A Social Semiotic Analysis of Social Actors in English-Learning Software Applications

Esmat Babaii
Associate Professor
Kharazmi University
ebabaii@gmail.com

Mahmood Reza Atai
Associate Professor
Kharazmi University
mahmood.atai@gmail.com

Masoumeh Kafshgarsouteh∗
PhD Student
Kharazmi University
kafshgarm@gmail.com

Abstract

This study drew upon Kress and Van Leeuwen's (2006, [1996]) visual grammar and Van Leeuwen's (2008) social semiotic model to interrogate ways through which social actors of different races are visually and textually represented in four award-winning English-learning software packages. The analysis was based on narrative actional/reactional processes at the ideational level; mood, perspective, social distance, and modality at the interpersonal level; and salience, framing, and vector at the compositional level. The findings revealed that although contemporary multimodal texts have tried to be unbiased and neutral in the verbal mode, there are still traces of discrimination, bias, and stereotyping in the visual mode. The results of this research can be of potential help and use for researchers, pedagogues, material developers, software designers, teachers , and students to become visually literate and get aware of the hidden messages that can be communicated by images in textbooks and multimedia.

Keywords: social actors, visual and verbal modes, compositional metafunctions, race, stereotyping

Images are the oldest form of recording and transmitting information, conveying messages, and communicating. While verbal expression is about 7000 years old, visual expression dates back to 30000 years ago (Hietala, 1996). Today in EFL/ESL materials, a shift of focus is obviously detected from textual content to visual presentation (Prodromou, 1988). Images combined with texts facilitate learning and make learners more aware of and inclined to thinking about the process of language. By bringing the real outside world to the learning contexts and by engaging learners in the act of finding ways to play with the language and its structure, images make the task of language learning appear more authentic (Canning-Wilson, 1999). Consequently, images are being consistently used in textbooks and educational materials including English-learning CD ROMs and DVD ROMs whether for the purpose of illustration, decoration, or information (Baldry, 2000; Kern, 2006; Kress & Van Leeuwen, 1996, 2006; Lim Fei, 2007; O'Halloran, 2009, 2011; O'Halloran, Podlasov, Chua & K.L.E, 2012; Pauwels, 2012; Tahririan & Sadri, 2013). Aiming at enhancing the learning processes of students, English-learning tools including interactive software packages enjoy a rich and coherent interplay of written text along with images, animation, music, sound and other graphic elements which turn them into highly multimodal texts.

It must be kept in mind, however, that images can never be innocent. The visual elements accompanying written texts play other roles than just complementing the verbal messages; they can affect the viewers emotionally more than words alone do and can have a detrimental lifelong effect on the viewers' minds (Lester, 2000). Visual images, in fact, can teach learners secondary information by misrepresenting the reality and implying stereotyping, prejudice, and bias regarding issues such as race, nationality, gender role, or social class. Thus one cannot simply claim that images are just used for pedagogical purposes and are playing a complementary role for written text; they can, consciously or unconsciously, play a significant role in shaping and influencing the development of self-concept, emotion, attitude, value, and behavior of people. As powerful agents of socialization, visual images can make people who they are and shape their attitudes and values (Belknap &

Leonard, 1991; Sheldon, 2004). People might be negatively affected and feel marginalized by poor, distorted, and undesirable portrayal of their different race, nationality, social class, or gender if they are represented humiliatingly in English-learning materials (Bishop & Jaworski, 2003; Caldas-Coulthard, 2003; Sunderland, 1992). Hence, the latent consequences of images —as part of a larger set of social, cultural, economic, historical, and political practices— conveying messages to learners need a serious consideration.

In spite of the significance of understanding the messages conveyed by images as a communication ability, or 'visual literacy' in Kress and van Leeuwen's (2006) term, it seems that 'reading images' is not taught in schools and universities, or in teacher education programs. This in turn will lead to producing illiterates, being a learner or a teacher. As Fairclough (1989) asserts "Not all photographs are equal: Any photograph gives one image of a scene or person from among the many possible images. The choice is important, because different images convey different meanings" (p. 52). However, neither the teachers in teacher education programs nor the learners are made aware and conscious of the fact that images can be used to position and manipulate the viewers (Belknap & Leonard, 1991; Berger, 1972; Bishop & Jaworski, 2003; Caldas-Coulthard, 2003; Lester, 2000; Sheldon, 2004; Van Leeuwen, 2008). They powerfully influence public opinions, values, and ideologies (Gamson, Croteau, Hoynes & Sasson, 1992) and may change the learners' attitudes toward certain groups of people with whom they have little or no direct contact. The way 'reality' is represented to the audience has a great social influence on people and they tend to receive these representations as natural through being constantly exposed to them (Griffiths, 2010).

In an attempt to uncover how certain groups of people with different races, gender, or ethnicity are represented to the public in multimodal texts, many researchers have drawn upon multimodal critical discourse analysis in analyzing various linguistic or non-linguistic semiotic resources such as photographs and other graphic elements (Iedema, 2003; Kress & van Leeuwen, 2006 [1996]), children's toys (Machin & Van Leeuwen, 2009), political cartoons (Mazid, 2008), music (Thompson, 2002), TV commercials (Babaii & Ansari, 2001; Baldry, 2000, 2004; Baldy &

Thibault, 2006; Lim Fei & O'Halloran, 2010; Thibault, 2000), websites and online communication (Adami, 2013; Chiew 2004; Jewitt 2004; Jones 2005; Norris 2004; O'Halloran 2004; Young 2009), newspaper (Coelho 2008; Knox 2009; Machin & Mayr, 2007), and video game (Jewitt 2005). Printed books have also attracted the attention of many multimodal discourse analysts. For instance, in an attempt to investigate the extent to which visual and verbal components collaborate to create meaning, Guijarro & Sanz (2008) conducted a multimodal analysis of a children's picture book. Perhaps the most widely attended issue in multimodal textbook analyses among the researchers is gender bias. Following Erving Goffman's (1979) classic work on Gender Advertisements, scholars such as Abbas-Nejad-Konjin (2012), Ansary & Babaii (2003), Gharbavi & Mousavi (2012), Graci (1989), Hartman & Judd (1978), Hellinger (1980), Peterson & Kroner (1992), Porreca, (1984), Poulou, (1997), and Reese (1994) to name a few, have done multimodal textbooks analyses in order to study gender bias.

However, when it comes to learning software packages, in spite of the widespread use of technology based media, especially language learning software packages, little research has been conducted into the subject of stereotyping and bias conveyed by visual and verbal elements. What adds to this shortcoming is that the very few researchers who worked on learning software packages have restricted their research to investigating the issue of gender stereotyping. For instance, Chappell (1996) analyzed seventeen mathematics educational software packages, Binns & Branch (1995) and Milburn, Carney, and Ramirez (2001) studied computer clipart, Drees and Phye (2001) investigated thirty four language arts software packages, Sheldon (2004) analyzed children's software, and Kordjazi (2012) analyzed two English learning software packages. Though investigating gender stereotyping in multimodal texts is of great importance and has already outnumbered those research studies that deal with other types of stereotyping, there are other crucial issues which are worth considering through social semiotic multimodal analysis such as how certain groups of people with different races are represented to the public through linguistic or non-linguistic semiotics.

To bridge this gap in the literature, the present study drew on the frameworks proposed by Kress and Van Leeuwen's (2006 [1996]) visual grammar, Van Leeuwen's (2008) social semiotic model, and Halliday's (1978, 1985) systemic functional linguistics to conduct a critical visual and textual analysis of four contemporary English-learning software applications. The aim was to explore the ways these software packages as educational multimedia had positioned, represented, and depicted people of different races through their images and to investigate and interpret the hidden ideological meanings the texts and images were, intentionally or unintentionally, trying to impose on the language learners. It is worth mentioning that these issues have already been investigated by other researchers (Brand, Knight & Majewski, 2003; Chappell, 1996; Drees & Phye, 2001; Hjorth, 1997; Horváth Futó, 2011; Jarić, 2002; Lee & Collins, 2006; Milburn, Carney, & Ramirez, 2001; Otlowski, 2003; Sheldon, 2004) but the focus was mostly on content analysis of either oral or written language. This research, on the other hand, is trying to interrogate photographs and illustrations through a visual analysis model to reveal any potential instances of stereotyping, bias, and prejudice in representing people of different races. The following research question is central to this inquiry:

- *In what ways are the people of different races visually represented in the discourse of English-learning software applications?*

In order to answer the main research question, the following sub-questions need to be addressed and taken into account.
Regarding people of different races:
- *Who is excluded in the images?*
- *Who is included in the images?*

For those who are included in the images:
- *Who is active?*
- *Who is passive?*
- *Who is involved in action as a patient?*
- *Who is involved in action as an agent?*

- ✓ *Are they shown as agents of actions which are held in low esteem?*
- *Who is shown as generic?*
  - ✓ *Are they shown based on cultural categorization or biological categorization?*
- *Who is shown as specific?*
- *Who is shown as individual?*
- *Who is shown as groups?*
  - ✓ *Are they shown as homogenization or differentiation?*
- *What roles have been allocated to the social actors?*

In case of representing the social actors:
- *How is the distance between participants and viewers shown in the images?*
- *What is the relationship between participants and viewers?*
- *What is the nature of interaction between participants and viewers?*

Theoretical Framework

This study draws upon Kress and Van Leeuwen's (2006 [1996]) visual grammar and Van Leeuwen's (2008) social semiotics, a branch of critical discourse analysis which is concerned with analyzing visual elements. In an attempt to develop a grammar for visual design, Kress and Van Leeuwen (2006 [1996]) in their book *Reading images: The grammar of visual design* and then later Van Leeuwen (2008) in his book *Discourse and Practice: New tools in Critical Discourse Analysis* offered a method for interpreting images involving context and formal elements. They were mainly influenced by Ronald Barthes's theory. However, they differed with him in that Barthes believed the meanings of images and texts are interdependent:

> These two structures are co-operative but, since their units are heterogeneous, necessarily remain separate from one another: here (in the text) the substance of the message is made up of words; there (in the photograph) of lines, surfaces, shades. Moreover, the two

structures of the message each occupy their own defined spaces, these being contiguous but not 'homogenized', as they are for example in the rebus which fuses words and images in a single line of reading (Barthes, 1977, p.16)

On the contrary, Kress and Van Leeuwen consider texts and images as independent modes of communication. Based on their visual design grammar, the image is an independent message which is connected to the text but not reliant on it nor is the text reliant on the image (Kress & Van Leeuwen, 2006; Royce, 2007).

The grammar of visual design offers ways for describing and analyzing how and in what ways the visual elements combine together and what overt or hidden meanings they are trying to convey just as the way the grammar of language describes how written words combine in texts to convey various meanings and messages (Caldas-Coulthard & Iedema, 2008; Camiciottoli, 2007; Iedema, 2003; Unsworth, 2008). The visual grammar, which is based on Michael Halliday's (1978, 1985) systemic functional grammar, makes use of the design elements of color, mood, perspective, social distance, framing and composition to show how visual design communicates meaning (Unsworth & Wheeler, 2002; Van Leeuwen, 2003).

Regarding the visual grammar, the functions of visual design are presented based on three metafunctions: ideational (patterns of representation), interpersonal (patterns of interaction), and textual (representation and communicative acts) which come together to give meaning to the visual elements (Baldry & Thibault, 2006; Lemke, 1998, 2002; Liu & O'Halloran, 2009; Lim Fei, 2007; O'Halloran, 2000, 2004, 2005; O'Toole, 2010; Royce, 1998, 2002, 2007; Unsworth & Wheeler, 2002; Ventola, Charles & Kaltenbacher, 2004).

The first metafunction of their theory, *ideational*, is mostly concerned with representing the interaction and relation between the people, places and things which are depicted in images. As stated by Kress & Van Leeuwen (2006 [1996]), it is not simply the case that images are showing us the visual structure of reality. The fact is that they represent the standpoint of the institution within which they are produced, distributed

and consumed. Images ideologically carry a "deeply important semantic dimension" (p. 47). Every semiotic act involves interactive participants - those who speak/listen or write/read, make/view images – and represented participants (or 'Social Actors' in Van Leeuwen's (2008) term) – those who are the subject of the communication. Narrative processes represent the social actors as being involved in doing something either to or for each other. The focus of these processes is on the social actors' transactions and reactions (Unsworth, 2008; Unsworth & Wheeler, 2002). Vectors, as lines of energy and direction such as eye-lines, bodies, limbs, or gestures indicating a line force in a particular direction, are used to link participants to one another or to the processes. On the basis of different vectors and the number and kinds of participants involved in the interaction, various narrative processes are identified such as (Kress & Van Leeuwen, 2006):

- Action processes: the participant who forms the vector or is the vector itself is called the Actor. Action processes can be either transactional or non-transactional. Transactional processes have two participants, an Actor and a Goal. Some transactional processes are bidirectional. In this case, a vector connects two interactors in a way that each participant is both an Actor and a Goal simultaneously. Non-transactional processes, however, have only one participant that is usually an Actor.
- Reactional processes: the vector is formed by an eye line. In reactional processes, instead of Actors, the word Reactor is used, and instead of Goal, the term Phenomenon is used. Like actions, reactions can be either transactional or non-transactional.
- Circumstances: narrative images may contain secondary participants that are not related to the main participants by means of vector but in other ways called circumstances. Circumstances of means are the tools with which the participants perform the actions. Locative circumstances relate the participants to a specific setting. Circumstance of accompaniment is the participant that has no vectorial relations with other participants.

The second metafunction, *interpersonal*, involves the interaction between the producer and the viewer of the image. The way in which images represent social interactions and relations is taken into account in

this metafunction. Where there is no face-to-face communication, producers and viewers of images are joined in interaction. In this case, the interactions are structured by three factors: gaze (the social interaction between the depicted people and the viewers), angle (the social relation between the depicted people and the viewers), and distance (the social distance between the depicted people and the viewers) (Caldas-Couldhard & Iedema, 2008; Camiciottoli, 2007; Iedema, 2001; Kress & Van Leeuwen, 2006 [1996]).

The interaction between the viewers of an image and the represented participants is formed and structured by the gaze of the represented social actors. If the participants do not look at the viewer, then it seems they have been offered to the gaze of viewers as objects of scrutiny. If they directly look at the viewers, on the other hand, it means they are making a demand and they want something from us. Connecting the viewer and the participant is done by vectors on a formal and imaginary level (Kress & Van Leeuwen, 2006 [1996]; Royce, 2007; Unsworth, 2008; Unsworth & Wheeler, 2002; Van Leeuwen, 2008).

Different relations between the viewers and the social actors can be implied by distance which can indicate an interpersonal relationship or closeness of our relationships. The close shot creates a kind of intimacy between the viewer and the depicted participants and makes us believe that the participant is 'one of us'. The long shot, on the other hand, creates impersonality and shows the participants as strangers to the viewers. The medium shot, however, is more social and objective (Kress & Van Leeuwen, 2006 [1996]; Moebius, 1986; Nodelman, 1988, 1999; Van Leeuwen, 2008). There are, in fact, varying degrees of distance among the represented and the interactive participants, each carrying a distinctive meaning. The extreme close-up shot which is anything less than the social actor's head and shoulders implies a sense of intimate distance. The close-up which shows the social actor's head and shoulders is an indication of a close personal distance. The medium close shot which cuts off the represented participants at the waist places them at a far personal distance. The medium shot, from knees up, is closer to a social distance. The medium long which shows the full figure of the participant is the indication of a close social distance. The long shot, where the human figure occupies

about half the height of the frame, implies a far social distance. And anything wider than that, which is the very long shot where a group of people are depicted in an image, implies a public distance (Kress & Van Leeuwen, 2006).

Regarding the angle from which we see the person, a horizontal angle is whether the viewer sees people from the side or front. The vertical angle means whether the viewer sees the participants from below, above, or at eye-level. Representing social relation between the viewers and social actors depicted in images, a vertical angle is an indication of power differences and a horizontal angle is an indication of involvement and detachment (Caldas-Couldhar & Iedema, 2008; Iedema, 2001; Kress & Van Leeuwen, 2006 [1996]; Royce, 2007; Van Leeuwen, 2008). In other words, looking down at somebody can indicate exerting an imaginary symbolic power over that person. Looking up at somebody, on the other hand, means somebody has power over the viewer. To look at somebody at the eye-level indicates equality. Moreover, seeing the depicted people from side or front might imply coming face to face with them, or confronting them (Camiciottoli, 2007; Royce, 2007).

Another factor which is of importance at the interpersonal level is modality which describes the degree of 'credibility' and 'truthfulness' of a visual text. Light, texture, and shade can give an image high modality, high truthfulness or credibility. Idealization, decontextualisation, and perspective are other influences on modality (Royce, 1998; Kress & Van Leeuwen, 2006). Visual images usually illustrate varying degrees of modality along a continuum from highest (most real or credible) to lowest (least real or credible) (Royce, 1998, p. 40).
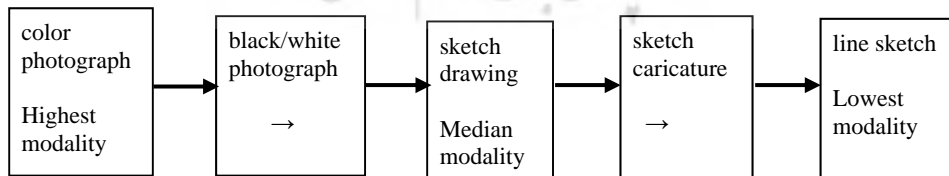
color photograph Highest modality	→	black/white photograph →	→	sketch drawing Median modality	→	sketch caricature →	→	line sketch Lowest modality

Figure 1. Modality in the naturalistic visual continuum, adapted from Royce, (1998, p. 40)

The third metafunction, *textual*, is a combination of the two previous metafunctions: it brings what is represented (ideational) together with interactive components (interpersonal). In other words, it looks at the way in which the first two metafunctions merge together to form a meaningful whole through three interrelated systems: information value, salience, and framing (Kress & Van Leeuwen, 2006 [1996]; Royce, 2007; Unsworth, 2008; Unsworth & Wheeler, 2002).

Information value means the placement of elements in relation to one another. For instance, depicting people on the left side of images might be an indication of old and given information, whereas people depicted on the right side might be considered as new information. Or elements placed at the top of the images might be considered as ideal and promising while those at the bottom part might focus more on the factual and real. Salience deals with how the elements depicted in the images are made to ask for the viewers' attention and retention. Framing also refers to the existence of any dividing lines or actual frame lines trying to connect or disconnect the elements in an image. This can be an indication of whether the elements in the image belong together or not (Kress & Van Leeuwen, 2006 [1996]; Moebius, 1986; Nodelman, 1988, 1999; Van Leeuwen, 2008).

Further, in an attempt to find out how social actors are depicted in images, Van Leeuwen (2008) adds the following dimensions and criteria in his grammar of visual design:

a) Exclusion: it means what kind of people have been included in the images and what specific people who are naturally living and working among us have been excluded from the images.

b) Roles: that means whether people in images are involved in performing an action, and if they are the agents and doers of the action or the patients and the ones to whom the action is being done?

c) Specific and generic: this dimension, which can specifically be used in identifying any potential racism, means whether people are depicted as specific people or are generally shown as generic people. That means whether we are supposed to generalize the qualities attributed to depicted people in images as a specific and individual person or to their certain social type and kind of people.

d) Individuals and groups: it means whether people are depicted as individuals or as groups. If people are depicted as groups, it can show that they are all the same and similar to each other with minute differences.

e) Categorization: that is whether people are categorized in terms of 'biological' or 'cultural' characteristics or based on a combination of both.

As stated by Van Leeuwen (2008), distance, gaze, and angle can be combined in multiple ways indicating many different ways of depicting people as 'others'. And he goes on to mention three possible strategies that can be used to visually represent people as 'others' and as 'not like us': *distanciation* showing social actors as strangers and not close to us; *disempowerment* depicting people as below us and downtrodden; and *objectivation* representing people as object of our scrutiny rather than as subjects addressing the viewers with their gaze (Van Leeuwen, 2008, p. 141). In other words, Van Leeuwen (2008) believes that these criteria can be used as strategies for visually depicting and representing people as 'others' (p. 147):

• the strategy of exclusion, not representing people at all in contexts where, in reality, they are present

• the strategy of depicting people as the agents of actions which are held in low esteem or regarded as subservient, deviant, criminal, or evil

• the strategy of showing people as homogeneous groups and thereby denying them individual characteristics and differences ("they're all the same")

• the strategy of negative cultural connotations

• the strategy of negative racial stereotyping

Method

Corpus of the Study

The four software applications of *Rosetta Stone, Tell Me More, Fairyland,* and *Your Baby Can Read* have been chosen as the sample of this research among only a very few software applications available on the market. These software applications have been chosen for this study for some reasons. They are the best-selling software packages in Iran and the most award-winning in the world. Being the most popular software applications among the ones available on the market, *Rosetta Stone, Tell*

Me More, Fairyland, and Your Baby Can Read enjoy attracting the largest audience of different ages, from children to adults, and various language proficiency levels ranging from beginners to advanced learners (*www.pipcenter.com*). Moreover, they have been verified and authorized by the Ministry of Culture and Islamic Guidance as the most appropriate and useful language learning software applications among the ones currently available on the market (*www.pipcenter.com*).

The Analytical Framework

In this research, Kress and Van Leeuwen's (2006 [1996]) and Van Leeuwen's (2008) visual grammar and social semiotic model have been drawn upon to interrogate ways through which social actors of different races are visually represented in the abovementioned software packages. The conceptual theoretical frameworks of the study are presented in figures 2 and 3 (pp. 13 & 14).

Design of the Study

Using a multimodal critical discourse analysis approach, the present study is an exploratory interpretive research including mainly qualitative investigations of the way people of different races are visually and verbally represented in four sets of English learning software applications. The analytical framework is primarily based upon the social semiotic model.
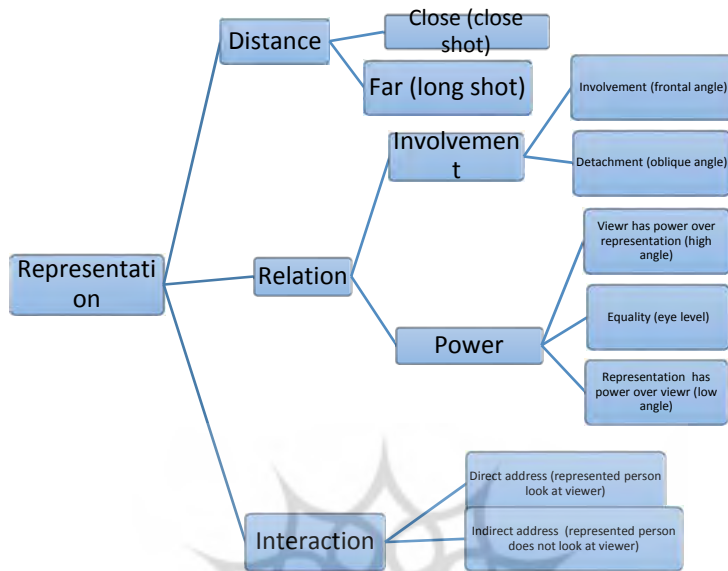
Figure 2. Representation and viewer network, adapted from Van Leeuwen (2008: 141): Discourse and practice
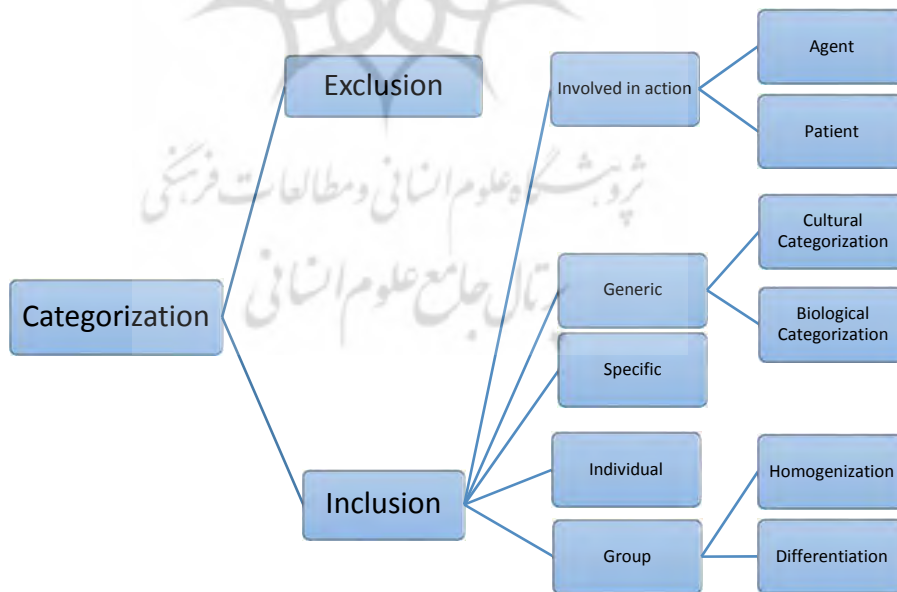

Figure 3. Visual social actor network, adapted from Van Leeuwen (2008: 141): Discourse and practice

Procedures

In order to investigate how social actors are visually and textually represented in the multimodal educational contexts, all the visual images which contained human participants -accompanied by their texts- were extracted from the four sets of software applications which constituted 4535 images altogether as the population of this study. For this research, the races of the represented participants were determined based on their visually identifiable attributes or verbally mentioned ones. Hence three general types of white/Caucasian, Asian/Mongolian, and black/Negroid were identified and the population was categorized into these three strata. Next, through a stratified sampling procedure, a random sample was drawn from each stratum which constituted 142 images altogether. For each software application, the social semiotic analysis of the selected sample of the three groups of race was carried out using Kress and Van Leeuwen's (2006 [1996]) and Van Leeuwen's (2008) social semiotic model.

As for the visual analysis at the ideational level, the interactions and relations between the social actors, places and things depicted in images were investigated. The social actors in the images participated in the interaction through two main processes of narrative and conceptual. In narrative processes the social actors were involved in doing something either to or for each other. Vectors were used to link participants to one another or to the processes. On the basis of different vectors and the number and kinds of participants involved in the interaction, the following narrative processes were identified.

In action processes the participants who formed the vector or were the vector itself were considered as the Actor. Then it was investigated whether the action process was transactional or non-transactional. In transactional processes, in addition to the Actor, the Goal was identified as well. In reactional processes, where the vector was formed by an eye line, the Reactors and Phenomena were identified. In circumstances, the secondary participants that were not related to the main participants by means of vector but in other ways including circumstances of means, locative circumstances, and circumstance of accompaniment were identified.

Regarding the visual analysis at the interpersonal level, the interaction between the social actors and the interactive participants was investigated. In order to do this investigation, four factors were considered (Kress & Van Leeuwen, 2006; Van Leeuwen 2008): gaze, angle, distance, and modality. Besides these factors, three possible strategies of visually representing people as 'others' and as 'not like us' were investigated: *distanciation*; *disempowerment*; and *objectivation*.

As for the visual analysis at the textual level, three interrelated systems of information value, salience, and framing were investigated. Moreover, in an attempt to find out how social actors are depicted in images, the following dimensions and criteria were taken into account: Exclusion, Roles, Specific and Generic, Individuals and Groups, and Categorization.

In order to maintain an acceptable index of reliability in data analysis, the data were analyzed by two other trained coders.

Results and Discussion

The analysis presented here focuses on the ideational, interpersonal, and textual meanings in multimodal texts. For this research, Kress and Van Leeuwen's (2006 [1996]) visual grammar and Van Leeuwen's (2008) social semiotic model have been drawn upon to interrogate ways through which social actors of different races are visually and textually represented in the English-learning software packages.

Rosetta Stone. Although *Rosetta Stone* has included a variety of different races, it seems that it is biased towards some specific ones. Tribal people, Arabs, Latin Americans, Afro-Asians and Far East Asians are 'others' in their own contexts. The only exception here is a Far East Asian from China, photo 1, where the image of the old man wearing Chinese costume is a 'demand'. However, they are not depicted as 'others' at all in American (western) contexts. For instance, in photo 2, where a group of children wearing American flag-like clothes are waving the U.S. flag, the most salient active participant is a Far East Asian girl who is involved with us because of her gaze and hand vector and is shown as belonging to our world.

This man is very old.

They are from the United States.

The Egyptian businessman is drinking tea.
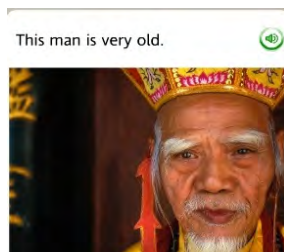
Image 1   Image 2   Image 3

In photo 3, where an Egyptian businessman is shown in a western-like context - wearing a formal suit and working on a computer in an office - he is not only a specific individual, but also has some power over the viewers since we are looking at him from a low angle. But when these people are represented in their own local contexts, they are shown as not belonging to our worlds, as strangers and disempowered ones. For instance, in case of Arabs, usually, the viewer has no involvement with them and they are offered as strangers and objects of scrutiny for the viewers to think about rather than as subjects addressing the viewers with their gaze. This has been done through the strategy of *objectivation* (Van Leeuwen, 2008). In photo 4, the children are welcoming their father, yet neither the father nor the children have any individuality. The children are seen from the back. The father's face is not seen clearly. The viewers see them from a high angle which suggests the interactive participants in power. This is done through the strategy of '*disempowerment'* which depicts social actors as below us and downtrodden (Van Leeuwen, 2008). They are shown generically as Arabs wearing dishdasha from a far social distance. In photo 5, two Arabs are greeting in the street. Due to the oblique angle, we have a side view of them and cannot see their faces clearly. They are once more offered as generic strangers and 'others' through the strategy of objectivation. The strategy of depicting social actors in a generic way as homogeneous groups, which in Van Leeuwen's (2008) idea can be used in identifying any potential racism, deprives them of any individual characteristics and differences and implies that the viewers are supposed to generalize the qualities attributed to the depicted people in images to their certain social type and kind of people as if

'they're all the same' (Van Leeuwen, 2008, p. 147). In photo 6, a father and his son are accompanied by their texts shown in a long shot from an oblique angle. It seems they do not belong to our world.

Image 4

Image 5

Image 6

Image 7

Even in cases where we are observing an Arab social actor from a medium shot, we are not involved either because he is wearing sunglasses (photo 7) or because he is shown from an oblique angle (photo 8).
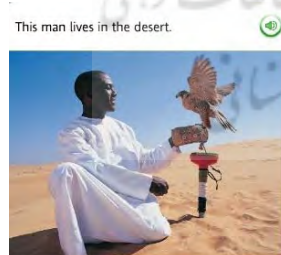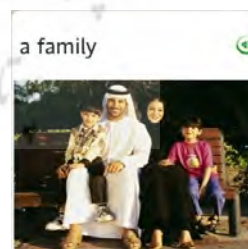

Image 8

Image 9

We as viewers are involved with Arabs only in one case (photo 9) where they are shown as a family. However, the long shot depicts them distant from us, therefore, as 'others' and strangers. Regarding Arabs, the only role allocated to them is that of a social one; they are shown either in

a family or on the street greeting or talking on the phone. In other words, they have been 'symbolically excluded' (Van Leeuwen, 2008) from certain roles and have been confined to some specific and predetermined ones.

Like Arabs, Latin Americans, Far East Asians, Afro-Asians and Tribal people are also depicted as strangers and 'others' through strategies of objectivation, distanciation, and disempowerment of the represented participants. In photo 10, a couple of Latinos are shown in Mexican traditional costumes dancing on the flag of Mexico. The viewers see them from a very high angle which indicates the disempowerment of the represented participants. Moreover, they have no individuality since their faces are not seen at all, and they are shown as generic Mexicans who are 'all the same', below us and downtrodden or as passive objects of scrutiny for the viewers to think about (Van Leeuwen, 2008). Photos 11, 12, and 13 portray Far East Asians as 'others', not belonging to our world. Although in photo 4.11 we have a close-up of the social participant, we are not involved with her since her gaze is away from the viewers looking down. The role allocated to her is that of passive object for the viewers to think about.

| Image 11 | Image 12 | Image 13 |

Image 10

Photo 12 depicts a Far East Asian dancer again in a cultural generic way having a local costume, makeup and hair style. Although the role given to her is an active one, she has been shown impersonally for the viewers to look at. Image 13 is a long shot of a Far East Asian woman distant from us in an oblique angle. She has a passive role and no involvement with the viewers. The mood in this image is an 'offer' like the previous two images. In photo 14, however, we are involved with a Far East Asian couple through their gaze, but we are looking at them from a high angle, therefore, having power over them. Afro-Asians in photos 15 and 16 do not have any individuality at all and are not involved with us. The extreme long shot pictures them as a cultural generic group of nomads having a far social distance from us.

Image 14   Image 15   Image 16

Image 17

In photo 17 where the participant is shown from a closer distance, we are not involved with him because his face is covered with a bandana. In all cases the horizontal angle is an oblique one which indicates that the represented participants do not belong to the world of the viewers. Photos 18 and 19 represent some tribal men in the generic role of hunters carrying darts who are for us to think about. There is one exception to this pattern where the represented participants are depicted as 'others' and strangers

who do not belong to the world of the viewer, yet they are powered over the viewers. For instance, photos 20 and 21 show two Jewish men in a synagogue performing their religious rituals who are not looking directly at the viewers, but are shown from a low vertical angle which implies a sense of power over the viewers on the part of the represented participants.
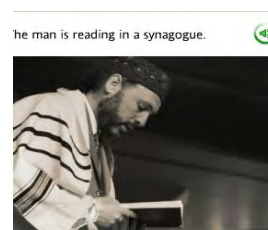

Image 19   Image 20   Image 21

Image 18

Images, as stated by Berger (1972), can be used to position and manipulate the viewers (Belknap & Leonard, 1991; Sheldon, 2004). Hence it is not unlikely that Rosetta Stone can powerfully influence public opinions, values, and ideologies through a distorted and biased way of representing Arabs, Latin Americans, Far East Asians, afro-Asians and tribal people (Caldas-Coulthard, 2003; Gamson et al., 1992) and can change viewers' attitudes toward these certain groups of people with whom they have little or no direct contact.

When it comes to occupations, however, not only are various races fairly included in almost all occupations including low and high class jobs, but they are also involved with us in most cases. Different races including white/Caucasian, Asian/Mongolian, and black/Negroid are depicted as managers, doctors, teachers, employees, police officers, waiters/waitresses, etc. They have the maximum involvement with their

viewers and there is no power distance between the represented and the interactive participants (photos 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, and 35).

Image 22	Image 23	Image 24	Image 25
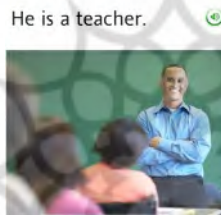
Image 26	Image 27	Image 28	Image 29
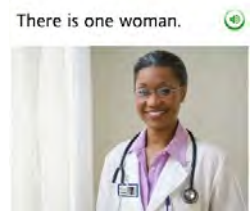
Image 30	Image 31	Image 32	Image 33

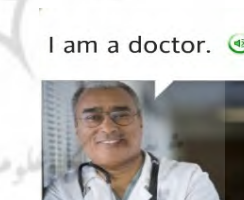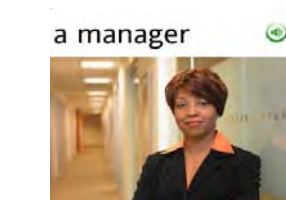|	
Image 34	Image 35

There are two exceptional cases here. One is regarding a black police officer whose gesture and pose for the photo shows him in power regarding

the viewers (photo 36). The other exception is white managers who are in power due to the way they look down at the viewers (photos 37, 38, and 39).
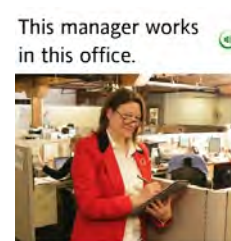

Image 36

Image 37

Image 38

Image 39

It should be noted that in all these cases, people of different races are shown in western-like contexts, wearing western-like clothes or formal and/or professional uniforms. As scholars such as Fairclough (1995) or Van Dijk (2001) believe, there are hidden ideologies in the way that 'reality' is represented in discourse and it is important to uncover these implicit ideologies behind the messages and understand what meanings are constructed by the readers and audience. The two completely opposing ways of representing different races in *Rosetta Stone* – 'others' and strangers in their own local contexts, and 'one of us' in western-like contexts – might imply that people of different races can be considered as 'one of us' providing that they conform to the rules and regulations of western society and culture.

Tell me More. *Tell me More,* unlike *Rosetta Stone,* has excluded almost all races except White Americans. That means only white Americans have been included in the images and people of other races who

are naturally living and working among us have been excluded and this act of including some specific social actors and excluding the others serves for the intentions and purposes of the software designers in relation to the readers and viewers for whom the images are produced (Van Leeuwen, 2008). In other words, it can be interpreted as a multimodal means of manipulating people and controlling their minds through biased or incomplete images and meanings they convey (Caldas-Coulthard, 2003; Van Dijk, 2006).

There are very few cases of Black and Far East Asians who have been represented in biased and distorted ways. Regarding Far East Asians, few images are included in this software (photos 40, 41, 42, 43). As the verbal mode indicates, image 40 is supposed to represent a Far East Asian having a business lunch with a white American in New York; what it actually shows is two figures in a restaurant from an extremely high and oblique angle having a public distance with us. There is no individuality in this image since their faces are not seen at all and it is almost impossible to detect the actor or goal of the action. Images 41 and 42 represent a medium close shot of two Far East Asians having a far personal distance with us. Although both participants play the role of 'actors' and the social actor in photo 42 is involved with us, they are shown from an oblique and high angle which represents the interactive participants' power over them. Photo 43 is different from the previous photos in that it represents a close-up of the social actor from a straight angle. Though we have an intimate distance with the represented participant, we are not involved with him as the mood is 'offer'.

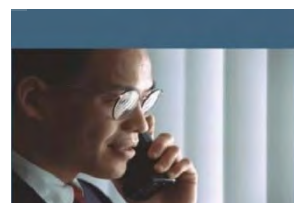Image 40   Image 41   Image 42   Image 43

As indicated by the image analysis, Far East Asians in *Tell Me More* are either shown from an oblique and high angle which represents the interactive participants' power over them or from a straight angle but in an 'offer' mood, implying that the social actor is not involved with the interactive participants. Hence, through the strategies of 'disempowerment' and 'objectivation', they have been represented as people who are below us and downtrodden and people who are strangers and 'others' and do not belong to our world but are objects of scrutiny for the viewers to look at (Van Leeuwen, 2008).

Regarding the black people, again very few cases have been included in this software (photos 44, 45, 46, 47, 48, 49, 50). In photo 44 there are two basketball players, a white one and a black one, equally salient in the image. The placement in the image and hand vector of the black one indicates his power over the white participant. Much in the same way, photo 45 shows us a cultural generic black musician from a low angle implying a sense of power on the part of the represented participant over the viewers. However, in both cases where the black participants seem to be dominant, they are shown stereotypically in typical roles associated with black people; that is, basketball players and jazz musicians. When it comes to black people in occupational contexts, only three job positions have been allocated to them, real estate agent, hotel receptionist, and doctor (photos 46, 47, 48). Nevertheless, they are shown as strangers and disempowered ones through the strategies of 'objectivation' and 'disempowerment' (Van Leeuwen, 2008). Photo 46 is a medium close shot of a real estate agent from a straight oblique angle who has no involvement with the viewers and is seen as a stranger. In the same way, the hotel receptionist in photo 47 is not only detached from the viewers because of the 'offer' mood but is also disempowered by the high angle from which he is shown.

Image 44   Image 45   Image 46
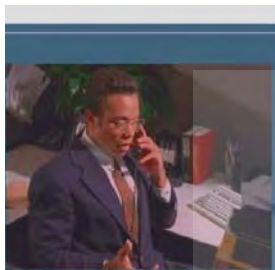

Image 47   Image 48   Image 49

When they are depicted in high class job positions, there is a lack of any circumstances of means which could indicate their professionality. Even though it seems image 48 represents a black woman in a high class job position (doctor), she is not shown as a professional since there are no circumstances of means indicating her professionalism such as medical uniform, stethoscope, spatula, etc. The only case where black people are involved with the interactive participants is when they are shown as students (photos 49 and 50). Once more, they have been represented as 'others' and strangers either through the strategy of 'symbolic exclusion' – by being confined to some specific roles and deprived of the others – or through the strategy of 'decontextualization' – by taking any contexts and signs of professionalism away from them (Van Leeuwen, 2008).

American-Indians have been represented in only three cases (photos 51, 52, and 53). Image 51 shows an Indian tepee with no mention of any Indians. That is, the Indians have been depicted as strangers and 'others' through the strategy of exclusion (Van Leeuwen, 2008).

| Image 50 | Image 51 | Image 52 | Image 53 |

Photo 52 illustrates a vague a line sketch of a homogenized group of American-Indians from a very far distance in such a way that they have no individuality, as if 'they are all the same' (Van Leeuwen, 2008), and the houses are more salient than the people. In image 53 where we have a nonrealistic portrait of a historical American-Indian (Pocahontas), there is no involvement between the represented participant and the viewers to imply a sense of detachment between her and the interactive participants (Kress and Van Leeuwen, 2006 [1996]).

Fairyland. This software has tried to include a good variety of different races especially as its main characters. Nevertheless, none of them is represented culturally different from the other characters and there is no visually or textually specific cultural sign indicating to which race they belong and they have all been illustrated in a western cultural setting. The only sign differentiating them from each other is their generic biological characteristics with some extent of exaggeration. For instance, four of the main characters are Emma, a white girl with blond hair and big blue eyes (image 54); Mona, a girl with dark complexion, curly crispy hair, and very thick protruding lips (image 55); Harry, a white boy with red hair and big blue eyes (image 56); and Lee, a Far East Asian boy with black hair and highly slanted eyes (image 57). The major character of the software is a female fairy called Erlina who is white with blond hair and big blue eyes (image 58).
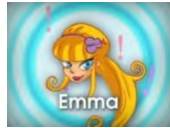
Image 54   Image 55   Image 56   Image 57   Image 58

When these characters are introduced to the viewers, they are all shown from a straight frontal angle and a close personal distance. The only characters who have involvement with the interactive participants through their gaze are Erlina and Emma (images 54, 58), the white girls. Nevertheless, children of different races are seen together as close friends, classmates, and teammates in everyday life settings including school, Animal Park, home, parties, etc. (images 59, 60, 61, 62, 63, 64, 65, 66, 67). In image 59, a group of soccer players of different races have been depicted who are all involved with the viewers through their gaze. In the same way, image 60 shows a group of mothers and their kids who are holding each other's hands. They are of different races and the picture is again a 'demand'. It is noticeable that the image of the globe behind the kids and their mums and the fact that they are all holding each other's hands implies that people of different races all over the world can be friends.

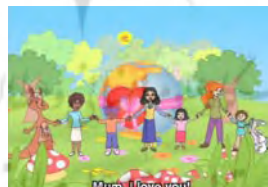Image 59   Image 60   Image 61

Image 61 illustrates the four main characters of the software on a safari boat from a far personal distance. As the participants are shown from a straight frontal angle and the image is a demand, the represented and interactive participants are involved with each other. Besides, Harry's hand vector invites the viewers to join. Images 62 and 63 illustrate the four main characters the same, frontally from a far personal distance, communicating with each other.

| Image 62 | Image 63 | Image 64 |

In image 64, not only are children of different races shown in the same setting, but also the black boy is more salient than the others, as he is depicted in the center of the picture. He is also the 'agent' of the act of singing. Images 65, 66, and 67 represent children of different races introducing each other as friends to the viewers.
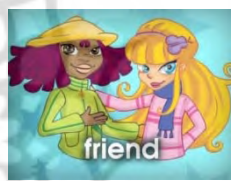

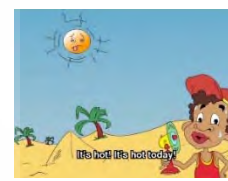| Image 65 | Image 66 | Image 67 | Image 68 |

It seems that *Fairyland* is following the same policy of *Rosetta Stone* regarding different races. Though people of different races are fairly included in this software, like that in *Rosetta Stone,* a critical semiological analysis reveals the implicit ideologies behind the messages both software packages are trying to convey and impose on their viewers (Fairclough, 1995; Van Dijk, 2001): a message which implies that people of different races can be considered as 'one of us' and involved with us to the extent they totally assimilate to and integrate with western culture and norms. One interesting point which needs to be noticed regarding race is an image showing a black man in a hot desert. Unlike *Rosetta Stone* that represented people in the context of desert culturally generic with no individuality, *Fairyland* illustrates a stereotyped depiction of a black man from a close personal distance who is involved with the viewers. His crispy hair, brown complexion, big flat nose, and very thick protruding lips make him biologically generic. Furthermore, *Rosetta Stone* always showed people in

this setting as primitive ones usually riding camels and carrying water pots on their heads. *Fairyland,* on the other hand, represents the black man with western clothing and an electrical hand fan.

Your Baby Can Read. Not a variety of different races are included in this application. The only two races included are a couple of white children and a few black ones. Hence this software might be considered as biased regarding people of different races. However, among the ones included, *Your Baby Can Read* does not seem to be biased regarding races. For instance, in order to introduce the word *bib*, both a black baby and a white one are shown from a close personal distance (photos 69, 70).
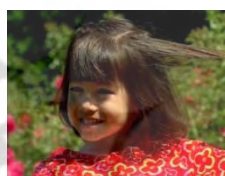

Image 69   Image 70   Image 71   Image 72

Likewise, to show the word *hair*, Your *Baby Can Read* has represented the close-up of both a white girl and a black one through straight frontal angle (photos 71, 72). In photo 73 four children who are of the two races invite the viewers to involve with them through their hand and gaze vectors.

Image 73   Image 74   Image 75

In photos 74 and 75, in the same way, two black and white children are shown from a personal distance. They are both involved with the interactive participants. The same pattern repeats itself in the whole

application, and almost every single concept is introduced by both white and black children equally. Both races have equal power distance with the interactive participants, are equally involved with the viewers, and are shown from the same social distance. This seems to agree with the findings of Ollivier (1992) who concluded that textbooks for children do not contain much stereotyping and are not as biased regarding gender as are the ones for adults.

Having reported the results of the study and offered their interpretations as to the ways images convey meaning at the three metafunctions of ideational, interpersonal, and textual, we now take a step back to see the whole picture created by these four software packages and to investigate their representing people of different races.

Conclusions

The aim of this study was to explore the ways software packages as educational multimedia have positioned, represented, and depicted people of different races through their images. The analysis of findings have revealed that when it comes to educational materials designed specifically for children such as *Fairyland* and *Your Baby Can Read*, we cannot find explicit traces of racial bias and stereotyping in either verbal or visual modes. This seems to agree with the findings of Ollivier's (1992) study on elementary school textbooks. She concluded that textbooks for children did not contain much stereotyping and were not as biased regarding gender as the ones for adults.

When it comes to teaching materials for adults, on the other hand, the results are different. Out of the four sets of educational software packages, *Tell Me More* is the only software which has bluntly shown its biased attitude towards races. It has manifested discrimination towards non-whites explicitly (Wodak, 1995, cited in Blommaert, 2005). As it was previously stated by Kordjazi (2012) that this software was biased towards gender, it seems that it is promoting inequity regarding race too. A quick look at the images and content of *Rosetta Stone* gives viewers the impression that it is not biased at all regarding different races. However, a deep and comprehensive analysis of the software has revealed hidden traces of discrimination, bias and stereotyping towards some specific races

(Wodak, 1995, cited in Blommaert, 2005) especially when they are represented in their own local and cultural contexts. It can be implied that the non-whites have been looked down by the producers and developers of these software applications which might lead to the same detrimental effect on the viewers' attitude as well (Lester, 2000).

Overall, what the current research conveys is that bias and stereotyping in educational media is strong and ever present. Although the multimodal educational contexts may no longer be explicitly biased and stereotypical regarding various races as it was the case with *Tell Me More*, they can still manifest bias and stereotypes implicitly in their visual mode. In the verbal mode of these four software packages, as the finding analysis shows, there is no sign of racial bias or discrimination and the texts are to a great extent neutral regarding different races. It seems that traces of racism are more easily detectable in the verbal mode than in the visual mode and people are more sensitive to verbal racism than to the visual one as Van Leeuwen (2008, p. 137) put it:

> If this reflects how contemporary popular culture ("naively or not") construes the differences between images and words, then it means, for instance, that visually communicated racism can be much more easily denied, much more easily dismissed as "in the eye of the beholder" than verbal racism.

Therefore, software developers seem to be more cautious in their verbal choices than their visual ones while the fact that visuals can have a great impact on viewers' perceptions by conveying stereotypical and biased messages in a subtle and implicit way may go unnoticed and taken for granted (Lester, 2000; Van Leeuwen, 2008). Educational materials can manipulate language learners and control their minds through biased or incomplete images and meanings they convey (Van Dijk, 2006), even if intentionally. They can also change language learners' attitudes toward these groups of people with whom they may have little or no direct contact. Being constantly exposed to biased, distorted, and stereotypical portrayals of various races and ethnic groups, the learners tend to receive these representations as natural (Griffiths, 2010).

The findings of this research can be of potential help and use for students and teachers to become visually literate and get aware of the hidden messages communicated by images in textbooks and multimedia. Material developers and software designers can also benefit from the results of this study by being aware of the messages conveyed by their chosen visual elements and the effects they might have on the learners and by being more cautious in producing and choosing images and illustrations for their software. Researchers and pedagogues, too, can employ the findings of this research in developing programs in schools or universities to make students, and teachers as well, visually literate and make them question learning tools by not taking them for granted.

References

Abbas-Nejad-Konjin, A. (2012). *Gender analysis of Iranian middle school textbooks.* (Unpublished MA thesis). UBC, Vancouver, Canada.

Adami, E. (2013). *A social semiotic multimodal analysis framework for website interactivity.* London: National Center for Research Methods.

Ansary, H., & Babaii, E. (2003). On the manifestation of subliminal sexism in current Iranian secondary school ELT textbooks. *Iranian Journal of Applied Linguistics, 6*, 57-69.

Babaii, E., & Ansary, H. (2001, March). The structure of and stricture on TV Commercials in Iran. *Paper presented at the 5th Conference on Theoretical and Applied Linguistics.* Tehran: Allameh Tabataba'i University.

Baldry, A. P. (2000). *Multimodality and multimodality in the distance learning age.* Campobasso, Italy: Palladino Editore.

Baldry, A. P. (2004). Phase and transition type and instance: patterns in media texts as seen through a multimodal concordancer. In K. L. O'Halloran (Ed.), *Multimodal discourse analysis* (pp. 83-108). London: Continuum.

Baldry, A. P., & Thibault, P. J. (2006). *Multimodal transcription and text analysis*. London: Equinox.

Barthes, R. (1977). *Image, music, text.* London: Fontana/Collins.

Belknap, P., & Leonard, W. M. (1991). A conceptual replication and extension of Erving Goffman's study of gender advertisements. *Sex Roles, 25*(3/4)103-118.

Berger, J. (1972). *Ways of seeing*. London: BBC/Harmondsworth: Penguin.

Binns, J. C., & Branch, R. C. (1995). *Gender stereotyped computer clip-art images as an implicit influence in instructional message design.* ERIC: ED 3080089

Bishop, H., & Jaworski, A. (2003). We beat 'em': nationalism and the hegemony of homogeneity in the British press reportage of Germany versus England during Euro 2000. *Discourse & Society, 14*(3), 243–271.

Blommaert, J. (2005). *Discourse: A critical introduction.* Cambridge: Cambridge University Press.

Brand, J. E., Knight, S., & Majewski, J. (2003). The diverse worlds of computer games: A content analysis of spaces, populations, styles and narratives. *Paper presented at Digital Games Research Conference,* The Netherlands. ISBN / ISNN: ISSN 2342-9666.

Caldas-Coulthard, C. R. (2003). Cross-Cultural Representation of 'Otherness' in Media Discourse. In G. Weiss & R. Wodak, (Eds.), *Critical discourse analysis:Theory and interdisciplinarity* (pp. 272-296). New York: Palgrave McMillan.

Caldas-Coulthard, C. R., & Iedema, R. (2008). *Identity trouble: Critical discourse and contested identities.* New York: Palgrave Macmillan.

Camiciottoli, B. C. (2007). *The language of business studies lectures.* Amsterdam / Philadelphia: John Benjamins Publishing Company.

Canning-Wilson, C. (1999). *Using pictures in EFL and ESL classrooms.* ERIC No: ED 445526. Retrieved from the ERIC database

Chappell, K. K. (1996). Mathematics computer software characteristics with possible gender-specific impact: A content analysis. *Journal of Educational Computing Research, 15,* 25–35.

Chiew, A. (2004). Multisemiotic mediation in hypertext. In K. L. O'Halloran (Ed.), *Multimodal discourse analysis: Systemic-functional perspectives* (pp. 131-162). London: Continuum.

Coelho, Z. P. (2008). Front page layout and reading paths: The influence of age on newspaper reading. *Estudos em Comunicação, 4,* 1-14.

Drees, D. E., & Phye, G. D. (2001). Gender representation in children's language arts computer software. *Journal of Educational Research, 95*, pp. 49–55.

Fairclough, N. (1989). *Language and power.* London: Longman.

Fairclough, N. (1995). *Critical discourse analysis: The critical study of language.* London and New York: Longman.

Gamson, W. A., Croteau, D., Hoynes, W., & Sasson, T. (1992). Media images and the social construction of reality. *Annual Review of Sociology, 18,*373-393.

Gharbavi, A. & Mousavi, S. A. (2012). The application of functional linguistics in exposing gender bias in Iranian high school English textbooks. *English Language and Literature Studies, 2*(1), 85-93.

Goffman, E. (1979). *Gender advertisements*. Cambridge, MA: Harvard University Press.

Graci, J. P. (1989). Are foreign language textbooks sexist? Exploration of models of evaluation. *Foreign Language Annals, 22*, 77-86. Retrieved from http://dx.doi.org/10.1111/j.1944-9720.1989.tb02771.x

Griffiths, Th. (2010). *Media representation theory*. Retrieved from *http://www.slideshare.net/Tomaskas/media-representation-theory*.

Guijarro, J. M., & Sanz, J. M. P. (2008). Compositional, interpersonal and representational meanings in a children's narrative: A multimodal discourse analysis. *Journal of Pragmatics*, *40,*1601–1619.

Halliday, M. A. K. (1978). *Language as social semiotic: The social interpretation of language and meaning*. London: Edward Arnold.

Halliday, M. A. K. (1985). *Introduction to functional grammar*. London: Arnold.

Hartman, P. L., & Judd, E. L. (1978). Sexism and TESOL materials. *TESOL Quarterly*, *12*(4), 383-393.

Hellinger, M. (1980). For men must work, and women must weep: Sexism in English language textbooks used in German schools. In C. Kramerae (Ed.)*, The voices and words of women and men* (pp. 267-274). New York: Pergamon Press.

Hietala, V. (1996). *The reality of the images. Introduction to the understanding and interpretation of visual culture.* Helsinki: Kirjastopalvelu.

Hjorth, J. (1997). *Presentation of females and males in an English textbook series for Finnish upper-secondary schools* (Unpublished Pro Gradu Thesis). University of Jyväskylä, Finland. Retrieved from *https://jyx.jyu.fi/dspace/handle/123456789/7366*

Horváth Futó, H. (2011). Textbooks and stereotypes. *Philologica, 3*(2), 265-278.

Iedema, R. (2001). Analyzing film and television: A social semiotic account of hospital: An unhealthy business. In T. Van Leeuwen, & C. Jewitt (Eds.), *Handbook of visual analysis* (pp. 183-204). London: Sage.

Iedema, R. (2003). Multimodality, resemioticization: extending the analysis of discourse as a multisemiotic practice. *Visual Communication, 2*(1), 29–57.

Jarić, I. (2002). Gender Stereotypes. *Nova srpska politička misao, 2*(Special issue), 5-19.

Jewitt, C. (2004). Multimodality and new Communication technologies. In P. LeVine & R. Scollon (Eds.) *Discourse and Technology: Multimodal discourse analysis* (pp. 184-195). Washington: Georgetown University Press.

Jewitt, C. (2005). Multimodality, 'Reading', and 'Writing' for the 21st Century. *Discourse: Studies in the Cultural Politics of Education, 26*(3), 315-331.

Jones, R. S. (2005). Sites of engagement as sites of attention: time, space and culture in electronic discourse. In S. Norris, & R. H. Jones (Eds.), *Discourse in action. Introducing mediated discourse analysis* (pp. 141-154). London: Routledge.

Kern, R. (2006). Perspectives on Technology in Learning and Teaching Languages. *TESOL Quarterly, 40*(1), 183-210.

Knox, J. S. (2009). Punctuating the home page: image as language in an online newspaper. *Discourse & Communication, 3*(2), 145–172.

Kordjazi, Z. (2012). Images matter: A semiological content analysis of gender positioning in contemporary English-learning software applications. *Novitas-ROYAL 6*(1), 59-80.

Kress, G., & Van Leeuwen, Th. (2006 [1996]). *Reading images: The grammar of visual design (2nd Ed.).* London: Routledge.

Lee, J. F. K., & Collins, P. (2006). *Gender Representation in Hong Kong English Textbooks.* Retrieved from http://www.eoc.org.hk/eoc/upload/2006711112336211184.pdf.

Lemke, J. (1998). Multiplying meaning: visual and verbal semiotics in scientific text. In R. Martin & R. Veel (Eds.), *Reading science: critical and functional perspectives on discourse and science* (pp. 87-113). London: Routledge.

Lemke, J. (2002). Travels in hypermodality. *Visual Communication, 1*(3), 299–325.

Lester, P. (2000). *Visual communication: Images with messages.* Belmont. CA: Wadsworth/Thomson.

Lim Fei, V. (2007). The visual semantics stratum: making meaning in sequential images. In T. D., Royce, & W. L., Bowcher (Eds.), *New Directions in the analysis of multimodal discourse* (pp. 195–213). Mahwah: Lawrence Erlbaum.

Lim Fei, V. & O'Halloran, K. (2010, Accepted for Publication). The ideal teacher: An analysis of a teacher-recruitment advertisement. *Semiotica.*

Liu, Y., & O'Halloran, K. L. (2009; accepted for publication). Inter-semiotic Texture: Analyzing cohesive devices between language and images. *Social Semiotics.*

Machin, D., & Mayr, A. (2007). Antiracism in the British government's model regional newspaper: the 'talking cure'. *Discourse & Society, 18*(4), 453–477.

Machin, D., & Van Leeuwen, T. (2009). Toys as discourse: Children's toys and the war on terror. *Critical Discourse Studies, 6*(1), pp. 51-63.

Mazid, B. M. (2008). Cowboy and misanthrope: a critical (discourse) analysis of Bush and bin Laden cartoons. *Discourse & Communication, 2*(4), 433-457.

Milburn, S.S., Carney, D. R., & Ramirez, A. H. (2001). Even in modern media, the picture is still the same: A content analysis of clipart images. *Sex Roles, 44,* 277-294.

Moebius, W. (1986). Introduction to picturebooks codes. *Word and Image*, *2*, 141–158.

Nodelman, P. (1988). *Words about pictures: The narrative art of children's picturebooks.* athens: The University of Georgia Press.

Nodelman, P. (1999). Decoding the images. How picture books work. In P. Hunt (Ed.), *Understanding children's literature* (pp. 129–139). London/New York: Routledge.

Norris, S. (2004). Multimodal discourse analysis: A conceptual framework. In P. LeVine, & R. Scollon (Eds.), *Discourse and technology. Multimodal discourse analysis* (pp. 101-115). Washington: Georgetown University Press.

O'Halloran, K. L. (2000). Classroom discourse in mathematics: a multi-semiotic analysis. *Linguistics and Education*, *10*(3), 359–388.

O'Halloran, K. L. (2004). *Multimodal discourse analysis: Systemic-functional perspectives.* London: Continuum.

O'Halloran, K. L. (2005). *Mathematical discourse: Language, symbolism and visual images*. London and New York: Continuum.

O'Halloran, K. L. (2009). Multimodal analysis and digital technology, In A. Baldry & E. Montagna (Eds.), *Interdisciplinary perspectives on multimodality: Theory and practice, proceedings of the third international conference on multimodality* (pp. 1-26). Campobasso: Palladino.

O'Halloran, K. L. (in press 2011). Multimodal discourse analysis. In K. Hyland, & B. Paltridge (Eds.), *Companion to discourse* (pp. 1-33). London and New York: Continuum.

O'Halloran, K. L., Podlasov, A., Chua, A. & K. L. E., M. (2012). Interactive software for multimodal analysis. *Visual Communication*, *11*(3), 363–381.

Ollivier, K. E. (1992). *Gender stereotyping in elementary school textbooks* (Unpublished Master's thesis). University of British Colombia, Vancouver, Canada.

Otlowski, M. (2003). Ethnic diversity and gender bias in EFL textbooks. *Asian EFL Journal 5*(3), 1-15.

O'Toole, M. (2010 [1994]). *The Language of displayed art (2nd Ed)*. London and New York: Routledge.

Pauwels, L. (2012). A multimodal framework for analyzing websites as cultural expressions. *Journal of Computer-Mediated Communication, 17,* 247–265.

Peterson, S., & Kroner, T. (1992). Gender biases in textbooks for introductory psychology and human development. *Psychology of Women Quarterly, 16*, 17-37.

Porreca, K. L. (1984). Sexism in current ESL textbooks. *TESOL Quarterly*, *18*, 704-707.
Retrieved from *http://dx.doi.org/10.2307/3586584*.

Poulou, S. (1997). Sexism in the discourse roles of textbook dialogues. *Language Learning, 15*, 68-73.
Retrieved from *http://dx.doi.org/10.1080/09571739785200141*.

Prodromou, L. (1988). English as cultural action. *ELT Journal*, *42*(2), 73-83.

Reese, L. (1994). Gender equality and texts. *Social Studies Reviews, 33*, 12-15.

Royce, T. D. (1998). Synergy on the page: Exploring intersemiotic complementarity in pagebased multimodal text. In *JASFL*

Occasional Papers 1 (pp. 25–49). Tokyo, Japan: Association of Systemic Functional Linguistics (JASFL).

Royce, T. D. (2002). Multimodality in the TESOL classroom: Exploring visual-verbal synergy. *TESOL Quarterly*, *36*(2), 191-205.

Royce, T. D. (2007). Intersemiotic complementarity: A framework for multimodal discourse analysis. In T. D. Royce, & W. L. Bowcher (Eds.), *New directions in the analysis of multimodal discourse* (pp. 63–110). Mahwah: Lawrence Erlbaum.

Sheldon, J. P. (2004). Gender stereotypes in educational software for young children. *Sex Roles*, *51*, 433-444.

Sunderland, J. (1992). Gender in the EFL classroom. *ELT Journal*, *46*, 81-91.

Tahririan, M. H., & Sadri, E. (2013). Analysis of images in Iranian high school EFL course books. *Iranian Journal of Applied Linguistics, 16*(2), 137-160.

Thibault, P. J. (2000). The multimodal transcription of a television advertisement: Theory and practice. In A. Baldry (Ed.), *Multimodality and multimediality in the distance learning age* (pp. 31-84). Lampo: Campo Basso.

Thompson, K. (2002). A critical discourse analysis of world music as the 'other' in education. *Research Studies in Music Education, 19*(14), 14-21.

Unsworth, L. (2008). Multiliteracies and metalanguage: describing image/text relations as a resource for negotiating multimodal texts. In D. L. D. Corio, M. Knobel, & C. Lankshear (Eds.), *Handbook of research on new literacies* (pp. 377- 405). New Jersey: Erlbaum.

Unsworth, L., & Wheeler, J. (2002). Re-valuing the role of images in reviewing picture books. *Reading: Language and Literacy, 36*(2), 68–74.

Van Dijk, T. A. (2001). Critical discourse analysis. In D. Tannen, D. Schiffrin & H. Hamilton (Eds.), *Handbook of discourse analysis* (pp. 352-371). Oxford: Blackwell.

Van Dijk, T. A. (2006). Discourse and manipulation. *Discourse & Society, 17*(2), 359–383.

Van Leeuwen, T. (2003). A multimodal perspective on composition. In T. Ensink & C. Sauer (Eds.), *Framing and perspectivising in discourse* (pp. 23-62). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Van Leeuwen, T. (2008). *Discourse and practice: New tools for critical discourse analysis.* New York: Oxford University Press.

Ventola, E., Charles, C, & Kaltenbacher, M. (Eds.) (2004). *Perspectives on multimodality*. Amsterdam: John Benjamins.

Young, K. (2009). Applying multimodal analysis to popular websites to develop students' digital literacy. *Proceedings of the 17th International Conference on Computers in Education. Hong Kong: Asia-Pacific Society for Computers in Education.*