

A Further Investigation of Facial Expressions and Body Gestures as metaphors in E-Commerce

DIMITRIOS RIGAS, NIKOLAOS GAZEPIDIS
 Department of Computing
 University of Bradford
 Bradford
 UNITED KINGDOM

Abstract: - This paper describes two sets of experiments that investigate facial expressions and some issues when using body gestures as communication metaphors in interfaces for e-Business applications. The purpose was to obtain an overall feedback for the use of these types of communication metaphors in e-Commerce. The first set of experiments involved three tasks with a short description of three products with each one presented using either textual or an avatar with facial expressions or body gestures. The second set of experiment involved an extended description of three different products but otherwise the methodology remained the same. The sample was 42 users for both sets of experiments. The results of experiments indicated that there is a prima facie case for the use of these metaphors in e-commerce applications and highlighted several issues for further research.

Key-Words: - E-Commerce, Multimodal metaphors, Facial expressions, Body language, Avatars, Usability

1 Introduction

Electronic Commerce (e-Commerce) is a general term that incorporates any process relating to buying, selling, or exchanging information electronically among buyers and sellers, goods or service providers and other third-party companies [1, 2]. User interfaces for e-Commerce (EC) applications typically use text and graphics with the occasional or limited use of multimodal metaphors such as speech, synthesised or environmental sounds, or avatars that make use of facial expressions and body gestures to deliver information. The most widely known EC model to the public is the Business-to-Consumer (B2C) and therefore we focus our research to that. In this paper we discuss some experimental results from our initial experiments using multimedia features for a B2C application in order to measure their role on users' decisions for purchasing online.

Other studies have found that many of these different types of modalities can be particularly useful in other applications. Examples include the use of earcons to communicate program executions [3] and to improve the usability of interfaces [4]. It is therefore believed that the use of different modalities in interfaces for B2C will provide additional benefits to the usability aspects to those interfaces. The paper describes the initial part of an empirical study that investigated facial expressions and body gestures in e-Business applications.

2 Relevant Work

The number of e-commerce websites has been increased in the Web over the last few years. Just few of them though, follow some basic guidelines as far as the Human Computer Interaction (HCI) aspects and multimedia metaphors (e.g. speech, auditory icons, and earcons) are concerned. Consequently, users tend to abandon their online purchases because some major characteristics are missing [5]. The challenge is to provide an efficient and effective electronic commerce application for the end-users that will attract their interest [6].

As e-commerce websites continues to expand the need for interaction and multimedia contents becomes noticeable [7, 8]. The use of speech, earcons and auditory icons enhance the browsing and the user interface improving the capabilities of the application. But what are these icons and sound symbols that are a major factor for EC websites?

2.1 Multimodal Metaphors

Auditory icons are non-speech sounds that simulate physical sounds representing an event or an action that took place [9]. Users are familiar to these sounds and they know what they represent since they derive from the natural environment.

Earcons are non-speech sounds that use abstract or synthetic sounds derived from instruments or synthesizers. Earcons (as well as other auditory

stimuli) can attract the attention of the user or announce to users an action or a result of an action [10]. They do not appear visually on the screen and they can vary on their level of sound intensity [11].

Speech metaphor is often used in multimedia user interfaces so as to provide users with feedback along with the graphical environment about system's current state [12] and it is a very useful tool especially for visually impaired users [13]. We distinguish two types of speech; natural and synthesised.

Computer-based systems offer a speech technology known as *Text-to-Speech (TTS)* synthesis technology [14]. TTS systems have the ability to read any arbitrary text, they analyze it and after converting it, they output it as a synthesised spoken message comprehensible by the user [15, 16].

Facial expressions give a more realistic interaction in human computer interfaces. The face is a mean of expressing emotions, feelings, and linguistic information and due to the improvement of computer hardware (high performance graphics and speed) instances of cartoon-like and human-like synthesized faces are under development and researched in depth for use in computer applications [17].

The movement of the body, of hands and head, play a major role in everyday communication. These are often used either to give emphasis to our speech, to point at an object or to illustrate the size or shape of it. *Gestures* are also known as "body language" communication [18].

2.2 Lack of Multimodality

In commercial web sites, there is a lack of a face to face communication between the customer and the seller. In addition, - referring to a website that is selling electronics for example - products are not tangible. Hence, it is necessary to experiment and evaluate some multimedia aspects that can be embedded in the designing of an EC site. We will measure the effect of audio-visual stimuli and multimodal interactions by introducing other modes of communication (see multimodal metaphors) between the user and the EC application. These metaphors will be used either on their own or simultaneously, or there will be some combinations so as to have precise results [19].

It has been speculated, that human-like interface agents with a combination of facial expressions like moving eyebrows, head movement, smiling, quiver of the eyelids, opening and closing eyes, or lips movement synchronised with a text-to-speech generation system [20, 21, 22] will have a positive

effect in consumers' decisions [23] as they simulate a real-life character interacting in a realistic way with them. Several projects have been developed such as the BEAT [20], the COGITO [23], SoNG [24].

2.3 Real Interaction

Multimedia technology can play an important role for the development of consumer to computer transactions, giving the advantages and the requirements needed for an easy navigation, providing various functionalities, and friendly user interface [6]. In spite of the fast growth of EC, consumers still abstain from online purchases. The metaphor of facial-to-facial approach is the key to attract mistrustful consumers [25]. The role of the semi-autonomous animation characters - in extend Intelligent Agents - along with the combination of TTS technology providing high-quality natural interaction will have a great effect not only for able users but also for disabled [21]. Natural face-to-face communication between the avatars and the users, involves as mentioned above multiple modes of facial expressions.

These characters will help to simplify and automate a purchase process as well by saving time, effort and money comparing to traditional shopping [26, 27]. They carry out most of the activities on behalf of the user and they are very useful for process purchasing tasks, interacting with the users [28] reflecting some very humans-like conversations, e.g. responding to verbal and non-verbal input from the user, generating a meaningful output and in general giving feedback along with face and body gestures (provided that the avatar has a full-body representation) [29, 30]. Maes [31] mentions the issue of trust related to these anthropomorphic avatars. Trust is an important element in EB transactions, for short term and long term customers [32]. In what extend are they trustworthy regarding to the information they provide and how they can guarantee they will persuade the users since they are the intermediate between the user and the online market. Accurate choices of human-like facial expressions and movements can lead to satisfactory outcomes by improving the interface and the capabilities of an EC site. However, they have not yet been widely in use.

Multimodal interfaces simplify and improve peoples' life. In electronic business, this technology will make the difference among the competitors [33]. As we mentioned above, auditory icons, earcons and other technologies such as text-to-speech (TTS), facial expressions and VoiceXML,

would frame an EC solution. In this research paper we will try to take advantage of the existing technologies and we will make a step further to experiment some new issues in B2C market [34]. Voice, data and facial expressions would be used for the development of a cohesive EC platform, and we suspect to be an advanced interactive tool between the consumer and the business.

3 Research Programme

Our aim in this research is to investigate the usability aspects of B2C interfaces that utilise speech, non-speech, avatars with facial expressions and gestures in addition to the typical visual graphical metaphors that are currently used.

User interfaces for B2C applications typically use visual metaphors with the occasional or limited use of multimedia features. More specifically, some of the research questions include:

1. Does the use of natural or synthesised speech have an effect upon the usability of an interface? If so, which of the two increases the usability of the interface?
2. Does the use of auditory icons and earcons have an effect upon the interface?
3. Does the use of avatars with human like facial expressions and different types of voice tones have an effect upon the usability of an interface?
4. Does the use of avatars with human-like gestures have an effect upon an interface?
5. Does multimodal overall provide less or more efficient options for a candidate buyer through the WEB upon his final decision?



Figure 1: The experimental platform used for the experiments. **A** front view, **B** side view, **C** texture preview, **D** face sculpting.

Taking into account the experiments took place in the past on different metaphors but basically based

on facial expressions and gestures we made the following hypothesis:

“In comparison with a textual-based E-Commerce website, multimedia E-Commerce website would be more effective and desirable for the users’ decisions in terms of accomplishing electronic purchases over the WEB”.

Fig. 1 shows the face model and its various stages of development that were used as a basis for the animated face and body. A front view and side view required in digital pictures format so as to create a texture preview that will be attached on the animation model and initiate the face sculpting.

Fig. 2 shows the application presented to the user into three different ways of presentation (textual, facial expressions and body gestures animations).

4 Experiment

The results of our previous experiment [35] showed us that we had to do an additional experiment to further explore our hypothesis. Again the experiment evaluated three different interface metaphors for a B2C application. These metaphors involved the use of speech, facial expressions and body gestures for the presentation of three products using three different ways (textual, face, and full body animation). The experiment measured aspects of usability, interactivity and likeliness of the animating application. The experiment was conducted with 42 users.

The B2C application developed involved three different products from the first set of experiment. The brand names of the products were removed in order to avoid brand name influencing user choice. Each product had a textual description which was either presented visually for the textual version or spoken for the face and full body animated versions. In the spoken version parts of the textual description were annotated with specific vocal tones, facial expressions and gestures.

4.1 Results

The *textual* presentations of the products were described by more than 50% of the users as poor (Fig. 3). However, only 33% of the users described as good and another 12% described it as very good. During the post experimental interviews, users commented upon the interactivity, attraction, eye-catching, interesting and impressive aspects of the multimodal approach. Some users also commented that it was easier to interact with a multimodal application compared to the one that is dominated

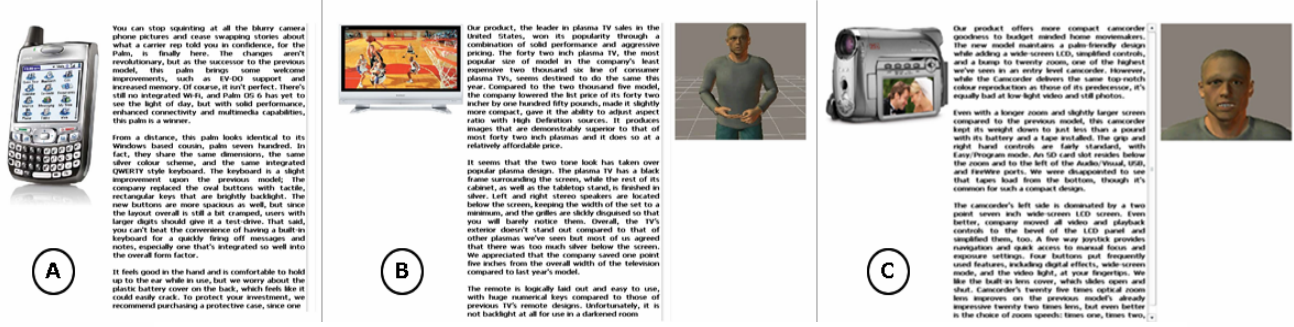


Figure 2: Textual vs. Multimodal approach presentations. A Text, B Body, C Face.

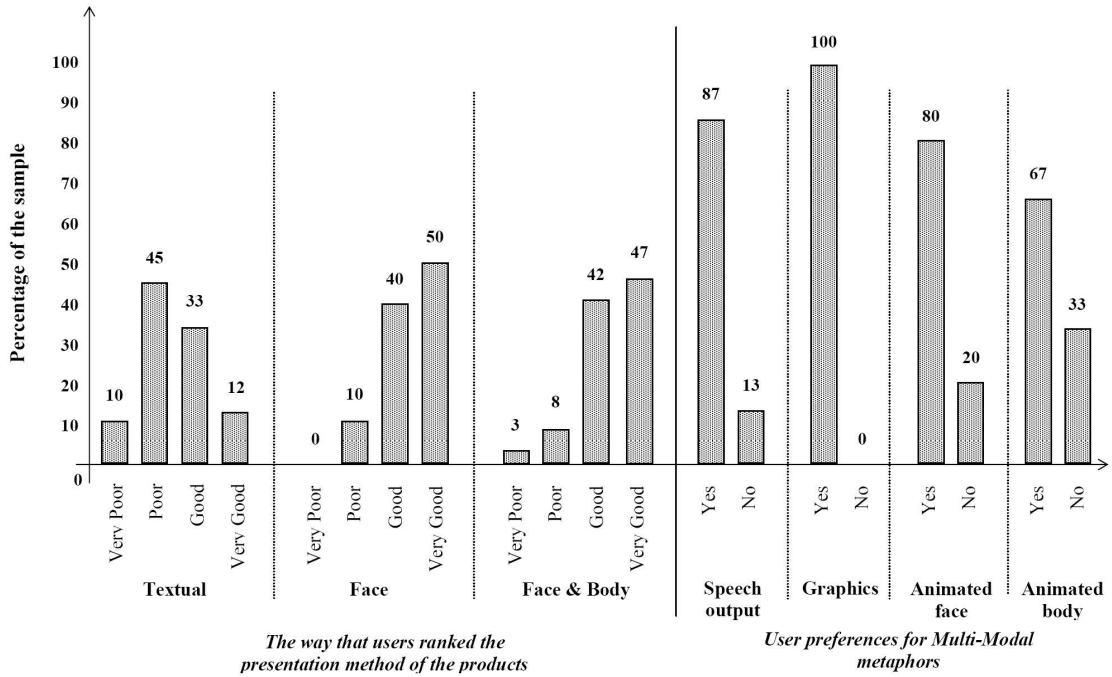


Figure 3: First set of experiment - Results in percentages of the way users selected the presentation according to product and regardless of it.

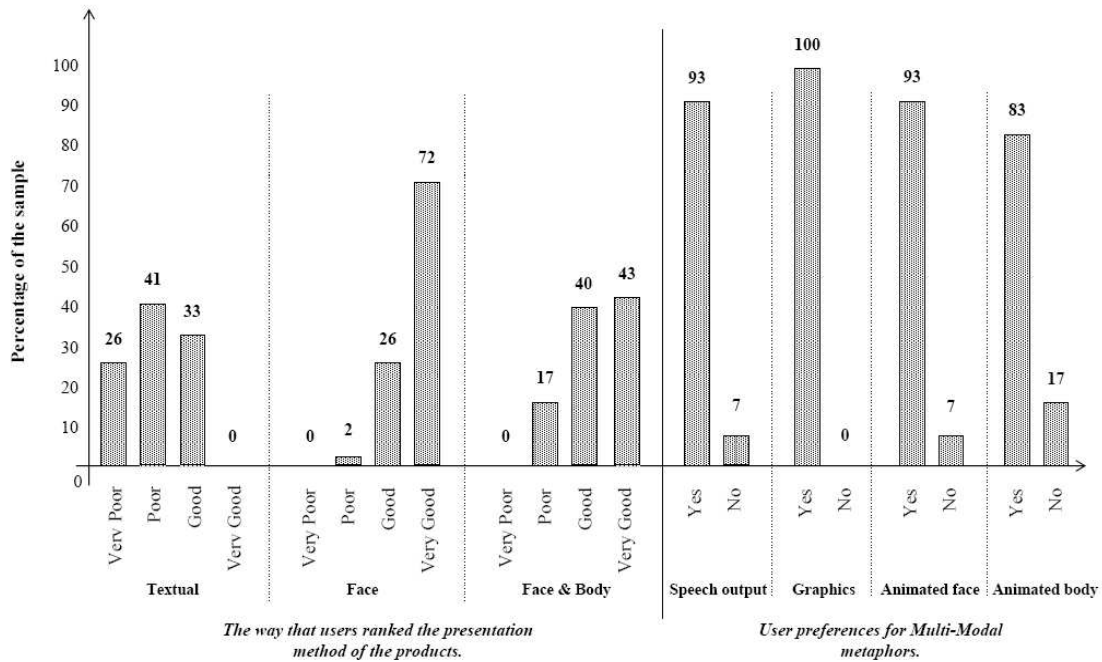


Figure 4: Second set of experiment - Results in percentages of the way users selected the presentation according to product and regardless of it.

with text and graphics only. The presence of the avatar gave a positive impression to users. Therefore the face-to-face approach was perceived by users that went some way to simulate a real face-to-face shopping experience.

As far as the multimodal metaphors are concerned, users stated that *speech output* was considered very important and it was selected by 87% of them. All 42 users regarded graphics as essential and that played an important role in the interaction of the e-Business transactions. *Animated face* was preferred by 80% of the users and *animated body* was preferred by 67% of the users.

Fig. 4 shows users' preference for each presentation method used in this phase of the experiment and their general impression for multimodal metaphors. More than 65% of the users participated in the experiment described the *textual* presentations of the products as negative method of presentation and only a 33% of them had a positive view. On the other hand the presentation of the products with the *animated face* was ranked positive by the majority of users. The total of good and very good opinion level reached the ultimate acceptance with 98% of satisfaction. It worth to mention that no user had a very poor view for this method of presentation and that only 2% described it as poor. Lastly, the *face & body* presentation also received positive judgements. The acceptance level was above 80% from the users and about 15% of them judged it as poor.

Users were also questioned about the multimedia metaphors used in their presentation; if they could play an important role when they do online shopping. *Speech output* is considered to be very important from the majority of users with a 93%. Furthermore, *graphics* are the most essential metaphor for a website and play an important role for all the 42 users of our experiment. Moreover, *animated face* was viewed positively from more than 90% of the users and it could help them when shopping online whereas *animated body* was preferred from 83% of the users.

5 Discussion & Concluding Remarks

Presentations of facial expressions followed by the body gestures were preferred by most of the users. These metaphors could help them when shopping online along with the use of speech and graphics. The results from the two sets of experiments have supported our initial hypothesis and showed us the need for further experiments, analyzing the usability

aspects of various facial expressions and body gestures on their own or combined in an interface.

The anticipated benefits of continuing this research will be to mainly produce a series of guidelines for the use of facial expressions and body gestures in B2C applications. Human-like avatars that incorporate the above metaphors give the potential interactivity between the user and the EC interface. A number of further experiments are currently under progress to investigate further the identified usability aspects of the different modalities.

References:

- [1] Barnes – Vieyra, P., Claycomb, C., *Business-to-Business E-Commerce: Models and Managerial Decisions*. Business Horizons / May-June, 2001.
- [2] Turban, E., King, D., 2003. *Introduction to E-Commerce*, Prentice Hall. Upper Saddle River, New Jersey, pp.3
- [3] Rigas, D., Alty, J., *Using Sound to Communicate Program Execution*. Euromicro, p. 20625, 24th.EUROMICRO Conference Vol. 2, 1998.
- [4] Brewster, S., *Using Earcons to Improve the Usability of a Graphics Package*. Glasgow Interactive Systems Group, Department of Computing Science, University of Glasgow, Glasgow, G12, 8QQ, UK, 1998.
- [5] P. Markellou, *Designing a Usable Shopping Cart*, Department of Computer Engineering and Informatics, University of Patras, Greece, 2001.
- [6] Fabio Nemetz, *The Missing Link: Multimedia and E-Commerce*, Department of Mathematical Sciences, University of Bath, 2000.
- [7] László Böszörményi, Hermann Hellwagner, Harald Kosch, *Multimedia Technologies for E-Business Systems and Processes*, Institut für Informationstechnologie, Universität Klagenfurt
- [8] Mehrdad Jalali-Sohi, Feza Baskaya, *A Multimodal Shopping Assistant for Home E-Commerce*, Fraunhofer Institute for Computer Graphics, 2001.
- [9] Jona J. Bjur, *Auditory Icons in an Information Space*, Department of Industrial Design, School of Design and Craft, Göteborg University, Sweden, 1998.
- [10] Stephen Brewster, Veli-Pekka Raty, Atte Kortekangas, *Earcons as a Method of Providing Navigational Cues in a Menu Hierarchy*, 1996.

- [11] Stephen A. Brewster, *Using Earcons to Improve the Usability of Tool Pallettes*, Glasgow Interactive Systems Group, Department of Computer Science, 1998.
- [12] Jenny Preece, Yvonne Rogers, Helen Sharp, David Benyon, Simon Holland, Tom Carrey, *Human-Computer Interaction*, Addison-Wesley, 1994.
- [13] Lorna Lines, Kate S. Home, *Older Adults' Evaluations of Speech Output*, Department of Information Systems & Computing, Brunel University, Uxbridge, Middlesex, U.K. 2002.
- [14] Thierry Dutoit, *A Short Introduction to Text-to-Speech Synthesis*, TTS research team, TCTS Lab., 1999.
- [15] J. Schroeter, J. Ostermann, H. P. Graf, M. Beutnagel, E. Cosatto, A. Syrdal, A. Conkie, Y. Stylianou, *Multimodal Speech Synthesis*, AT&T Labs – Research, New Jersey, USA, 2000.
- [16] Johan Wouters, Brian Rundle, Micahel W. Macon, *Authoring Tools for Speech Synthesis Using the Sable Markup Standard*.
- [17] Jonas Beskow, *Talking heads—communication, articulation and animation*, Department of Speech, Music and Hearing, KTH, 1996.
- [18] David Benyon, Phil Turner, Susan Turner, *Designing Interactive Systems – People, Activities, Contexts, Technologies*, Addison-Wesley 2005.
- [19] Jörn Ostermann, David Millen, *Talking Heads and Synthetic Speech: An Architecture for Supporting Electronic Commerce*, AT&T Labs–Research, New Jersey, USA, 2000.
- [20] Justine Cassell, Hannes Högni Vilhjálmsson, Timothy Bickmore, *BEAT: the Behavior Expression Animation Toolkit*, MIT Media Laboratory, 2001.
- [21] A. Nijholt, D. Heylen, R. Vertegaal, *Inhabited interfaces: attentive conversational agents that help*, Deptment of Computer Science, University of Twente, The Netherlands, 2000.
- [22] Brigitte Krenn, Barbara Neumayr, Erich Gstrein, Martine Grice, *Life-like Agents for the Internet: A Cross-Cultural Case Study*.
- [23] A. Paradiso, M. L'Abbate, *A Model for the Generation and Combination of Emotional Expressions*, GMD – German National Research Center for Information Technology Institute (IPSI), Darmstadt, Germany, 2005.
- [24] F. Guerin, K. Kamyab, Y. Arafa, E. Mamdani, *Conversational Sales Assistants*, Intelligent and Interactive Systems, Department of EEE, Imperial College of Science, 2000.
- [25] Carey E. Heckman, Jacob O. Wobbrock, *Put Your Best Face Forward: Anthropomorphic Agents*, E-Commerce Consumers, and the Law, 2000.
- [26] Aleksander Pivk, Matjaž Gams, *E-commerce Intelligent Agents*, Jozef Stefan Institute, Department of Intelligent Systems, Ljubljana, Slovenia, 2000.
- [27] Pattie Maes, Robert H. Guttman, Alexandros G. Moukas, *Agents that Buy and Sell: Transforming Commerce as we Know It*, Software Agents Group, MIT Media Laboratory, 1999.
- [28] Robert H. Guttman, Alexandros G. Moukas, Pattie Maes, *Agents as Mediators in Electronic Commerce*, Software Agents Group, MIT Media Laboratory, 1998.
- [29] Yasmine Arafa, Abe Mamdani, *Electronic Personal Sales Assistants: An Interface to e-Commerce Applications*, Intelligent and Interactive Systems, Department of EEE, Imperial College of Science.
- [30] Justine Cassell, *Embodied Conversational Interface Agents*, Communications of the ACM, 2000.
- [31] Pattie Maes, *Agents that Reduce Work and Information Overload*, MIT Media Laboratory, 1994.
- [32] Florian N. Egger, *Evaluating the Customer Trust Experience in B2C E- Commerce Environments*, Geneva, Switzerland, 2003.
- [33] Bob Woods, "What is Multimodality", 2002, <http://www.instantmessagingplanet.com/wireless/article.php/976511>.
- [34] José Rouillard, *A multimodal E-commerce application coupling HTML and VoiceXML: E-commerce – User Interface and Interactions – Web applications*, Laboratoire TRIGONE, CUEEP, 2002.
- [35] Dimitrios Rigas, Nikolaos Gazepidis, *Facial Expressions and Body Gestures as Metaphors for B2C Interfaces: An Empirical Study*, University of Bradford, Bradford, U.K., 2006.