBTITLES	***(1)*** *Strongly*	***(2)***	***(3)***	***(4)***	***(5)*** *Strongly*
I thought that the video	*Disagree*	*Disagree*	*Neutral*	*Agree*	*Agree*
a) was interesting		1	5	3	3
b) was easy to understand	1	5	3	3	
If your video had subtitles...					
c) subtitles were helpful					
d) subtitles were unhelpful					
STANDARD SUBTITLES	***(1)*** *Strongly*	***(2)***	***(3)***	***(4)***	***(5)*** *Strongly*
I thought that the video	*Disagree*	*Disagree*	*Neutral*	*Agree*	*Agree*
a) was interesting		1	2	6	4
b) was easy to understand		3	4	5	1
If your video had subtitles...					
c) subtitles were helpful		1		4	8
d) subtitles were unhelpful	10	2	1		
KARAOKE SUBTITLES	***(1)*** *Strongly*	***(2)***	***(3)***	***(4)***	***(5)*** *Strongly*
I thought that the video	*Disagree*	*Disagree*	*Neutral*	*Agree*	*Agree*
a) was interesting			1	7	4
b) was easy to understand			4	8	
If your video had subtitles...					
c) subtitles were helpful				3	9
d) subtitles were unhelpful	9	3			
KARAOKE+DUAL SUBTITLES	***(1)*** *Strongly*	***(2)***	***(3)***	***(4)***	***(5)*** *Strongly*
I thought that the video	*Disagree*	*Disagree*	*Neutral*	*Agree*	*Agree*
a) was interesting			2	4	6
b) was easy to understand		1	3	5	3
If your video had subtitles...					
c) subtitles were helpful			1	4	7
d) subtitles were unhelpful	7	4	1		

The initial impressions of the video-clip, and the subtitles (if present), show the general trend that the subtitled videos were thought to be both "more interesting" and "easier to understand" than the No-Subtitled video-clip.

Table 6 below displays the results from Survey B. This survey, as explained previously, was administered after all of the groups watched the Karaoke Subtitles version of the video together, and as such seeks their impressions comparing the video-clip that they had watched first with the Karaoke Subtitles clip.

TABLE 6
Collected Results from the Comparison Survey, Survey B

SURVEY B RESULTS	***(1)*** *Strongly*	***(2)***	***(3)***	***(4)***	***(5)*** *Strongly*
I thought that the Karaoke Subs video	*Disagree*	*Disagree*	*Neutral*	*Agree*	*Agree*
a) was easier to listen to		5	12	19	13
b) was easier to understand		7	7	21	14
c) was easier to read		5	13	14	17
d) helped me to better notice words		2	14	16	17
	Karaoke+Dual		*Karaoke*		*Standard*
e) Which video did you prefer?	8		22		19

The results in from this survey show a somewhat favourable impression of the Karaoke Subtitles, though clearly, as the spread of the results show and the fact that 19 out of the 49 participants reported preferring Standard Subtitles, the synchronised style of subtitling is far from being seen as an unequivocal improvement. Indeed, as with the empirical results from the noticing and perception test, while the Karaoke Subtitles proved to be more facilitative, they were not categorically so. Similarly, the participants' impressions and preferences appear to reflect the same.

This sentiment is also echoed by the comments volunteered in the open question section at the end of Survey B. Many of the participants wrote about how they preferred the Standard Subtitles:

「Karaoke jimaku no supiido to jibun ga sono jimaku wo yomu supiido ga chigau tam, sukoshi rikai suru no ga okureteshimau」*"Because the speed that I can read the Karaoke subtitles and the speed of the subtitles themselves is different, my understanding comes a little bit late."*

"I think there are only [a] little bit different, Karaoke and Standard subtitles."

「Karaoke jimaku ha ima doko no bubun wo hanashiteiru no ka, tango tanni de wakatte sono ten ha yokatta ga, imi wo bunshou toshite jyanaku hitotsu hitotsu ni kangaeteshimau ki ga suru」*"The Karaoke Subtitles were good because I was able to see which part they were speaking, but regarding the meaning I think I might tend to think of it word-by-word rather than a sentence."*
「Karaoke jimkau ni ki ga torareru」*"The Karaoke Subtitles distract my focus."*

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

「Iroduke ni me ga itai no de, futsuu jimaku yori karaoke jimaku ni ki ni natta」*"The different colours hurt my eyes, so the Karaoke Subtitles bothered me more than the Standard Subtitles."*

However, a large number also supported the Karaoke Subtitles, as evidenced by the following remarks:

「Karaoke jimaku no hou ga, rikai shiyasukatta desu. Doko wo hanashiteiru no ka ga wakaru kara da to omoimasu」 *"The Karaoke subtitles were easier to understand, because I felt I understood where and when they were saying something."*

「Eizou to jimaku to miteiru to jimaku no doko wo shabeteiru no ka ga wakaranaku naru koto ga ookatta no de karaoke jimaku no hou ga wakariyasui to kanjita」 *"While watching the screen and the [Standard] subtitles I lost where they were speaking many times, so I felt that the Karaoke Subtitles were easier to understand."*

「Gakushuu no tame ni miru nara karaoke jimaku de imi wo tsukande, jimaku nashi de mitai to omou」 *"If I were to watch to for learning, I would like to use the Karaoke Subtitles to grasp the meaning, and then watch without subtitles."*

「Watashi ha, karaoke no hou ga yori wakariyasui no de yakunitatsu to omoimasu. Tango wo miteoboeru renshuu ni mo naru shi, bunshou no hatsuon ya supiido ga wakarimasu. Mata, karaoke jimaku no bideo wo mitai desu」 *"For me, the Karaoke Subtitles were easier to understand so I thought that they were useful. I can practice looking at the words and remembering them, and I can understand the sentence speed and pronunciation. I want to watch videos with this kind of subtitles again."*

「Karaoke jimaku de ima dono bubun wo kiiteiru no ka ga hitome de wakaru no de yomiyasui」 *"With Karaoke Subtitles I can see at a glance which part I am hearing, so they are easy to read."*

"I would like to watch more videos like this, because it was easier to read and understand the story. And, Karaoke style [subtitles] makes us learn the pronunciation of each word."

「Karaoke jimaku ha totemo yakunitatsu to omou. Moshi futsuu no eiga ni mo kono kinou ga areba, eigo no benkyoni naru to omotteta. Karaoke jimaku to yaku, mata ha karaoke jimaku nomi nado, sentaku dekiru DVD ga ii desu」 *"The Karaoke Subtitles were very useful, I think. If a normal movie had this function, I thought it might help English learning. A DVD where you can choose Karaoke+Dual Subtitles, or even just Karaoke Subtitles, would be good."*

According to the impressions of the participants, it would seem that the 'Karaoke'-style subtitles, despite their unfamiliar, experimental and sometimes "focus taking" nature, were relatively well received by the participating English learners. Additionally, as supported by the comments, many of the participants reported that the Karaoke Subtitles enabled them to better "hear" or "see" or "read" when specific

words in the native-speed audio stream were being said, suggesting that the word salience was favourably improved by this synchronised style of subtitling.

CONCLUSION

In summation, the results gathered, analysed and summarised above, would seem to suggest that the 'Karaoke'-style L2 subtitled digital-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. This matches the trend of the results from the previous investigation, and hopefully the fact that this replication of the experiment yielded similar results should encourage further exanimation of the potential of subtitled digital-video in the future. The results from the impressions surveys also suggest that the 'Karaoke'-style subtitles were generally received positively, though as with other research there remain issues of distraction and pacing to consider.

Additionally, in this investigation, the 'Karaoke'-style subtitles in this study were able to help participants achieve a statistically significant improvement in word perception and noticing when compared to the experiment participants who view the non-subtitled video, revealing that the synchronised subtitles could have enhanced some of the saliency properties of standard L2 subtitles. Despite this, and despite displaying an overall higher average of correctly-noticed words in the word perception test, the Karaoke Subtitles group did not perform unequivocally better than the Standard Subtitles group.

However, the newly added Karaoke+Dual Subtitles, those with the added target vocabulary translated into the participants' L1, demonstrated not only the highest performance on the word perception test but also a statistically significant improvement over both the non-subtitled and standard-subtitled video groups' results. Perhaps if this result can be examined further and hopefully replicated in future investigations, it might eventually facilitate foreign language learners' use of digital video for study purposes.

There were several points that require further attention in this study; chiefly, the small sample size (n=49) does not lend much statistical validity to the findings of this investigation. Furthermore, given the time constraints, and also that three of the participants fell asleep, it is quite possible that video-clip used was a little too long for experimental purposes. However, in spite of these points to consider for future studies, the results of the investigation do seem to be promising.

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). Intergrative 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. *Language 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.

Kihlstrom, J. (1984). Conscious, Subconscious, Unconscious: A Cognitive Perspective. In K. S. Bowers and D. Meichenbaum (Eds.), *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., 25*, 15-35.

Lees, D. (2014). To What Extent Can 'Karaoke'-style Subtitles On Digital Video Help Learners Of English As A Foreign Language? *Kwansei Gakuin University: Language Center Annual Research Report, 17*, 57-74.

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.
OUP. (2015). ELT Catalogue: Stretch Series Textbooks Retrieved 30/12/2015, 2015, from https://elt.oup.com/catalogue/items/global/adult_courses/stretch/?cc=global&selLanguage=en&mode=hub
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. (2003). *Using Authentic Video in the Language Classroom*. Cambridge: Cambridge University Press.

Sherman, J. (2008). Where are we with Media Video?, *IATEFL CALL Review*.
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.
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.
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 & Continuity of Television Programs: Effects on Childrens Attention & Comprehension. *Developmental Psychology, 20*.

APPENDIX 1: Word Perception Test

Look at the words below. Did these words appear in the video? Circle "Yes (Y)", "No (N)" to answer.
下にある言葉を見てください。ビデオに出て来たか、「Y」（はい）、「N」（いいえ）と丸にして答えてください。

	はい	いいえ		はい	いいえ
1. emergency	Y	N	11. insane	Y	N
2. ordinary	Y	N	12. same	Y	N
3. craving	Y	N	13. speak	Y	N
4. lobby	Y	N	14. draft	Y	N
5. soaking	Y	N	15. backwards	Y	N
6. perfectly	Y	N	16. split	Y	N
7. scanner	Y	N	17. fairly	Y	N
8. disgusting	Y	N	18. breach	Y	N
9. rubbish	Y	N	19. snap	Y	N
10. poison	Y	N	20. brand new	Y	N

Notes:

APPENDIX 2: Survey A

1. What did you think of the video? Please answer the questions below by using the scale: このビデオはどう思いますか？下にある質問を答えてください。

I thought that the video…
見たビデオは。。。と思います

	Strongly disagree 強く賛成しない					Strongly agree 強く賛成する
a) interesting 面白かった	1	2	3	4	5	
b) was easy to understand 分かりやすかった		1	2	3	4	5

If your video had subtitles / もし見たビデオは字幕があったら

c) subtitles were helpful 字幕は役に立つ	1	2	3	4	5	
d) subtitles were unhelpful 字幕は役に立たない		1	2	3	4	5

2. Imagine that you were assessing whether the video style would be suitable for helping learning English in classes or in your free time. Please write down your impressions and feelings about the video, in your own words. Was the English too quick to understand? Were the subtitles helpful? Would you like to watch more videos like this?
あなたはこんなビデオ・タイプが授業でも自由な時間でも英語を勉強することに適切かどうか調査していることを想像してください。自分の言葉でさっきに見たビデオの考えと印象を書いてください。こんなビデオに出た英語は理解するのに速すぎましたか？ぜひ、自由に書いてください。

Notes:

APPENDIX 3: Survey B

3. You have now seen the video twice. The second viewing used a “Karaoke Style” subtitling method. Which video type do you generally prefer out of the following?
これでビデオを二回見たことがあります。二回目のほうが「カラオケ字幕」です。
一般的に考えば、これらのビデオ・タイプの中で、どれが一番好きですか？

Dual+Karaoke ( )　　Karaoke ( )　　Standard ( )　None ( )
通訳+カラオケ字幕　カラオケ字幕　普通字幕　字幕なし

4. What did you think of the “Karaoke Style” video? Please answer the questions below
この「カラオケ字幕」ビデオはどう思いますか？下にある質問を答えてください

I thought that the “Karaoke Style” video…
見た「カラオケ字幕」ビデオは。。。と思います

	Strongly disagree 強く賛成しない				Strongly agree 強く賛成する	
a) was easier to listen to より聞きやすかった	1	2	3	4	5	
b) was easier to understand より分かりやすかった		1	2	3	4	5
c) was easier to read より読みやすかった		1	2	3	4	5
d) allowed me to better notice individual words 一つずつの言葉を見やすくしてくれた	1	2	3	4	5	

5. Please write down your impressions and feelings about the video, in your own words. Were the subtitles helpful? Were the subtitles confusing? Would you like to watch more videos like this?
自分の言葉でさっき見たビデオの考えと印象を書いてください。こんなカラオケ字幕は役に立つと思いますか？また、こんなカラオケ字幕のついたビデオをもっと見たいと思いますか？ぜひ、自由に書いてください。

Notes: