Abstract—The purpose of this study is to investigate the effects of modality principles in instructional software among first grade pupils’ achievements in the learning of Arabic Language. Two modes of instructional software were systematically designed and developed, audio with images (AI), and text with images (TI). The quasi-experimental design was used in the study. The sample consisted of 123 male and female pupils from IRBED Education Directorate, Jordan. The pupils were randomly assigned to any one of the two modes. The independent variable comprised the two modes of the instructional software, the students’ achievement levels in the Arabic Language class and gender. The dependent variable was the achievements of the pupils in the Arabic Language test. The theoretical framework of this study was based on Mayer’s Cognitive Theory of Multimedia Learning. Four hypotheses were postulated and tested. Analyses of Variance (ANOVA) showed that pupils using the (AI) mode performed significantly better than those using (TI) mode. This study concluded that the audio with images mode was an important aid to learning as compared to text with images mode.

Keywords—Cognitive theory of Multimedia Learning, Modality Principle, Multimedia, Arabic Language learning

I. INTRODUCTION

In the course of finding ways to improve the teaching-learning process, many techniques were developed by mankind to deliver the contents to the learners [1]. The most prominent of such techniques were particularly the sound, images, text and film in the form of multimedia procedures [2]. However, computers today have revolutionized the ability of dealing with such techniques and procedures where many software programs are designed to employ such aids in the teaching-learning process [3].

Over the last two decades, the instructional technology researchers’ interests have shifted from the environment and external stimuli to the cognitive processes that occur inside the learner so as to present content in a way similar to his mental processing of knowledge and to have the new knowledge linked with the prior knowledge [4]. In short, the focus has shifted from studying the effects of external factors in education to studying the effects of the external stimuli on the internal learning processes [5]. As a result many theories, propositions and models have been developed that address intellectual functions, mental processes, memory types and functions [6].

Designers today are more inclined to follow such an approach, that is, to design a learning environment instead of designing a teaching environment, the difference here would be that in the learning environment, the learner's mental processes, learning patterns and activities take centre stage whereas in the teaching environment the external stimuli garner the most attention in the learning design [7]. This, however, does not mean that one disregards the role of the teacher and the surrounding environment; rather that the external stimuli play a secondary role [8].

II. THEORETICAL FRAMEWORK

According to Mayer [7] when watching a multimedia presentation, the information processing system in human beings uses both words (printed text, spoken text) and pictures (drawing, charts, graphics, maps, photos, animation and video) together rather than words alone. The design of multimedia environments should be compatible with how people learn. Mayer [7] presented a cognitive model of multimedia learning to present the human information processing system (Figure 1).

In Figure 1 there are three frames, "second, third, fourth", of human memory that store Sensory, Working and Long-term memory. Pictures and words come in from the outside world through a multimedia presentation and enter sensory memory through the eyes and ears (included in the sensory memory frame). In Figure 1 there are three frames, "second, third, fourth", of human memory that store Sensory, Working and Long-term memory. Pictures and words come in from the outside world through a multimedia presentation and enter sensory memory through the eyes and ears (included in the sensory memory frame). Sensory memory allows for pictures and printed text to be held as exact visual images for a very brief time period in the visual sensory memory and for spoken words and other sounds to be held as exact auditory images for a very brief time period in the auditory memory.
The arrow from the pictures to the eyes corresponds to a picture being registered in the eyes; the arrow from the words to the eyes corresponds to the printed text being registered in the eyes. The central work of multimedia learning takes place in the working memory. Mayer [7] advocated that the working memory is used for temporarily holding and manipulating knowledge in active consciousness. The visual sensory memory and the auditory memory go into the working memory, that is, the visual images of pictures and sound images of words as represented in the left hand side of the box labeled working memory in Figure 1 The arrow from sounds to images represents the mental conversion of a sound (such as the spoken word “cat”) into a visual image (such as an image of a “cat”), that is, when you hear the word “cat”, you might also form a mental image of a “cat”. The arrow from images to sounds represents the mental conversion of a visual image into a sound image, that is, when you see a picture of a “cat” you may mentally hear the word “cat”. These processes may occur by mental association in which the spoken word “cat” primes the image of a “cat” and vice versa. In contrast, the right side of the working memory box represents the knowledge constructed in working memory pictorial and verbal mental models and link between them. The major cognitive processing required for multimedia learning is represented by the arrows labeled selecting images, selecting sound, organizing images, organizing words, and integrating.

The box labeled long-term memory is the learner’s storehouse of knowledge. Unlike the working memory, long-term memory can hold large amounts of knowledge over long periods of time. For a person to actively think about material in the long-term memory, it must be brought into the working memory (as indicated by the arrow from long-term memory to working memory). The learner represents an active agent in the learning process via multimedia. He attempts to construct a meaning for the information presented through the following three major mental processes: Selection, Organization and Integration. Selection is a mental process where the individual learner pays attention to relevant information presented to him verbally or non-verbally. Mayer [9] defined selection as choosing the right information and adding it to the working memory. On the other hand, Paivio [10] defined selection as a process that involves selecting verbal stimulation "words" to construct verbal base "text base" and selecting non-verbal stimulation to build non-verbal base "pictorial base" information to be sent to "short-term memory". After the learner has selected the verbal and non-verbal information, the next step is organization. This process involves ordering and organizing information that have been selected meaningfully and logically. Clark and Mayer [5] states that organization is a mental process performed on text "selected words" in order to organize them in a verbal model that is capable of interpreting this verbal information.

The organization process is also performed on image base "selected images" to organize them in an image model capable of interpreting them. As the verbal model and visual model are constructed, the integration process follows. Integration means making connection between the verbal and image models. Moreno & Mayer [11] indicated that such processes involve connecting organized information in the verbal and image models with relevant and similar information stored in long term memory. Paivio [10] indicated that integration involves making referential connections between both verbal and image models. As such, we can see that in multimedia, information is presented to the learner in more than a single form such as words, images, motion pictures, sounds and other forms. Therefore, it is necessary to know how the learner processes information presented to him via Multimedia that fits his mental processing style and which will facilitate his construction of correct and good mental models.

III. MODALITY PRINCIPLE

The modality principle, use of audio narration to explain a visual presentation enhances understanding, when practical [7]. Cognitive theory of multimedia show that spoken text (audio) can use to explain a image [4]. The auditory text is of great importance, as the learners can listen to an auditory presented text while they are trying to study a picture. Moreover, the auditory presented text will also draw the learner’s attention to that part of the picture, which is just mentioned in the text. Thus, the information can be processed and interpreted in parallel without the need of transformation of the previous parts as in the visual-only format. The above concept is known as the modality principle [12]. Which is concerned with the presentation of the information to the learner as portrayed in Figure 2 the principle itself indicates that the auditory explanations are usually better compared to the written explanations especially if it is used along with illustrative learning material [13]; [11] & [14]. Thus, it is suggested that pictures should go together with a synchronized auditory explanatory recitation rather than a synchronized explanatory text, see Figure 2.

Mayer & Mereno [15] found that the learner’s understanding when watching animated images on a certain phenomenon (the lightning phenomenon) accompanied with an auditory explanation is much better than the learner’s understanding when watching animated images on the same phenomenon while verbally reading an onscreen-text explanation.

At the same time, this principle is in tandem with the cognitive learning theory using multimedia techniques. The cognitive learning theory suggests that reading an onscreen-
text along with animated pictures will result in interference in the rendered information especially when this information is to be processed verbally [16]. However, this will add more load on the visual working memory because both are addressed in the form of verbal memory and in visual approach while the auditory text is processed in the verbal model and audio working memory [17]. Furthermore, animated images are processed in the visual model, thus, providing an auditory text along with watching an explanatory film that does not overlap [18]. Therefore, this also fits the modality principle by combining the visual channel for graphical material and the verbal channel for the explanation of this material to increase the working memory capacity effectiveness and facilitate learning.

The results of Hamtini’s [19] study to investigate the effect of using instructional software in other academic disciplines recommended the use of computers in teaching art education where higher levels of achievement can be pursued if art education is held with the aid of a personal computer. The Computer-based multimedia learning environments consisting of pictures (such as animation) and words (such as narration) offer a potentially powerful venue for improving student understanding [20]. The modality principle represents the human information processing system, the cognitive theory of multimedia learning, and supports the use of free cognitive resources [21].

Alherish, Ababneh & Aldalalah [22] conducted a study to examine the effects of using different code systems in the instructional software on third graders’ achievements in science. The results indicated a significant difference in achievement for the group which was taught using audio and animated pictures in comparison to the other groups. Aldalalah & Fong [23] study found that the sounds and pictures used in the educational learning process in the multimedia are treated in the memory through visual and auditory channels. Haruskamp, Mayer & Suhre [24] conducted a study that demonstrated that the modality principle applies to multimedia learning of regular science lessons in school settings. Similar results were found in a study conducted by Moreno, Mayer, Spires, & Lester [25] when they examined how college students learned to design a plant in a multimedia environment. Results indicated a significantly better performance on tests of retention and problem-solving transfer when students received instruction that contained narrated words rather than on-screen text. Tindall-Ford, Chandler & Sweller [26] found that students learned electrical engineering concepts better when instruction contained both visual information (diagrams and tables) and auditory information (words presented through audio cassettes-tapes) than when instruction contained only visual information (diagrams and words presented as text). The authors explained that the engineering concepts that contained a high level of element interactivity placed a significant burden on students’ work.

Finally, the modality effect in multimedia suggests accompanied text and pictures to be presented auditorily rather than presenting them visually. This combination between text, picture and audio explanation or presentation will significantly help prevent the split of attention dilemma as suggested by Seufert, Schu & Nken [27] and Aldalalah & Fong [28]. Therefore, spoken explanations are proposed to accompany animations in order to prevent overloading the visual channel (i.e. the modality effect) [18].

IV. STATEMENT OF THE PROBLEM

We live in a world of knowledge explosion and technological advancements. It is not surprising that the multimedia technologies have found their place in various educational programs which capitalized on their educational capabilities particularly in the use of computers as an instructional instrument. The real challenge is how to employ the processing capabilities of computers to support some of the theories when designing an instructional software program so that the designing process of instructional software programs are built on a sound theoretical basis and not merely dependent on the designer's taste [2]; [29] & [8]. The Ministry of Education in Jordan is cognizant of the need to develop a holistic curriculum incorporating all elements relying on local experiences in line with the country’s objective of achieving a knowledge-based economy in Jordan.

The present study explores the most effective way of learning the Arabic language based on the cognitive theory of multimedia learning technologies in an effort to test some of the principles derived from this theory. This pursuit complements the efforts of earlier researchers to test for such principles in their quest to develop a theory on designing instructional software programs.

V. SIGNIFICANCE OF THE STUDY

The present study is deemed important as it is in tandem with the aims of the Jordanian Ministry of Education to see a greater and more effective role of the use of computers and the Internet in the teaching and learning process of various subjects in the education system. Students, in the early basic instructional stages of language learning, learn the basic language skills in the various fields as well as being socialized in a way that would help them acquire higher skills in Arabic language literacy later on. This study would be an addition to the research efforts of developing new Arabic language teaching methods. It is expected that this study will be important as it goes beyond the traditional teaching-learning method that is not anchored on a theoretical framework which could only lead to confusing results and ineffective learning by pupils. Instead this study presents an attempt to present information in a way that allows learners to handle new information effectively based on sound theoretical constructs.

VI. RESEARCH OBJECTIVE

There are two purposes in this study: The first is to design and develop two modes of courseware, namely, text and images (TI), and, Audio and images (AI). The second purpose is to evaluate the impact of the Modality Principle on computer-assisted learning.

VII. RESEARCH HYPOTHESES

1. Will High achievers pupils using the (Audio, images) mode attain significantly higher post-test score than High degree pupils using (text, images) mode?
2. Will Medium achievers pupils using the (Audio, images) mode attain significantly higher post-test score than Medium achievers pupils using (text, images) mode?
3. Will Low achievers pupils using the (Audio, images) mode attain significantly higher post-test score than low achievers pupils using the (text, images) mode?
4. Are there interaction between gender and the 2 modes of software’s?

VIII. LIMITATION

1. The program will only involve designing a program with two treatments, (Audio, images) mode and (text, images) mode.
2. Subjects in the current study will include primary first graders within IRBED Provincial School District in Jordan.
3. This study will be limited to Arabic language classes.
4. This study will be limited to modality principle of cognitive theory of multimedia learning.

IX. METHODS

A. Sample

The study sample consists of 123 pupils from the first grade of the IRBED Education Directorate of the 2009/2010 academic year. The simple random sampling method was used to identify the primary first grades from the schools therein. According to Gay and Airasian [30] ‘‘all the individuals in the defined population have equal and independent chance of being selected. The first grade pupils were divided into three groups depending on their achievement in the Arabic language class in the first semester following the equation:

\[ Z = \frac{\text{Highest Mark} - \text{Lowest Mark}}{\text{Number of Levels}} \]

Where

- \( H_{\text{min}} \) = Highest Mark
- \( L_{\text{min}} \) = Lowest Mark
- \( Z \) = The difference between a level and the other.

\[ L = \text{Low, } M = \text{Medium, } H = \text{high} \]
\[ L \text{ (range)} = [L_{\text{min}} - L_{\text{min}} + Z] \]
\[ M \text{ (range)} = (L_{\text{min}} + Z - L_{\text{min}} + 2 Z) \]
\[ H \text{ (range)} = (L_{\text{min}} + 2 Z - H_{\text{min}}) \]

Highest Mark = 99, Lowest Mark = 56, Number of Levels = 3
\[ Z = \frac{99 - 56}{3} = 14 \]
\[ L \text{ (range)} = [56 - 70] \]
\[ M \text{ (range)} = (71 - 85) \]
\[ H \text{ (range)} = (86 - 99) \]

B. Experimental Condition

The pupils’ distribution within the treatment groups was conducted randomly. Then the treatment groups were exposed to the treatment consecutively. The two treatment groups are as follows:
1. First treatment: (text, images)
2. Second treatment: (audio, images)

C. Instruments

1) There is one instrument in this study, Arabic language class achievement test. The Arabic language class achievement test that was administered onto the two groups’ participants in this study was adapted from the Arabic language class competency test developed by the researcher. The Arabic language class achievement test consists of 30 multiple-choice Arabic language class items based on regarding reading and writing. There are 15 items on reading and 15 items on writing, and the reliability coefficient was measured by applying the test on a pilot sample. The total score of the Arabic Language achievement test is 30; a student receives a score of either 1 (correct answer) or 0 (missing data incorrect answer or didn’t answer). The reliability of the test questions was calculated using the Cronbach Alpha procedure to calculate the internal consistency. The Cronbach Alpha of the test was 0.80; the internal consistency of the test was 0.93.

2) Instruments Validity: Validity of the instruments are important aspects that should be taken into account when conducting a research. Validity consists of two different aspects that is face and content validity. According to Gay & Airasian [31] face validity relates to the “degree to which a test appears to measure what it claims to measure”. Face validity was judged by a panel of experts in the field of Arabic language teaching. Content validity refers to the “degree to which a test measures an intended content area” [31]. Content validity of the instruments in this research was justified by the panel. The feedback and comments received from the panel of experts were employed to establish the necessary clarifications, changes, and modifications before and after piloting the study.

D. Study Design

This study followed the quasi experimental design to measure the impact of the 2 modes of treatments on the total achievement score in reading and writing in the Arabic language class.

E. Research Variables

The present research contains two types of variables (independent and dependent) that are presented as follows:

- Independent Variables:
  1. The modes of presentation:
     - Multimedia computer-based learning courseware with Arabic language presented in text and image (TI)
     - Multimedia computer-based learning courseware with Arabic language presented in audio and text (AI)
  2. The students’ achievement levels in the Arabic Language class has three levels (High, Medium, Low)

- Dependent Variable
  1. The total achievement score in reading and writing in the Arabic language class
VII. RESULTS AND DISCUSSION

The analyses of the collected data were carried out through ANOVA statistical techniques, the data were compiled and analyzed using the Statistical Package for the Social Science (SPSS 16) for Windows computer software. In this section we discuss the results and recommendations of the study.

A. Results related to first question

<table>
<thead>
<tr>
<th>Source</th>
<th>Type II Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>21437.837 a</td>
<td>4</td>
<td>5359.459</td>
<td>1387.399</td>
<td>.000</td>
</tr>
<tr>
<td>Gender</td>
<td>2.624</td>
<td>1</td>
<td>.679</td>
<td>.417</td>
<td></td>
</tr>
<tr>
<td>Group</td>
<td>11.073</td>
<td>1</td>
<td>2.866</td>
<td>.102</td>
<td></td>
</tr>
<tr>
<td>Gender* Group</td>
<td>.264</td>
<td>1</td>
<td>.068</td>
<td>.796</td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td>108.163</td>
<td>28</td>
<td>3.863</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>21546.000 a</td>
<td>32</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

It was shown that there are no significant differences between pupils using (text, images) and pupils using (audio, images) and this finding is in tandem with the cognitive theory of multimedia learning [7]. When information is presented through two channels, the learner can build two mental models: a verbal model and an image model and he can build referential connections between them.

There is no statistical difference between the two processes. This could be that the pupils who have a high achievement level are able to build mental model, regardless of the existence of a good or bad design. The learning requires building simple relations between the introduced information and the information resides in the long-term memory. Thus, the high achievers will be able to build the station model depending on the extracted ideas from the information shown with no need for a good model existence.

B. Results related to the second question

<table>
<thead>
<tr>
<th>Source</th>
<th>Type II Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>26138.557 a</td>
<td>4</td>
<td>6534.639</td>
<td>932.340</td>
<td>.000</td>
</tr>
<tr>
<td>Gender</td>
<td>0.298</td>
<td>1</td>
<td>.043</td>
<td>.837</td>
<td></td>
</tr>
<tr>
<td>Group</td>
<td>31.411</td>
<td>1</td>
<td>4.482</td>
<td>.039</td>
<td></td>
</tr>
<tr>
<td>Gender* Group</td>
<td>15.110</td>
<td>1</td>
<td>2.156</td>
<td>.148</td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td>350.443</td>
<td>50</td>
<td>7.009</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>26489.000 a</td>
<td>54</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In regard to total accomplishment, there were some significant statistical indications between the two multimedia applications. The results were in favor of the second application which studies the effect of the (modality principal) audio, and images mode. The pupils could build better understanding since they received the information via the images and audio symbols which enabled them to construct interrelationships between the two channels and the existing symbols in the working memory.

According to Clark and Mayer[5] redundant onscreen text to a multimedia presentation could overload the visual channel because the image enters the pupils’ cognitive system through the eyes and is processed in the visual channel since; the printed text enters through the eyes and must be processed in the visual channel. Since the limited cognitive resources in the visual channel must be shared in processing both the image and the text and hence causing an overloaded.

According to the cognitive theory of multimedia learning, pupils have limited cognitive capacity so, if they use their cognitive capacity to reconcile printed and spoken text, they cannot use it to make sense of the presentation. This is called cognitive load; also those pupils might pay so much attention to the printed text that they pay less attention to the image. When their eyes are on the printed text, the pupils cannot be looking at the image. In addition, the pupils may try to compare and reconcile the printed text and the narration text which requires cognitive processing extraneous to learning the content.

When information is presented to the eye (image and text), humans begin by processing that information in the visual channel; when information is presented to the ears (audio), humans begin by processing that information in the auditory channel. According to Mayer [7] humans have limited capacity in the amount of information that can be processed in each channel at any one time. In this way the information processed through the two channels are balanced, that is, neither one of the channel is cognitive overloaded. The picture enters through the eyes (processed in the visual/ pictorial channel) while the spoken word enters through the ears (processed in the auditory / verbal channel).

C. Results related to the third question

<table>
<thead>
<tr>
<th>Source</th>
<th>Type II Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>6553.816 a</td>
<td>4</td>
<td>1638.454</td>
<td>186.971</td>
<td>.000</td>
</tr>
<tr>
<td>Gender</td>
<td>84.480</td>
<td>1</td>
<td>84.480</td>
<td>9.640</td>
<td>.004</td>
</tr>
<tr>
<td>Group</td>
<td>4.983</td>
<td>1</td>
<td>4.983</td>
<td>.006</td>
<td>.940</td>
</tr>
<tr>
<td>Gender* Group</td>
<td>5.561</td>
<td>1</td>
<td>5.561</td>
<td>.635</td>
<td>.431</td>
</tr>
<tr>
<td>Error</td>
<td>289.184</td>
<td>33</td>
<td>8.763</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>6843.000</td>
<td>37</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The total accomplishment of the pupils was different in the two multimedia applications. The pupils’ performance was better for the second application where more senses were involved in the learning process. According to cognitive load theory humans have limited capacity in the amount of information that can be processed in each channel at any one time [7]. Hence, it is not surprising that low achievement level pupils using AI mode reported lower levels of mental effort than low achievement level pupils using TI mode. Since AI assists in the reduction of intrinsic cognitive load, pupils with
low achievement levels will experience lower levels of cognitive load when processing incoming information. In this way the information processed through the two channels are balanced, that is, neither one of the channel is overloaded.

Recent study has examined the possibility that low achievement level pupils may interfere with working memory. In the TI mode the image and the text are transformed through the visual channel only. This channel has limited capacity of the channel leading to a condition of cognitive overloaded. As such the auditory channel was unused, and this may affect the text and image in the working memory. There the information will be partly effective.

D. Results related to first question

It was shown that there are no significant differences from a statistical point of view due to gender or the interaction between gender and the application. This conforms to the results in different literature studies which demonstrated the inexistence of any differences in the approach used to process the information between males and females [32].

VIII. CONCLUSIONS

This study found that the use of modality principle reduced the cognitive overload and increased the working memory capacity resulting in better Arabic language learning among low and medium level pupils. In this study it was shown that there are no significant differences from a statistical point of view due to gender or the interaction between gender and the application. There is no statistical difference between AI and TI among high level pupils.

REFERENCES


Osamah (Mohammad Ameen) Ahmad Aldalalah is with the School of Educational Studies, Universiti Sains Malaysia. He has published many papers on educational technology in many journals. His research interests are in the areas of Aptitude Treatment Interactions, Educational Technology, Multimedia and Music Education.

Soon Fook Fong is an Associate Professor of the School of Educational Studies, Universiti Sains Malaysia. He received his PhD in 2001 from Universiti Sains Malaysia. He has published many papers on multimedia in
education in many journals. His research interests are in the areas of Aptitude Treatment Interactions, Instructional Designs and e-Learning Designs. Ziad Waleed Mohamed Ababneh is with the Ministry of Education, Jordan. He received his master's degree in 2003. His research interests are in the areas of Educational Technology, Multimedia and teaching. He has published many papers on educational technology in many journals.