Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Int. J. Human-Computer Studies 62 (2005) 211–229
International Journal of Human-Computer Studies
www.elsevier.com/locate/ijhcs

Variations in gesturing and speech by GESTYLE

Han Noot[a,*], Zsófia Ruttkay[b]

[a] *Center for Mathematics and Computer Science, INS, Kruislaan 413, 1090 GB Amsterdam, The Netherlands*
[b] *University of Twente, EWi-HMI, P.O. Box 217, 7500 AE Enschede, The Netherlands*

Abstract

Humans tend to attribute human qualities to computers. It is expected that people, when using their natural communicational skills, can perform cognitive tasks with computers in a more enjoyable and effective way. For these reasons, human-like embodied conversational agents (ECAs) as components of user interfaces have received a lot of attention. It has been shown that the style of the agent's look and behaviour strongly influences the user's attitude. In this paper we discuss our GESTYLE language making it possible to endow ECAs with style. Style is defined in terms of when and how the ECA uses certain gestures, and how it modulates its speech (e.g. to indicate emphasis or sadness). There are also GESTYLE tags to annotate text, which has to be uttered by an ECA to prescribe the usage of hand, head and facial gestures accompanying the speech in order to augment the communication. The annotation ranges from direct, low level (e.g. perform a specific gesture) to indirect, high level (e.g. take turn in a conversation) instructions, which will be interpreted with respect to the style defined. Using style dictionaries and defining different aspects like age and culture of an ECA, it is possible to tune the behaviour of an ECA to suit a given user or target group the best.

Keywords: Embodied conversational agent; Multimodal communication; Style; Mark up language

*Corresponding author. Tel.: +31 20 5924330; fax: +31 20 5924199.
E-mail addresses: han@cwi.nl (H. Noot), z.m.ruttkay@ewi.utwente.nl (Z. Ruttkay).

1071-5819/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ijhcs.2004.11.007

1. Introduction

1.1. About ECAs

Empirical studies suggest that users respond to complex interactive devices as they respond to humans (Reeves and Nass, 1996). This has given rise to the CASA (Computers Are Social Actors) paradigm (Nass et al., 1994). A striking example of such behaviour is reported in Nass (2004). In the experiments, computer software was used for explanatory purposes. After a session with the system, the users were asked to evaluate the software by answering a series of questions. It mattered whether both tasks did run on the same or on different computers. Users rated the service better if they had to do the evaluation on the same computer, which had been used to help them. This parallels the tendency that in every-day life, because of politeness, we express less criticism about a service directly to the person who assisted, than to somebody else. So users expressed politeness towards the computer.

On the other hand, for people the natural way of communicating is speech, accompanied by subtle gestures, facial expressions and postures. These two observations gave rise to human-like characters, so called embodied conversational agents (ECAs) in man–machine communication. It is expected that people, when using their natural communicational skills, can perform cognitive tasks with computers in a more enjoyable and effective way.

An ECA is some creature which resides on the computer screen, which resembles a living creature in look and behaviour, and assists the user in the task at hand (Cassell et al., 2000). Most often human-like characters are used, but agents with embodiments as animals (Isbister et al., 2000) or even animated objects (Microsoft's paperclip) do occur. When utilizing ECAs, many design questions and evaluation issues need to be taken care of (Massaro et al., 2002; Ruttkay et al., 2004, to appear). We mention only a few: How should the ECA look like: 2D or 3D, realistic or cartoon like, what gender and culture does it have, should it posses a complete body or only have a (talking) face? How should it be dressed? What should be its communicative abilities? Does it indicate turn giving/taking, does it show idling behaviour (blinking, drumming its fingers), does it display emotions? What nonverbal signals are used to indicate these states? What are the motion characteristics of the gestures? Is the ECAs nonverbal behaviour fully repetitive, or are some variances possible? Can it adapt to the (static or changing) characteristics of a specific user?

The believability of ECAs highly depends on their nonverbal communicational skills: the richness of the used modalities and gestures, and the correctness and consistency of choosing and performing a gesture. Different persons, depending on their cultural, social and professional background and their personality, use different gestures or exploit different modalities in the same situations while communicating (McNeill, 1991; Kendon, 1993). Also, there is evidence that the user's response to the ECA depends on subtle characteristics like ethnicity and personality of the ECA (Walker et al., 1994; Nass et al., 2002). In general, it seems that the ECA should resemble the user in order to be appreciated most. For instance, the virtual real estate