

Evaluation of Individual Multimodal Behavior of 2D Embodied Agents in Presentation Tasks

BUISINE Stéphanie¹ ABRILIAN Sarkis¹ MARTIN Jean-Claude^{1 & 2}

(1) LIMSI-CNRS, BP 133, 91403 Orsay Cedex, France +33.1.69.85.81.04

(2) LINC-Univ. Paris 8, IUT de Montreuil, 140 Rue de la Nouvelle France, 93100 Montreuil, France

<mailto:{buisine, sarkis, martin}@limsi.fr>

<http://www.limsi.fr/Individu/martin/>

ABSTRACT

Individuality of Embodied Conversational Agents (ECAs) may depend on both the look of the agent and the way it combines different modalities such as speech and gesture. In this paper, we describe a study in which male and female users had to listen to three short technical presentations made by ECAs. Three individual multimodal strategies of ECAs were compared: redundancy between speech and arm gestures, complementarity, and speech-specialization – in this last case, arm gestures were non-semantic. These strategies were randomly attributed to three different-looking agents, in order to test independently the effects of multimodal strategy and agent's appearance. The variables we examined were subjective impressions and recall performance. Although multimodal strategies were hardly perceived explicitly, they proved to influence subjective ratings of quality of explanation, in particular for male users. On the other hand, appearance affected likeability, but also recall performance. These results stress the importance of both multimodal strategy and agent's appearance to ensure pleasantness and effectiveness of presentation ECAs. Such a study might be useful for future recommendations regarding both the design and the evaluation of ECAs individuality.

Categories and Subject Descriptors

H.5.2-H.5.1 [Information Interfaces and Presentation]: User Interface – *interaction styles, standardization, ergonomics, user interface management systems*. Multimedia Information Systems.

General Terms

Algorithms, Human Factors, Languages.

Keywords

Embodied Conversational Agent, Evaluation, redundancy, complementarity, specialization.

1. INTRODUCTION

In order to make Embodied Conversational Agents (ECAs) more believable [13] and more comfortable [4], attempts are made to give them some aspects of emotions and personality during the interaction with human users (see [4] for a review; [1]). Personality contributes to a large extent to defining ECAs as individuals: extraversion, agreeableness or friendliness are some personality traits that have been most studied. They affect all verbal and nonverbal modalities of communication: content of speech, intonation, facial expression, body posture, arm movements, etc.

Personality can be given to ECAs whatever their function. In assistance tasks, [3] use ECAs whose behavior combines presentation acts, which are not based on individual characteristics, and specific behaviors depending on their personality (on the dimensions of extraversion and agreeableness). However, to increase again agents' believability, presentation acts themselves could have been associated to individual strategies. In human behavior, speech-accompanying arm movements can be considered as an integral part of individual communicative style [8] and their occurrences could depend on the tactic of expression temporarily preferred by the speaking person ([10], cited by [15]).

During presentation tasks, ECAs have to relate speech and pictorial information. In such a context, cooperation between modalities observed in humans could be used to specify ECAs' behavior. [9] have identified redundancy, complementarity and specialization – among others – as being different types of cooperation between modalities in human behavior. In a presentation context, redundancy would consist in giving verbal information and repeating it either with an iconic gesture or a deictic gesture toward an object. Although not explicitly named, this kind of strategy seems to be most frequently adopted for animated presenters or pedagogical agents [3, 14]. Conversely, cooperation by complementarity would aim at reducing the amount of information given by each modality. For example, the agent could talk about an object and give information (e.g. shape or size) by hand gesture without mentioning this information by speech.

Some other presentation agents could be designed to give verbally the whole content of the presentation. In this case, arm gestures would be non-semantic (e.g. beat gestures) and modalities would cooperate by speech-specialization. This type of cooperation corresponds to the "elaborate speech-style", which is likely to

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Conference '00, Month 1-2, 2000, City, State.

Copyright 2000 ACM 1-58113-000-0/00/0000...\$5.00.

occur when the discourse content is distant from personal experience, conventional, abstract, and objective [15].

The primary goal of this study was to determine whether individual multimodal strategies, when exhibited by ECAs, would be perceived by a human listener and/or would have an impact on the effectiveness of presentation. In these cases, what strategy would be the best one? We decided to test the effect of three multimodal strategies – cooperation by redundancy, complementarity and speech-specialization – in ECAs short presentations. We have selected these three strategies as they are rather different from one another and thus one could expect significant results when comparing them (although we did not make any preliminary hypothesis about which one would be perceived best). The three selected strategies were randomly attributed to three different-looking agents: we were thus able to test independently the effects of multimodal strategy and appearance of the agent on subjective impressions (in a post-experimental questionnaire) and recall performance of the listeners. We think that the coherent design of an ECA's appearance and multimodal strategies is of importance to the user. Thus, we believe that both should be considered in evaluation studies. By means of this experimental design, our goal was to identify optimal combinations of agent's appearance and multimodal behavior as a function of the characteristics of application and users. Finally, we included in the post-experimental questionnaire items about agents' personality, in order to test whether multimodal strategy and/or appearance influenced user's perception of agent's personality. In order to fully control the parameters of the agents' behavior, the users could not interact with them. Thus, the users' task consisted in listening to three short technical explanations (60 to 75 seconds), trying to recall the maximum of information, and then filling out a questionnaire.

The next section presents our methodology in detail. The results are described in section 3 and discussed in section 4. A few concluding remarks are presented in section 5.

2. METHOD

2.1 Participants

Two groups of users from our laboratory participated in the experiment: 9 male adults (age range 23 to 51, mean = 30.7) and 9 female adults (age range 22 to 50, mean = 29.2). These two groups did not differ in age ($F(1/16) = 0.129$; N.S.).

2.2 Apparatus

Animations were presented on a 19" computer screen (1024*768 resolution) and loudspeakers were used for speech synthesis with IBM ViaVoice. In addition to speech synthesis, the text of the agent's presentation was displayed sentence by sentence on the top of the screen (see Figure 1 to 3; the initial text was in French).

2.3 Scenarios

The presentations were three short technical explanations, dealing with the functioning of a video-editing software, a remote control and a copy machine. The main difficulty lay in ambiguities of position, color and shape of keys which are on the three objects. The explanations consisted of a sequence of relevant information (e.g. position of the button, its function...). The agents appeared in front of a black background and a whiteboard. Each

explanation was associated with a single picture displayed on this whiteboard (see Figures 1 to 3).

2.4 Independent Variables

The primary variable tested was the multimodal strategy of the agents. It comprised (the rules are detailed in the next section):

- Cooperation by redundancy (see Figure 1): relevant information (e.g. position, shape, size of objects) was given both by speech and arm gesture (deictic gesture toward the picture or iconic gesture).
- Cooperation by complementarity (see Figure 2): half of relevant information was given by speech, and the other half was given by gesture (deictic gesture toward the picture or iconic gesture).
- Cooperation by speech-specialization (see Figure 3): all information was given by speech. Arm movements were limited to beat gestures.

The appearance of the agents was the second variable investigated in this experiment. We used three 2D cartoon-like Limsi Embodied Agents that we have developed. The 2D agents technology that we used is described in [2]. Multimodal behavior of all agents was specified using a low-level XML language. In this experiment, we used one female agent and two male agents, namely Lea, Marco and Julien (see Figures 1 to 3). A demonstration is available on the Web¹. Combinations between agents' appearance, multimodal strategy and content of presentation were counterbalanced: each agent used each strategy and presented each object the same number of times across the users' sample.



Figure 1: Our Lea agent presenting the software with a redundant strategy.

Finally, the influence of users' gender on dependent variables was tested. Additional variables such as the content of the presentations or the order of presentations were considered as control variables, and were neutralized. The presentations were equivalent in duration for the three contents (75 seconds for redundant and speech-specialized scenarios, 60 seconds for complementary scenarios) and complexity. The presentation order

¹ <http://www.limsi.fr/Individu/martin/research/projects/lea/>

of the three explanations, of the three strategies and of the three agents were all randomized.

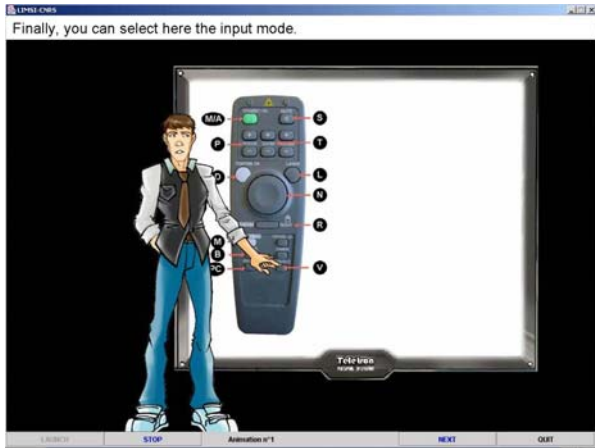


Figure 2: Our Marco agent presenting the remote control with a complementary strategy.



Figure 3: Our Julien agent, presenting the copy machine with a speech-specialized strategy.

2.5 Generation of multimodal behavior

In this section, we present the rules we followed when specifying the agents' behavior whatever their appearance. We first describe rules that were common to the three strategies, and then specific rules for each strategy.

2.5.1 Common rules for animation

- Eye blink: one eye blink was displayed every 0.7 second on an average. The eyes were never immobile more than 1.4 second.
- Lip-sync: when the agent was talking, the shape of the mouth changed every 0.2 second.
- Head moves: the agents slightly turned the head every 20 seconds on an average.

- Eyebrows: emphasis was displayed via eyebrows on certain words (ex: "on the **right**", or "the **blue** button"); in addition, eyebrow movements were randomly inserted every 3 seconds on an average.
- Intonation: this parameter was fixed as neutral.
- Gesture: the number of gestures was the same for all strategies. There was an arm/hand movement every 1.2 second on an average. The rate of beat gestures among them was minimal in redundant scenarios, intermediate in complementary scenarios, and maximal in speech-specialized scenarios.

2.5.2 Rules for generating redundant multimodal behavior

- Speech: for items of interest, absolute localization (e.g. "on the top left side") was used whenever it was possible; otherwise the agent used relative localization (e.g. "just below, you will find..."). Shape, color and size of items were given whenever it was a discriminative feature.
- Hand and arm gestures: shape and size were displayed via an iconic gesture when possible (with both hands). A deictic gesture was used for every object (finger or palm hand shape). Beat gestures were used when no other gesture was possible.
- Gaze: the agent glanced at target items during 0.4 second at the beginning of every deictic gesture.
- Eyebrows: shape of big objects was sometimes displayed via raised eyebrows.
- Locomotion: if needed, the agent moved closer to the target item before deictic gesture.

2.5.3 Rules for generating complementary multimodal behavior

- Speech: in comparison with redundant scenarios, information concerning localization, shape, color or size was given for half of the items.
- Hand and arm gestures: deictic or iconic gestures were used every time the information was not given by speech. Beat gestures were used the rest of the time.
- Gaze: the agent glanced at target items during 0.4 second at the beginning of every deictic gesture.
- Locomotion: if needed, the agent moved closer to the target item before deictic gesture.

2.5.4 Rules for generating speech-specialized multimodal behavior

- Speech: the same information as in redundant scenarios was given by speech (localization, shape, color, size).
- Hand and arm gestures: only beat gestures were displayed.

2.6 Dependent Variables

2.6.1 Performance data

After the three presentations, the users were given the three pictures used in the experiment. On this basis, they had to recall

the maximum of information they remembered. The experimenter marked out the performance (between 0 and 10) according to the number of information recalled (ex: "this is the start button").

2.6.2 Subjective variables

Finally, the users filled out a questionnaire in which they had to grade the three agents for the following questions:

- Which agent gave the best explanation?
- Which agent do you trust the most?
- Which agent is the most likeable?
- Did the agents have the same personality? Which one had the strongest personality?
- Which agent was the most expressive?

The users could also add free comments, and were particularly prompted to explicit their observations about the way each agent gave explanations.

The whole experiment lasted about 20 min for each user.

2.7 Analyses

Subjective variables as well as performance data were submitted to analyses of variance with user's gender as the between-subject factor. For each dependant variable, the analysis was successively performed using agent's strategy and agent's appearance as the within-subject factor. By way of control, the effects of the content of explanation were also tested. All the analyses were performed with SPSS².

3. RESULTS

The results described in this section will be discussed globally in the next section.

3.1 Subjective variables

3.1.1 Quality of explanation

The main effect of agent's strategy on ratings of quality of explanation proved to be significant ($F(2/32) = 5.469$; $p = 0.009$; see Figure 4). Indeed, agents with a redundant or a complementary strategy obtained equivalent ratings ($F(1/16) = 1.000$; N.S.) but were both rated better than agents with a speech-specialized strategy (respectively $F(1/16) = 13.474$; $p = 0.002$, and $F(1/16) = 4.102$; $p = 0.060$).

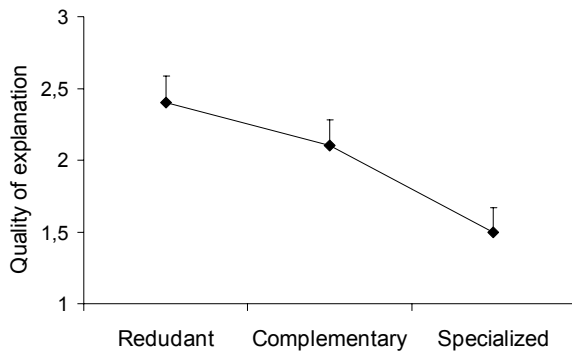


Figure 4: Ratings of the quality of explanation as a function of agent's multimodal strategy.

However, an interaction between strategy and user's gender ($F(2/32) = 4.980$; $p = 0.013$; see Figure 5) showed that the strategy effect was significant for male users ($F(2/16) = 19.000$; $p < 0.001$) but not for female users ($F(2/16) = 0.757$; N.S.). Male users rated the agents with a redundant strategy better than the others ($F(1/8) = 12.000$; $p = 0.009$ for complementary strategy and $F(1/8) = 100.000$; $p < 0.001$ for speech-specialized strategy). They also tended to rate complementary strategy better than speech-specialized strategy ($F(1/8) = 4.000$; $p = 0.081$).

No effect of agent's appearance or content of presentation was observed.

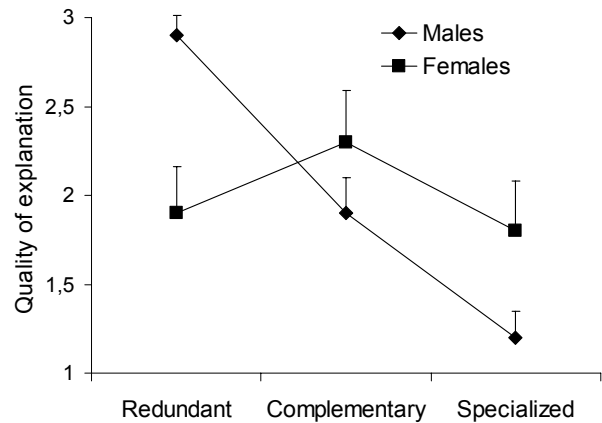


Figure 5: Ratings of the quality of explanation as a function of agent's multimodal strategy and user's gender.

3.1.2 Trust

No main effect of agent's strategy arose in subjective ratings of trust, but an interaction between strategy and user's gender appeared ($F(2/32) = 3.735$; $p = 0.035$; see Figure 6). In a similar way as for quality of explanation, the effect of agent's strategy tended to be significant for male users ($F(2/16) = 2.868$; $p = 0.086$). Although Figure 6 seems to show that female users had more trust in agents with complementary strategy, the overall effect of strategy was not statistically significant for female users ($F(2/16) = 2.500$; N.S.).

A positive linear correlation was found between this variable and ratings of quality of explanation (Pearson's correlation between 0.630 and 0.757, $p < 0.005$ for the three strategies).

No effect of agent's appearance or content of explanation was observed on ratings of trust.

² <http://www.spss.com>

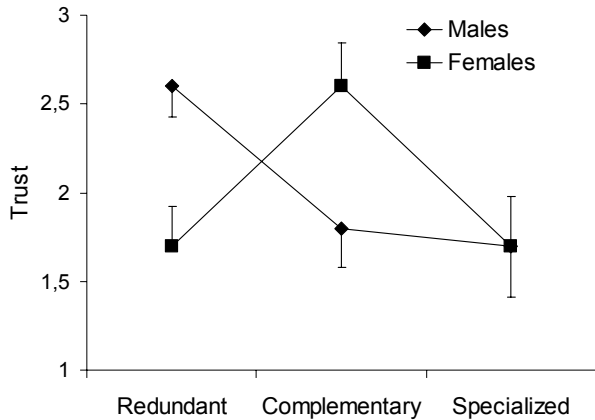


Figure 6: Ratings of trust as a function of agent's multimodal strategy and user's gender.

3.1.3 Likeability

Analyses on this variable yielded no effect of agent's strategy, but a main effect of appearance proved to be significant ($F(2/32) = 3.328$; $p = 0.049$; see Figure 7). It showed that no preference arose between Marco and Lea ($F(1/16) = 0.471$; N.S.), but Julien appeared less likeable than Marco ($F(1/16) = 6.479$; $p = 0.022$) and than Lea (in trend: $F(1/16) = 3.390$; $p = 0.084$). This effect did not vary with user's gender. Moreover, if Marco and Julien's scores are combined, no interaction between agent's gender and user's gender appears.

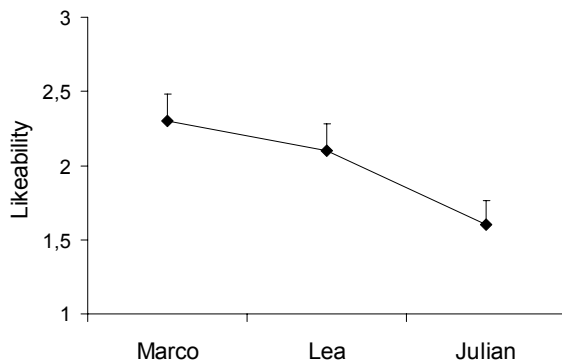


Figure 7: Ratings of likeability as a function of agent's appearance.

3.1.4 Personality and expressiveness

No effect of agent's strategy or appearance was observed on these variables.

3.2 Performance data

A main effect of user's gender on the amount of information recalled was significant in trend ($F(1/16) = 4.174$; $p = 0.058$), suggesting that female users recalled slightly more information (7.1 / 10) than male users (5.8 / 10).

Agent's strategy did not influence recall performance, but a main effect of agent's appearance neared significance ($F(2/32) = 3.215$; $p = 0.053$; see Figure 8), suggesting that recall was slightly better

when Marco had given the explanation, and slightly worse with Julien – recall with Lea being intermediate. This decrease of performance seems to follow the ratings of likeability, but no significant correlation between these two variables was found.

Concerning the influence of the content of explanation, no main effect arose, but an interaction between content and user's gender proved to be significant ($F(2/32) = 5.150$; $p = 0.012$). The effect of the content of explanation on recall performance was significant for female users ($F(2/16) = 9.838$; $p = 0.002$) but not for male users ($F(2/16) = 0.683$; N.S.). Actually, female users recalled more information about the copy machine than the two other objects. This effect, which constitutes a bias in our experiment, could come from a better previous familiarity of females with this object, although our two groups of users were homogeneous regarding socio-professional category.

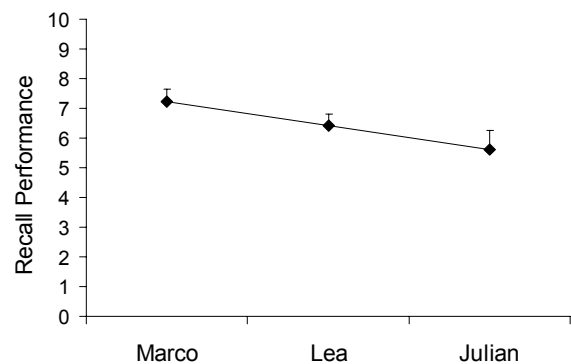


Figure 8: Recall performance as a function of agent's appearance.

4. DISCUSSION

4.1 Effects of multimodal strategies

The effect of multimodal strategy on ratings of quality of explanation proved to be globally significant and showed that redundant and complementary explanations were rated better than speech-specialized explanations. However, considering the interaction with user's gender, multimodal strategy of the agent appears to influence only male users' judgments. This effect seems to be partly implicit. The analysis of free comments given after the experiment shows that only 5 male users among 9 reported that they had observed differences about the way the three agents gave explanations. Moreover, they noticed that some agents made deictic gestures, but none of them mentioned cooperation between speech and gesture. Yet, male users' ratings not only proved to be massively influenced by multimodal strategies, but also clearly discriminated between redundant and complementary strategies. The fact that this effect appeared partly implicit in our experiment is consistent with Rimé's figure-ground model [15] in which the speaker's nonverbal behavior is usually at the periphery of the listener's attention.

To explain the unexpected gender difference, one could assume that female users were more focused on the object of explanation and less on the agents. However, they also made a lot of comments about agents' appearance and as many females as males (5 among 9) mentioned differences in agents' strategies of

explanation. Gender differences in decoding of nonverbal behaviors have been described in the literature, but they concerned facial expressions and reported greater decoding skills for females [7]. We will thus need further investigation of psycholinguistic literature to interpret our data.

Finally, this result is not clarified by performance data, since agent's strategy had no effect on user's recall. The fact that agent's strategy influenced subjective variables without affecting performance does not in any way detract from the importance of these multimodal strategies. Indeed, we think that subjective variables remain a crucial factor of engagement and determine, to a certain extent, the success of such multimedia tools.

Ratings of trust yielded the same kind of interaction between agent's strategy and user's gender. Actually, this result proved to be linked to the perceived quality of explanation. The fact that, in our experiment, multimodal strategy influenced trust (whereas agent's appearance did not) could be exploited in contexts for which trust is required (e.g. e-commerce...).

4.2 Effects of agents' appearance

Agent's appearance had no effect either on ratings of quality of explanation or ratings of trust. However, it had a significant effect on likeability, which was not influenced by user's gender. This result showed that Marco and Lea were preferred to Julien. Marco's smile happened to be designed broader than smiles of the other agents, and this was appreciated by the users, as they indicated after the experiment. Comments about Lea were more contradictory, because of her white coat: some users found her nicer and more serious; some others found her too strict. Finally, the fact that Julien's eyes were not so visible through his glasses was negatively perceived by most of the users. Besides, his position at rest consisted in having his arms folded, and several users found it unpleasant.

Agent's appearance also tended to influence recall performance of the users. Although this result lacks statistical significance, it warns us about the consequences of agent's design not only on user's satisfaction, but also on the effectiveness of the application. Performance was not shown to be correlated to ratings of likeability. In a similar way, [11] found that pedagogical efficacy of agents varied with their appearance, but they failed to find a link with any subjective variable (likeability, comprehensibility, credibility, quality of presentation, and synchronization of speech and animation). Further experiments are thus needed to confirm and interpret the influence of agent's appearance on recall performance.

4.3 Additional results

No effect of multimodal strategy or appearance of agents arose in perceived personality or expressiveness. Comments given by users at the end of the experiment indicated that three dimensions influenced their judgments: agent's appearance, amount of movements, and voice. This last parameter was not controlled in our experiment. It should also be noticed that 4 users (1 male and 3 females) did not find any personality differences between the three agents.

Finally, the bias produced by the content of presentation (better recall for females about one of the objects) could possibly explain the overall better performance of female users (obtained in trend).

5. CONCLUSIONS AND FUTURE DIRECTIONS

Our results stress the importance of both multimodal strategy and appearance to ensure the design of pleasant and effective presentation agents.

Taken as a whole, males and females subjective ratings show no preference between redundant and complementary scenarios. The advantage of complementary strategy lies in the possible reduction of the amount of information that needs to be transmitted: it enables avoiding both an overload of verbal information and an exaggerated gesticulation, which can be perceived as unnatural [5]. Complementary scenarios could also save time (20% in our experiment). However, we must keep in mind that male users found redundant strategies better, which suggests the interest of using these scenarios when target users are males or when the duration of presentation matters little. This recommendation might be valid for presentation tasks with spatial aspects (in our experiment, positions of elements were very important), but results could have been different in a more narrative context. Benefits of redundancy in pedagogical applications were previously observed (e.g. [6, 12]), but they concerned multimedia presentations (addition of text to auditory material) rather than multimodal behavior of agents.

Users' comments about agents' appearance suggested avoiding features such as a white coat or glasses, and behaviors such as folding arms. Conversely, a cartoonish broad smile seemed to be a predominant factor of likeability.

We intend to carry out further experiments within the same methodological framework, in order to complement this study with data on more subjects. Our 2D agents technology will be improved by going up from manual specification of behavior to higher-level specification language. We will also include different speech intonations, different energies and temporal patterns in movements, and some idiosyncratic gestures. Although 2D agents with individual behavior can be of interest for mobile applications, we are also considering the design of 3D agents.

We also suggest building agent's individuality from corpora of individual human behaviors. We believe that ECAs look as if they came from the same mould because they are usually specified by the same set of general psycholinguistic rules. So far, both the literature on individual multimodal behavior and the automatic extraction of context-dependent and individual rules from corpora annotation were neglected in the field of ECAs.

We expect to get more experimental results and then formulate a few recommendations for agent design in various application areas such as games or educational tools, which could also include teams of agents having each their own multimodal behavior. One issue will be the granularity of such design guidelines which should not be too specific in order to be useful to ECA designers.

6. ACKNOWLEDGMENTS

The work described in this paper was developed at LIMSI-CNRS and supported by the EU / HLT funded project NICE (IST-2001-35293): <http://www.niceproject.com/>. Our agents were designed by Christophe Rendu. The authors wish to thank their partners in the NICE project as well as William Turner, Frédéric Vernier and the reviewers for their useful comments.

7. REFERENCES

- [1] Aamas'2002. in Proceedings of Workshop on "Embodied conversational agents - let's specify and evaluate them!", in conjunction with The First International Joint Conference on "Autonomous Agents & Multi-Agent Systems", Bologna, Italy, 2002.
- [2] Abrilian, S., Busine, S., Rendu, C., and Martin, J.-C. Specifying Cooperation between Modalities in Lifelike Animated Agents in Proceedings of International Workshop on "Lifelike Animated Agents: Tools, Functions, and Applications", held in conjunction with the 7th Pacific Rim International Conference on Artificial Intelligence (PRICAI'02), Tokyo, Japan, 2002, pp. 3-8.
- [3] André, E., Rist, T., Van Mulken, S., Klesen, M., and Baldes, S. The automated design of believable dialogues for animated presentation teams in Embodied conversational agents, J. S. J. Cassell, S. Prevost, E. Churchill (Eds.). Ed.: MIT Press, 2000, pp. 220-255.
- [4] Ball, G. and Breese, J. Emotion and personality in a conversational character in Embodied Conversational Agents, J. Cassell, Sullivan, J., Prevost, S. and Churchill, E. (eds.), Ed.: Cambridge, MA: MIT Press, 2000, pp. 189-219.
- [5] Cassell, J. and Stone, M. Living hand to mouth: psychological theories about speech and gesture in interactive dialogue systems in Proceedings of Proceedings of the AAAI 1999 Fall Symposium on "Psychological Models of Communication in Collaborative Systems", North Falmouth, MA, 1999, pp. 34-42.
- [6] Craig, S. D., Gholson, B., and Driscoll, D. Animated pedagogical agents in multimedia educational environments: effects of agent properties, picture features, and redundancy. *Journal of Educational Psychology*, vol. 94 428-434, 2002.
- [7] Feldman, R. S., Philippot, P., and Custrini, R. J. Social competence and nonverbal behavior in *Fundamentals of Nonverbal Behavior*, R. S. F. B. Rimé, Ed.: Cambridge University Press, 1991, pp. 329-350.
- [8] Kendon, A. Gesticulation and speech: two aspects of the process of utterance in *The relationship of Verbal and Nonverbal Communication*, M. R. Key, Ed.: Mouton Publishers, 1980, pp. 207-228.
- [9] Martin, J. C., Grimard, S., and Alexandri, K. On the annotation of the multimodal behavior and computation of cooperation between modalities in Proceedings of Proceedings of the workshop on "Representing, Annotating, and Evaluating Non-Verbal and Verbal Communicative Acts to Achieve Contextual Embodied Agents" in conjunction with the 5th International Conference on Autonomous Agents (AAMAS'2001), Montreal, Canada, 2001, pp. 1-7.
- [10] Mcneills, D. *Psycholinguistics: a new approach*. New York: Harper & Row, 1987.
- [11] Moreno, K. N., Klettke, B., Nibbaragandla, K., and Graesser, A. C. Perceived characteristics and pedagogical efficacy of animated conversational agents in Proceedings of Proceedings of the 6th International Conference on Intelligent Tutoring Systems (ITS'2002), Biarritz, France and San Sebastian, Spain, 2002, pp. 963-971.
- [12] Moreno, R. and Mayer, R. E. Verbal redundancy in multimedia learning: when reading helps listening. *Journal of Educational Psychology* (94), pp. 156-163, 2002.
- [13] Nijholt, A. Towards multi-modal emotion display in embodied agents in Proceedings of Proceedings of the 5th Biannual Conference on Artificial Neural Networks and Expert Systems, Dunedin, New Zealand, 2001, pp. 229-231.
- [14] Rickel, J. and Johnson, W. L. Animated agents for procedural training in virtual reality: Perception, cognition, and motor control. *Applied Artificial Intelligence* (13), pp. 343-382, 1999.
- [15] Rimé, B. and Schiaratura, L. Gesture and speech in *Fundamentals of Nonverbal Behavior*, R. S. F. B. Rimé, Ed.: Cambridge University Press, 1991, pp. 239-284.