The Role of Visual Elements in Understanding News Videotext: the Case of Syrian EFL Learners at the Intermediate Level

Dr. Dima Dayoub[*]

□ ABSTRACT □

This study explores how Syrian EFL students at the intermediate level comprehend news videotext. It particularly focuses on the role of visuals in comprehension. To this end, three groups totaling 41 students of English as a Foreign Language at the Higher Institute of Languages viewed BBC news clips during three 45-minute sessions. News videotexts were repeatedly paused, and participants were invited to write down immediately retrospective individual reports on what elements of the news videotext helped or hindered comprehension. Data analysis was enriched by principles of multimedia learning. Overall, the mediating role of the visual elements in comprehension was manifest in the data. Evident too was the interplay between audio and visual processing of news content, which underlines the need for strengthening learners' media literacy as part and parcel of news videotext comprehension.

Keywords: news videotext, multimedia, dual coding theory, introspective methods of data collection

[*] **Assistant Professor Educational Technology & ELT at the English Language Teaching department, the Higher Institute of Languages, Tishreen University**
Email: dima.dayoub@tishreen.edu.sy

دور مكونات الفيديو البصرية في فهم فيديو الأخبار: حالة الطلاب السوريين في المستوى المتوسط للإنكليزية كلغة أجنبية

د. ديما ديوب*

□ ملخّص □

تتحرى هذه الدراسة استيعاب طلاب اللغة الإنكليزية السوريين في المستوى المتوسط لمقاطع فيديو الأخبار وتركز بالأخص على دور العناصر البصرية في الاستيعاب. لهذه الغاية تم عرض ثلاثة مقاطع من أخبار 546 لبي بي سي على ثلاث مجموعات مكونة بالإجمال من 41 طالب و طالبة لغة انكليزية في المستوى المتوسط قي المعهد العالي للغات وزعوا على ثلاث جلسات مدة كل منها 45 دقيقة. تم إيقاف مقاطع فيديو الأخبار بشكل متكرر وطلب من كل مشارك كتابة تقارير فورية فردية إسترجاعية فيما يتعلق بعناصر الفيديو التي ساعدت في فهم فيديو الأخبار أو حالت دون استيعابه. أغنت مبادئ التعلم عبر الوسائط المتعددة عملية تحليل البيانات و بالخلاصة فقد كان الدور الوسيط للعناصر البصرية في الفهم جليا" كما كان من الواضح وجود تفاعل بين المعالجة البصرية و السمعية لمحتوى الأخبار مما أكد على ضرورة تعزيز ثقافة المتعلمين في فهم الوسائط المتعددة كجزء لا يتجزأ من استيعاب مقاطع فيديو الأخبار .

الكلمات المفتاحية: مقاطع فيديو الأخبار ، تعددية الوسائط، نظرية الترميز المزدوج، طرق جمع المعطيات الاستبطانية

***مدرس ، قسم تعليم اللغة الإنكليزية، المعهد العالي للغات ، جامعة تشرين**
البريد الالكتروني: dima.dayoub@tishreen.edu.sy

1. Introduction: The Research Objective:

This study explores how intermediate students of English as a Foreign Language at the Higher Institute of Languages at Tishreen University make sense of the visual elements of BBC news videotext and whether news visuals, and other multimedia components, help or hinder the comprehension of news content, according to students' accounts. Inextricably interwoven with comprehension is the notion of media literacy that has become multimodal and multidimensional in "the emergence of post-typographic forms of text production, contribution, and reception" (Koltay, 2011: 211). Increased accessibility to major television news channels on social forums such as FaceBook and YouTube and the rise of citizen journalism where news clips are user-generated and shared on such generic forums underline the need to cultivate comprehension skills, at the deep level of critical analysis. This is particularly challenging when the presented material is a real-time, multimodal authentic video posing extra processing demands on the memory. Promoting comprehension of vivid visually-mediated news videotext is critical, yet research in this area remains scant in general and in the Syrian context in particular. What follows sheds light on current literature on multimodality and its relevance to learning before detailing the research questions, methods and findings.

2. Literature Review:
2.1. Multimodality:

Simply put, a "multimodal interactive system is a system that employs at least two different modalities for input and/or output" (Bernsen, 2008: 9). A modality is a "way of representing information in some physical medium" and each modality varies in terms of its "expressiveness" and the "properties of the recipient of information represented" (*ibid.*, 7-8).

A multimedia presentation refers to any combination of words (written on-screen text that people read or spoken text that people listen to) and graphics (static such as images and charts or dynamic such as animation and video) (Clark & Mayer, 2016: 70). Different modes for presenting information can be combined in different ways to achieve a particular purpose. Modality redundancy refers to the presentation of more or less the same information in two different modalities, which is useful to reinforce the presented information. In modality replacement, one modality is replaced by another to achieve functionality as in substituting spoken discourse by speech-to-text for the deaf and hard of hearing (Bernsen, 2008: 27). Even when modality redundancy is an inbuilt feature of a presentation, one modality may be filtered out as a result of a disability or intentionally via the controls as in turning on and off captions and subtitles or muting the sound.

2.2. News Videotext as a Multimodal Text:

Videotext is defined as "multimodal text consisting of contiguous, dynamic, and interwoven sounds (verbal, musical and/or background) and visual images (still, moving, text and/or graphic) which can be presented using a range of media". News videotext is distinct from other videotext genres in terms of its highly structured informative nature (Cross, 2011: 45-46).

2.3. Multimodal Combinations and Learning:

It is contended that combining two or more modalities creates a sum of their expressive strengths that is larger than their individual affordances. A multimodal experience also

affords accommodating individuals who have different perceptual, cognitive, and emotional capabilities and diverse competencies, skills and preferences (*ibid.,* 24). In multimedia presentations, deeper learning happens as learners actively engage in making associations between pictorial and verbal representations. This is referred to as the multimedia principle (Clark & Mayer, 2016: 70-71) and is rationalized in the light of dual coding theory.

According to dual coding theory, human cognition comprises two subsystems: verbal for processing the language and non-verbal, also called imagery, for processing and generating mental images. Although these subsystems are functionally independent in the sense that each can be active on its own or simultaneously with the other subsystem, they are functionally dependent in the sense that "activity in one system can initiate activity in the other" (Paivio، 1990: 53-54). This means that a learning material that presents linguistic and graphical information or stimulates mental imagery or imaginal processing can aid the mnemonic power of memory. The "modality-specific nature" of dual coding theory evident in its focus on how "separate verbal and nonverbal mental representations are collectively processed" (Cross, 2011: 45), makes it an apt theoretical ground to understand audio-visual processing. That said, multimedia presentations do not automatically lead to better learning experiences, as will be next detailed.

To start with, Clark and Mayer (2016: 79-84) argue that the effectiveness of graphics depends on a number of factors including the relevance of the graphic, its non/dynamic nature, and the learner's proficiency level. A decorative illustration (a neutral illustration that is not relevant to instructional goals) has no effect on learning while an "instructive illustration" (one that is directly related to instructional goals) aids learning. High-knowledge learners might benefit better from one mode of presentation than low-knowledge learners. Static graphic frames are more engaging to learners as they stimulate them to mentally animate the presented information without adding extraneous cognitive load. As such, static frames are more suited for instructional purposes unless the content requires dynamic representations.

Moreover, concurrent audio and visual messages do not always completely match, which has an impact on comprehension. The degree of correspondence between audio and visual messages is classified according to the correspondence continuum proposed by Walma van der Molen (2001). The continuum comprises four major categories: direct when the visual and concurrent audio information semantically overlap, indirect when audio and visual information thematically overlap as in standard scenes and hand gestures, indirect where audio and visual information mismatch or diverge, and talking head when the visual is the top half of the news anchor. These categories provided a reference framework for some studies on the comprehension of news videotext. For example, Cross draws on the correspondence continuum to examine the audio-visual characteristics of BBC news videotexts and concludes that the direct category aided comprehension, the indirect one had a positive impact on comprehension, the convergent one had an adverse effect on comprehension, while the talking head category had no impact on comprehension (2011: 64).

To take visual elements as facilitative of listening comprehension is a reductionist approach that inhibits and downplays the learner's critical thinking skills. A better alternative proposed by Gruba is to "inform students that visual elements offer *potential opportunities* for developing understanding in tandem with verbal elements" (2006: 87). As such, the visual modality, in terms of its pedagogical affordances, is on par with the verbal one. Having said that, visual elements can also be verbal as is the case of captions.

2.4. Captions and Comprehension:

The effect of multiple-mode input on the development of various language skills has been the focus of a growing body of research and extends beyond the effects moving or still images have on comprehension. For example, there has been a rising surge in the number of studies concerned with the impact of captions on comprehension. Some studies conclude that captioned videos, as opposed to non-captioned ones, are associated with "novel vocabulary recognition" and "overall comprehension of the videos" (Winke et al. 2010: 79). Others report that captions fall short of facilitating the transformation of audiovisual input into a comprehensible one and attribute this finding to the difficulty of catching up with reading in L2 within given time constraints (Hosogoshi, 2016: 168). Questions have also been raised regarding the transferability of caption-assisted listening skills to real-life authentic listening settings where students have to rely solely on "acoustic signals and relevant contextual factors" (Vandergrift, 2004: 7&9). Although this is a valid concern, this research aims to foster listening to the news videotext genre, which is multimodal and pictorially mediated. Additionally, news clips do not include word-by-word transcriptions of the spoken word streams; therefore, news captions do not compromise compensatory strategies and the real-time demands of information processing.

3. The Research:
3.1. The Research Questions:

The broad research question is: *How do EFL Syrian students at the intermediate level at the Higher Institute of Languages in Tishreen University make sense of visually mediated BBC news videotext?* This question was broken down into the following sub-questions:

1. Do moving and still images play a role in aiding or hindering comprehension of news videotext, according to students' reports?
2. Do captions play a role in aiding or hindering comprehension of news videotext, according to students' reports?
3. What other multimedia elements do students mention as salient in the process of news videotext comprehension?

Throughout the discussion any evidence of media literacy that students demonstrate in handling the videotext will also be referred to.

3.2. The Research Sample:

Two intermediate classes at the Higher Institute of Languages were approached for this study. Students were already grouped according to their proficiency based on a placement test they had taken at the start of their courses. There are two reasons why students at the intermediate level were selected for this study. Firstly, they are less likely to encounter linguistic complexity than lower-level proficiency learners. Secondly, they are more capable of verbalizing internal cognitive processes such as their strategies of making associations between audio and visual messages than lower-ability learners or higher-level learners since the latter might automatically employ such strategies (Charters, 2003: 72; Gruba 2006: 81).

3.3. Data Collection Methods:
3.3.1. Introspective Methods:

For data collection, this study used introspective methods defined as "the various ways of soliciting self-reflections from respondents" "about unobservable mental processes such as thoughts, feelings, and motives" (Dörnyei, 2007: 147). In other words, an introspective

method is metacognitive practice: thinking about thinking. The tool used in this study is analogous to the think-aloud technique in terms of its use in tandem with the language-based task of viewing news videotexts. It diverts from it in the output of participants as they were requested to produce written reports of their thinking processes. Written reports might elicit more data since Syrian students traditionally tend to feel more at ease writing than speaking, especially since they were invited to provide accounts on lack of comprehension, not only on successful comprehension. Anonymous written reports also ensure that respondents do not provide answers that are necessarily desirable by the researcher. While display, the selected videotexts were also broken down into shorter streams by pausing so as to guard against overwhelming the participants' working memory (Charters 2003:72),

3.3.2. Piloting:

Two free-to-air authentic BBC news videotexts were downloaded: one about a crime in an English village and the other about the 2019 Oscar ceremonies. In the first session, the 3-4 minute long videotexts were displayed in class pausing at regular intervals so as allow the students enough time to write down responses to the questions :
1. What is the video about? and
2. What helped you understand it? What made it difficult to understand? (images, sounds, captions, body language, music, other, or a combination of these multimedia elements).

It was stressed that the "how" is the main interest of the research and that writing in Arabic was welcome if participants encountered any difficulties in expressing their thinking processes in English. Students were encouraged not to adhere to the prompts provided but to take their answers in any directions they wished since there were no right or wrong answers.

Some alterations were made in the second session, which was run with another class: the Oscar news videotext was replaced by a recent news clip on an Ethiopian plane crash that contained clearer language and was slower-paced than the other two videos. More prompts were provided verbally and collectively prior to the immediately retrospective written responses so as to allow participants to brainstorm. The videotext was re-played this time to encourage learners to get familiar with the task and feel more at ease with writing down their thinking processes. 17 students took part in the pilot; 18 others participated in the second session.

Finally, six students who had already participated in session one volunteered in a further data collection stage since they felt more knowledgeable about the research objectives and more capable of expressing their thinking processes. As such the total number of respondents was 35 while the total number of responses was 41. No comparisons were made between the six volunteers' answers in the first and last stages of data collection, since the aim of this research is not to explore the development of the participants' processing of the news. In addition, the time between the two data collection sessions was too short to allow any significant changes in the participants' processing strategies and no instructions on how to process audio-visual information was provided to them. In sum, the extra session simply broadened the scope of the data gathered. The six responses were allocated the numbers 36-41.

3.4. Analysis:

It is not possible to isolate multimedia elements in authentic vivid videos such as news clips, neither does isolation mirror the naturally complex comprehension processes that characterize daily interactions with the news. Analysis, however, was restricted to the

mediating role of the visual modality in comprehension and on the interplay between visual and audio processing, as surfaced in the responses. The term visuals in this study refers to moving graphics or video, still images and captions.

Analysis was carried out inductively in the sense that all responses were read systematically and classified manually into headings in order to generate in-house coding categories. Thematic analysis was conducted to move from the stage of coding to that of thematizing. A theme is a category that stems from the generated codes and relates to the research focus. The core theme in this study was the influence of visual elements on comprehension. Each visual modality (moving and still images, captions, facial expressions) was tabulated with specific details of:

1) its *perceived utility* in comprehension (evident in using such words as "helped", "don't/can't understand", "understand slightly", "a bit/difficult", etc), and
 2) extracts from the participants' reports.

Relationships between the perceived utility of the specified visual element and cognitive load were significantly informed by core principles of multimedia learning, mainly the reverse modality principle, the redundancy principle and the signalling principle – all will be thoroughly discussed in the analysis section.

3.5. Discussion and Findings:
3.5.1. Graphics as a Substitute for Audio Information:

One aspect of visual literacy documented in the literature is students' ability to re-direct attention to aural processing only when the visual input negatively interfered with understanding. This was reported to be triggered by mismatching visual and aural elements in Gruba's research (2006; 86). Similarly, in this study, one student reported abandoning viewing and attending to listening only:

It was difficult for me to concentrate about the picture and the sound together. I just listened. I understand only because of sound. I couldn't link between the picture and the sound. It was just difficult. (participant 30)

Conversely, 15 out of 17 students reportedly resorted to visuals when the audio narration of the Soham crime news videotext was challenging: fast-paced and/or difficult due to an accent or novel vocabulary items. In such cases, the visuals aided partial comprehension or gave the gist of the news only. Participants used images to make up for comprehension gaps or lack of comprehension:

Both the words and photos show that there is a problem, but I wasn't interested in speaking. (response 40).

She [the newscaster] was very fast but the images helped me to build a simple idea (participant 1)

The picture show me that girls got out of their home but I can't understand the words or what they say. The British accent is very difficult. (participant 4)

I don't understand everything because she [the newscaster] talks very quickly. The headline help me to understand a small idea. (participant 11)

The above excerpts are in line with "the reserves modality effect", according to which the superiority of visual-only over audio-visual presentations is attributed to "lengthy, complex, auditory textual material" that creates working memory overload, especially since speaking is "transitory" in nature (Sweller et al., 2011:137).

3.5.2. Short Captions as key Anchors for Understanding Details:

While the moving and still images gave a general idea of the news content, short captions – as a form of visual - played an important part in enriching comprehension by making details noticeable. This was particularly true of numerical graphics referring to time:

It was clear from the picture and words like (crashed at 8.44). (response 36)

She [the newscaster] said that [the plane crashed with different nationalities on board] and the picture of the plane. The numbers of nationalities were written. (response 37)

The above excerpts corroborate Cross's finding that graphics presenting numerical details allow students to make associations with the more complex audio content and facilitate comprehension (Cross, 2011: 55).

Overall, using text sparingly played a positive part in comprehension. Other examples than numerical graphics were the use of keywords such "missing," "live", and location names (the latter is referred to by participants 19, 20, 21, 29, 30) in addition to the use of minimally captioned maps with dotted lines indicating routes (figure 1).

(Figure 1: a screenshot illustrating an instance of image organization: maps and dotted lines in the Ethiopian plane crash videotext)

It may be safe to conclude that visual elements that are simplifications and summary representations of complex audio input were perceived to aid comprehension.

[I understood] the nationalities of people who were in the plane [through] the words which were on the screen. (response 40)

[What helped] was writing names of countries and numbers next to them to indicate the number of victims from each country, using of maps to show the route of the plane, and writing names and departure times. (participant 18)

The word (live-Soham) → *the murder place.* (participant 32)

"missing" word it was the first step that made me start understand. (participant 34)

Captions in the final excerpt initiated comprehension of the audio text while in the first three excerpts attention to captions was initiated by lack of comprehension. Previous research also underlined both the captions-as-"crutches" metaphor and their positive role in

understanding the more complex aural input emphasizing that each modality reinforced the other (Winke et al. 2010: 80).

Indeed, captions in this study served a "signalling" function as they worked as "cues" to guide the participants and minimise the cognitive load produced by the irreducibly rich nature of authentic videos. As manifest above, signalling, as an aid to comprehension, took on the forms of "stressing keywords in speech", using coloured arrows, word organization via captions, and image organization via maps (Mayer and Moreno, 2003:48). This is also aligned with Sweller et al's contention that narration should not simply mean re-description of graphics. Audio and visual modalities should be complementary and not intelligible in isolation (129&139); otherwise, the video risks creating a redundancy effect, as the next example of longer text confirms.

Even when longer text overlapped semantically with the audio narration, comprehension was at risk when the start of the text did not match the start of the audio narration:

I know that the prime minister is talking about the accident but I don't know what he says because they don't read all of the picture, they read one sentence from the middle. (participant 36)

(**Figure 2: a screenshot illustrating one example of textual graphics in the *Ethiopian Plane Crash* news videotext**)

The verbatim transcript of the concurrent audio content reads as follows: "The Ethiopian Prime Minister's office put out a statement expressing its deepest condolences to the families of those who've lost loved ones".

Eliminating the search for information characterizes effective multimedia presentations. Search can increase working memory load by consuming attentional resources in extraneous processing. This is what is referred to as split attention, i.e. the division of attention in an attempt to process and mentally integrate related representations of information that are separated temporally (in time) and/or spatially (in location) (*ibid.,* 111). The above is an example of spatial split attention: although part of the visual content (text) is synchronized with the audio content, the student still had to look for that part.

Another student echoed this perceived difficulty of comprehending longer text:

When they spoke about the Prime Minister, the title of Prime Minister was clearer than the rest of the text, and I focussed more on it than on the text. (response 41)

In addition to emphasising that poor attention is the result of memory overload, this excerpt corroborates Cross's finding that captions play a role in speaker identification, i.e. orienting the viewer to the name and position of speakers (2011: 64).

3.5.3. Graphics as Cues to Infer the News Content:

Interestingly, when asked if they had noticed the "police reconstruction" caption that initially appeared on the top right corner of the video where two young girls from a drama company took part in the reconstruction of the crime, many students replied that they had not. Three students signalled out the posters of the two real victims in the crime reconstruction scene as the young actresses took the same routes the victims took in the village. They concluded that this video could not have been the real footage.

(Figure 3: a screenshot illustrating the segment containing the posters of the missing victims in the reconstruction video, in the Soham crime news videotext)

Another student stated that the content was hard to understand but inferred that the news segment was related to a crime because of the images of flowers in the graveyard. The flowers were actually placed in the crime scene, a tradition that is not common in the Syrian culture. In this vein, although images generally aided comprehension, culturally laden images did not. One student, for instance, referred to a scene that was hard to understand in the Soham crime videotext. The scene shows specialist police officers digging in search of clues. It is meant to emphasize that, alongside the police reconstruction, "intense searching continues" and widens:

I didn't understand anything from the workers [specialist police officers] and the lake that appeared in the video. (participant 23)

The part of the river [shown in the videotext] was very strange. (participant 21)

This segment may have been hard to understand because it is not a tradition to show crime investigation processes and details on Syrian news. The un/familiarity of graphics,

their implications and their cultural connotations is an interesting venue to follow in understanding the impact of visuals on comprehension. The scene of intense investigation might have been presented to counterbalance the negative atmosphere and to alleviate the feelings of frustration and anxiety British viewers and village locals might have had. This conclusion was reinforced when this crime news segment was immediately followed by news about plans to tighten the laws against paedophiles. In other words, the promise of serious action to control crime marked the transition from the crime news segment. This sequence was not, however, noticed by students. Solicited interpretations of it deemed it as a warning for potential offenders to halt crime (participant 33) or as an inadequate move as far as the Soham crime is concerned:

The anchor said that the government work to tighten the law but I think that is not enough to find the girls. (participant 28)

3.5.4. Familiar Graphics as Cues to Infer the News Context:

While unfamiliar graphics did not aid comprehension, graphics representing common themes did. In the Oscar news videotext, a student referred to the "red carpet" and "famous actors", so inference was based on the learners' background knowledge of the graphic presented and this helped her to "understand slightly" (participant 6). Another example of how graphics contextualized the news and set the stage for the verbal content is the note made by participant 32:

The man is wearing formal clothes → He is a policeman or a detective in the crime

3.5.5. Facial Expressions, Voice Tone and Music as Indicators of the Mood of the News:

Other multimedia elements that students referred to were body language, voice tone and background music. These elements were perceived as an aid in establishing the mood of the news. Effective listening includes attending not to only linguistic aspects such as grammar and vocabulary but also to paralinguistic or nonverbal aspects of speech such as facial expressions and body language (Sobouti & Amiri, 2014: 342). Although Sobouti & Amiri advocate direct strategy instruction of such skills in their study, examples of such strategies surfaced in this research unsolicited:

[What helped me understand was] the facial expression of the newscaster. (participant 32)

Music and voice intonation played a similar role:

I understand that there is something terrible had happened. [What helped the participant understand was] the music of the background and the serious voice of them [of the newscaster, the reporter and the chief constable]. (participant 8)

The picture of two girls and the music help me to know that there is something dangerous happened to those girls. (participant 32)

Although the tone of voice helped the participants to get a feel of the news mood, it hindered understanding at times. Participant 32 mentioned that the reporter's tone "was strong" when he said the victims are "still alive", which he viewed as supportive of understanding. In fact, the student missed out the word "believed" since the words "still

alive" were indeed amplified. The transcript of this part of the reporter's speech is: "they [detectives] believe the girls are still alive, are probably being held together and very much against their will". In this case, the rising tone hindered, rather than aided, comprehension as other equally important words were toned down in comparison. While pictorial elements played a complementary role in understanding, the verbal modality took precedence over the visual one when the introduced information was abstract: what detectives believed.

One concluding remark is that in addition to the aspects of literacy demonstrated in decoding nonverbal language, utilizing their background knowledge, making inferences and alternating or orchestrating audio-visual processing channels, participants were also able to diagnose "gaps" in their proficiencies, a conclusion that echoes Richards's suggestion that watching television allows exposure to casual everyday English and authentic social interactions, including the expression of emotions. Such exposure, especially outside the classroom, equips learners with the capacity of "noticing gaps" in their knowledge (2014: 15). Participants in this study demonstrated such awareness of gaps:

I think I need more practice in watching news in English language and the problem again is the fast way they talked in and the intonation of the British accent. (participant 6)

I guess that will be better by practice. This is my first time I listen to news [in English]. (response 39)

It may be true that participants could have had varying degrees of exposure to news clips in English and might have drawn on certain strategies while watching and reporting on their audiovisual processing. This fact does not contradict, but rather enriches, this research as it is mainly concerned with understanding the associations participants make between audio and visual messages and the status quo of the depth of their audiovisual processing. Additionally, although participants might have watched news clips in Arabic, the "macrostructure and content of BBC news videotext are culture-bound" (Cross, 2011: 63-64).

4. Conclusion and Directions for Future Research:

This study was an endeavour to fathom Syrian EFL learners' sense-making processes and cognitive interactions with multimedia, especially visual, elements of BBC news videotexts. Students reported trade-off between audio-visual processing at times, orchestration of audio-visual processing at others. The visual modality -- in its diverse forms of moving and still images, short and long captions, and facial expressions – diversely impacted the comprehension of news videotext. This research mainly revolves around comprehending the presented content of news videotext based on students' perceptions and self-reports. Repeated sessions with the same participants might yield deeper understanding of their implicit thinking processes. Finally, although the data shows sporadic instances of participants' media literacy, future research can draw on findings to systematically train learners to consciously manipulate the various multimedia elements presented in the news and to reflect on the deeper messages underpinning news clips. This is by no means an easy undertaking given the intricate and real-time nature of such multimodal presentations in reality, but it is a worthwhile one for promoting a new approach to listening as an active process of utilizing not only knowledge but also media literacy and visual literacy.

References

BBC News, *BBC Six O' Clock News*, 10 Aug. 2002. Accessed on 25 Feb. 2019. https://www.youtube.com/watch?v=eGor_qB6noA]

BBC News, *Ethiopian Airlines: 'No Survivors' on Crashed Boeing 737*, 10 Mar. 2019. Accessed on 12 Mar. 2019. https://www.youtube.com/watch?v=nOIfgmp5Mic

BERNSEN, N. O. Multimodality Theory in TZOFARAS, D. Ed., 2008, pp.5-28.

CHARTERS, E. *The Use of Think-aloud Methods in Qualitative Research: An Introduction to Think-aloud Methods*. Brock Education. Vol. 12, Nº. 2, 2003, 68-82.

CLARK, R.C. & MAYER, R.E. *E-Learning and the Science of Instruction: Proven Guidelines for Consumers and Designers of Multimedia Learning*. 4th. ed. Wiley, Hoboken, New Jersey, 2016.

CROSS, J. *Comprehending News Videotexts: The Influence of the Visual Content*. Language Learning & Technology. Vol. 15, Nº. 4, 2011, 44-68.

DÖRNYEI, Z. *Research Methods in Applied Linguistics: Quantitative, Qualitative and Mixed Methodologies.* OUP, Oxford, 2007.

GRUBA, P. *Playing the Videotext: A Media Literacy Perspective on Video-Mediated L2 Listening*. Language Learning & Technology. Vol. 10, Nº. 2, 2006, 77-92.

HOSOGOSHI, K. *Effects of Captions and Subtitles on the Listening Process: Insights from EFL Learners' Listening Strategies*. JALTCALL. Vol. 12, Nº. 3, 153-178, 2016.

KOLTAY, T. *The Media and the Literacies: Media Literacy, Information Literacy, Digital Literacy.* Mind, Culture & Society. Vol. 33, Nº. 2, 2011, 211-221.

MAYER, R. E. & MORENO, R. *Nine Ways to Reduce Cognitive Load in Multimedia Learning*. Educational Psychologist. Vol. 38, Nº. 1, 2003, 43-52.

PAIVIO, A. *Mental Representations: A Dual Coding Approach*. Oxford UP, New York, 1990.

RICHARDS, J. C. *The Changing Face of Language Learning: Learning beyond the Classroom*. RELC. 2014, 1-18.

SOBOUTI, F. & AMIRI, M. *The Impact of Listening Strategy Instruction on the Improvement of Iranian Intermediate EFL Learners' Comprehension of News Videotext*. Academic Research International. Vol. 5, Nº. 2, 2014, 334-344.

SWELLER, J. AYRES, P. & KALYUGA, S. *Cognitive Load Theory*. New York, Springer, 2011.

TZOFARAS, D. Ed. *Multimodal User Interfaces: From Signals to Interaction*. Springer-Verlag, Heidelberg, 2008.

VANDERGRIFT, L. *Listening to Learn or Learning to Listen?* Annual Review of Applied Linguistics. Vol. 24, 2004, 3-25.

VAN DER MOLEN, W. *Assessing Text-Picture Correspondence in Television News: The Development of a New Coding Scheme.* Journal of Broadcasting and Electronic Media. Vol. 45, Nº 3, 2001, 483-498.

WINKE, P. GASS, S. & SYDORENKO, T. *The Effects of Captioning Videos Used for Foreign Language Listening Activities*. Language Learning & Technology. Vol. 14, Nº. 1, 2010, 65-86.