

Feminity, Masculinity and Androgyny: How Humans Perceive the Gender of Anthropomorphic Agents

Andreea Niculescu, Frans van der Sluis, and Anton Nijholt

Human Media Interaction Group, University of Twente,
P.O. Box 7500, Enschede, The Netherlands
{niculescuai, f.vandersluis, a.nijholt}@utwente.nl

Abstract. In this paper we present the first experimental step of a larger study investigating the effects of agents' gender identification on the perceived interaction quality. The purpose was to determine how test subjects judge the 'gender' of virtual humans. 48 test subjects rated 7 agents according to their degree of feminity, masculinity and gender ambiguity. The results showed that subjects' gender combined with the order in which the agents were presented have a strong impact on their perceived degree of feminity, masculinity and gender ambiguity. The agents identified as having strong feminine, masculine and androgynous¹ traits will be used in the next experimental session to represent a multimodal dialogue system for medical queries.

Keywords: Anthropomorphic agents, physical appearance, gender perception.

1 Introduction

Physical characteristics, such as age, gender and ethnicity are important cues in human social perception, cognition and behavior [1]. They represent a basic 'business card' that tells people how to approach a potential conversational partner.

Research studies showed that humans prefer to engage in conversation with those whose physical appearance can be labeled consistently [2]. The reason is the human tendency to simplify the interlocutor's representation by framing her into pre-defined categories (e.g. young, female, Asian) [3]. This framing lightens the cognitive load and gives a secure feeling of dealing with predictable situations [4].

Among all salient visual cues that coins physical appearance the gender seems to be of fundamental importance, being one of the first visual information people exchange. The explanation goes beyond the cognitive load lightening and relates to our evolutionary history where gender related information assured the correct orientation toward a potential mating candidate. Since the decoding of such information has powerful impact in social interactions we believe that its lack would be perceived as unpleasant. In other words we assume humans would prefer to interact with those whose gender they can label consistently and they would maintain this preference even when they interact with artificial entities such as embodied conversational agents (ECAs).

¹ We use the term 'androgynous' to refer to a person whose look makes it difficult to determine her gender.

2 ‘Virtual’ Gender Problematic in HCI

The virtual gender problematic became a highly interesting topic in the HCI field since many computer media systems started to use human avatars to represent a certain application. Previous research demonstrated that humans treat computers as they were social actors even if they don’t exhibit anthropomorphic traits [5]. By adding face and embodiment the social relationship between user and computer becomes even more explicit: clothing, hair style, facial expressions, age, gender, etc displayed by agents bring the rich and complex world of social interactions into the interface [6].

A number of researchers have studied the effects of the ‘virtual’ gender on the way people perceive conversational agents and build relationships with them. For example Zimmerman et al. [7] concluded that people’s preferences correlate with agents’ conformity to gender stereotypes referring to specific roles: female agents were preferred for tasks traditionally undertaken by women (librarian, matchmaker), male agents for task undertaken by men (fitness trainer). They also found that men prefer embodied agents more than women do and that both male and female users prefer female over male agents.

Another study by Catrambone et al. [8] suggested that male and female users might have a different way to personify agents: 54% of the female subjects used a personal pronoun (he/she) to refer to an agent, while only 13% of the male users did the same.

De Angelli and Brahnan [3], who did research on chatterbots found that ‘virtual’ genders impact the incidence of sex talk: agents clearly signaling their gender (female, male) were more prone to be verbally abused by human users than those that didn’t do so (robot).

However, except for [3] no other studies focusing on the effect of gender ambiguity in HCI could be found. This is surprising, since there is a common practice for computer applications to display agents whose physical appearance doesn’t point to any particular gender - the designer’s intention behind is that both male and female users could relate to the agent.

Therefore we propose a research study to investigate the impact of agents’ ambiguous look on the perceived conversational interaction quality. In this paper we present the results of the first experimental step.

3 Experiment Design and Methods

In the experiments we plan users are interacting with three talking heads of a multi-modal information system for medical queries (the system was partly developed in our department²); users can interact with it using speech, text type or mouse pointing. Each head has a particular gender marked appearance, i.e. masculine, feminine and androgynous. The androgynous agent, called Ruth³ could be classified either as a young male or a female with slight masculine traits.

² http://www.nwo.nl/nwohome.nsf/pages/NWOP_652GXD

³ The head was developed at Rutgers University, New Jersey (USA), see: <http://www.cs.rutgers.edu/~village/ruth/>; interestingly the name ‘Ruth’ is Dutch ambivalent: it designates a female person, if spelt ‘th’ (Ruth) or a male, if spelt without (‘Rut’).



Fig. 1. A 'portrait' of Ruth

The masculine and feminine heads were developed based on the androgynous version by adding masculine, respectively feminine characteristics. The modifications were kept minimal in order to make the comparison between the heads sustainable, i.e. factors such as beauty or face symmetry should not influence the agent's perception. The female versions are wearing earrings, have narrow eyebrows, a lighter skin color and a feminine hair style; each head has a different hair color.

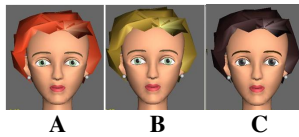


Fig. 2. Female versions of Ruth

The masculine heads have a much darker skin color, moustache and/or beard and short hair; one of the heads is blond, the other two are brunettes.

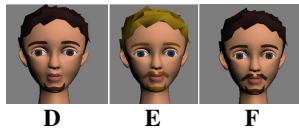


Fig. 3. Male versions of Ruth

Pitch and frequencies of the TTS voices are adapted to fit the agents' appearances - for the androgynous agent a female voice with low pitches is chosen. Wizard-of-Oz simulation is going to be used to ensure a homogeneous interaction.

Test subjects are informed they are interacting with three different systems. In fact they are interacting with only one system, whose speech recognition module is replaced by the Wizard. Subjects are asked to assess the interaction quality and the agent's appearance. Subjects' degree of femininity, masculinity and androgyny are measured in advance using the BEM test. [9]. The aim is to determine whether gender related personality traits correlate with the preference for a certain virtual gender type.

In the first experimental step the agents' images (Fig. 1,2,3) were shown to 48 test subjects (24 males and 24 females), aged between 16 and 73 years, originating from 10 different countries across Europe, Latin America, Asia and Middle East. They had to rate the agents' gender belonging as well as their degree of femininity/masculinity on a 5 point scale. No direct question addressed the gender ambiguity in order to avoid priming effects.

For the feminine and masculine heads letter identifiers (A, B, C, etc) were used in order to exclude any semantic association test subjects might unconsciously made between name and look.

To remove any potential bias from being exposed to one image category (gender ambiguity) before the other (gender explicitness) test subjects were organized into two equal groups: one half had to rate the gender marked images first and afterwards the image representing Ruth (test order 1: TO1); the other group did the opposite (test order 2: TO2).

The purpose of this experimental step was to determine: 1) whether Ruth is indeed perceived as gender ambiguous and 2) which of the gender marked agents best represent the femininity and masculinity concept as opposed to gender ambiguity.

4 Results

The results confirmed Ruth's gender ambiguity: even there was a slight trend to consider the agent a 'male' the answers of both test groups were given randomly ($\chi^2(1, N=48)=3.00, p >.05$). However the trend was not particularly strong: measuring the degree of masculinity Ruth was considered 'less masculine' ($M=3.43$) than 'so-so' ($M=3.00$), $t(46)=3.42, p <.001$. For the femininity degree there was no statistically significant difference between the middle value 'so-so' ($M=3.00$) and the mean value of the answers chosen by the test subjects ($M=3.07$), $t(44)=.50, p >.05$. This proves again there was no trend toward a particular gender direction.

No statistically significant difference could be found on how male and female test subjects decided on Ruth's gender belonging. However when the test order was taken into account the subjects' gender became statistically significant: females found Ruth more feminine in TO1 ($M=3.00$) compared TO2 ($M=3.83$) and more masculine in TO2 ($M=2.86$) compared to TO1 ($M=3.60$); males found Ruth more feminine in TO2 ($M=2.69$) compared to TO1 ($M=3.33$), and more masculine TO1 ($M=3.00$) compared to TO2 ($M=3.75$). Thus, males and female subjects reacted oppositely when confronted with a different test order. The interaction effect between order and gender was confirmed to be statistically significant by a MANOVA test (Ruth's perception of femininity, $F(1,41)=7.71, p <.01$ and masculinity, $F(1,43)=9.11, p <.01$).

Regarding the most feminine agent a slight trend towards the blond (B) and brunette (C) agents could be observed. However, the trend was statistically not significant. No negative values ('less' or 'no' feminine) were found. Concerning the most masculine agent a significant difference ($\chi^2(2, N=48)=41.63, p <.001$) was found in the favor of the brunette agent displaying beard and moustache (F).

5 Conclusions and Further Work

In this paper we investigated how humans judge agents concerning their femininity, masculinity and gender ambiguity. Our results showed that female and male test subjects had opposite perceptions on Ruth's gender depending on the order they had to rate it: before seeing the gender marked agents or afterwards. However test order and gender, as influence parameters taken separately, were statistically not significant;

only when combined together in a single significance test they were statistically meaningful concerning Ruth's gender perception. The strong interaction effect between test order and subjects' gender supports the idea that female and male have quite different perceptions on masculine, respectively feminine traits.

The results of the study also indicated the femininity and masculinity degree of the gender marked agents. This outcome is especially useful for our next research step when we intend to use particular strong gender marked agents as opposed to gender ambiguous in a more dynamic context (those of spoken conversational interactions). Additionally we plan to perform experiments with female/male agents having an elderly look in order to determine whether the age has an impact on the perceived interaction quality.

Acknowledgments

The authors are grateful to all subjects for participating in the test, to Herwin van Welbergen for proving useful hints regarding the 3D modelling tool and to Betsy van Dijk for useful comments regarding the statistical data analysis.

References

1. Hogg, M.A. and Vaughan, G.M. *Social psychology*. Pearson Prentice Hall, Harlow, England (2005)
2. Nass, C., Isbister, K., Lee, E., J.: *Truth is Beauty: Researching embodied conversational agents*. In: J. Cassell, J. Sullivan, S. Prevost, E. Churchill (eds.): *Embodied conversational agents*, MIT Press, Cambridge, MA, USA (2000)
3. De Angeli, A., Brahnam, S.: *Sex stereotypes and conversational agents*. In: *Proc. of Gender and Interaction: real and virtual women in a male world*, Venice, Italy (2006).
4. Fiske, S. T., S.E. Taylor. *Social Cognition*. New York (1991)
5. Reeves, B., Nass, C.: *The media equation: how people treat computers, television, and new media like real people and places*. Cambridge University Press/CSLI, New York (1997)
6. Forlizzi, J., Zimmerman, J., Mancuso, V., Kwak, S.: *How interface agents affect interactions between humans and computers*. In: *Proc. of Designing Pleasurable Products and Interfaces*. ACM Press (2007) 209–221
7. Zimmermann, J., Ayoob, E., Forlizzi, J., McQuaid, M.: *Putting a face on embodied interface agents*. In: *Proc. of Designing Pleasurable Products and Interfaces*. Eindhoven Technical University Press, (2005) 233–248
8. Catrambone, R., Stasko, J., Xiao, J.: *Anthropomorphic agents as a UI paradigm: experimental findings and a framework for research*. Technical Report GIT-GVU-02-10. Georgia Institute of Technology (2002)
9. Bem, S. L.: *The BSRI and gender schema theory: A reply to Spence and Helmreich* *Psychological Review*, 88, 369–71, (1981)