Conversational Behavior Reflecting Interpersonal Attitudes in Small Group Interactions

Brian Ravenet[1], Angelo Cafaro[2], Beatrice Biancardi[2], Magalie Ochs[3], and Catherine Pelachaud[2]

[1] Institut Mines-Télécom; Télécom ParisTech; CNRS-LTCI
[2] CNRS-LTCI; Télécom ParisTech
[3] Aix Marseille Université, CNRS, ENSAM, Université de Toulon, LSIS UMR7296,13397, Marseille, France
{ravenet,cafaro,biancardi,pelachaud}@telecom-paristech.fr
ochs@lsis.org

Abstract. In this paper we propose a computational model for the real time generation of nonverbal behaviors supporting the expression of interpersonal attitudes for turn-taking strategies and group formation in multi-party conversations among embodied conversational agents. Starting from the desired attitudes that an agent aims to express towards every other participant, our model produces the nonverbal behavior that should be exhibited in real time to convey such attitudes while managing the group formation and attempting to accomplish the agent's own turn-taking strategy. We also propose an evaluation protocol for similar multi-agent configurations. We conducted a study following this protocol to evaluate our model. Results showed that subjects properly recognized the attitudes expressed by the agents through their nonverbal behavior and turn taking strategies generated by our system.

1 Introduction

In a conversing group there might be from three up to twenty participants [2]. All participants adhere to specific social norms governing, for example, their distance and body orientation in order to coordinate and make it easier to interact with each other [15][32]. Goffman classifies participants into different roles (e.g. speakers and listeners) [11]. The role and the attitude that each participant aims at expressing towards the others determine the verbal and nonverbal behavior that are exhibited in such group interactions [20]. For virtual agents, the expressions of attitudes in groups is a key element to improve the social believability of the virtual worlds that they populate as well as the user's experience, for example in entertainment [19] or training [14] applications. This paper presents a model for expressing interpersonal attitudes in a simulated group conversation. The intended attitudes are exhibited via nonverbal behavior and impact the management of the group interaction (i.e. group formation), the conversational nonverbal behavior and the turn-taking strategy of the agents. The model is grounded on human and social sciences literature. We use the Argyle's representation of Status and Affiliation for describing interpersonal attitudes [1].

The main contributions of this paper are (1) a model that allows an agent to express interpersonal attitudes through its turn-taking strategies and nonverbal behavior while interacting in a small group, and (2) a study protocol designed to evaluate this model and similar scenarios involving small group interactions.

2   Related Work

ECAs gathering in groups. Prada and Paiva [24] modeled groups of autonomous synthetic virtual agents that collaborated with the user in the resolution of collaborative tasks within a 3D virtual environment. Rehm and Endrass [28] implemented a toolbox for modeling the behavior of multi-agent systems. In [23], the authors combined a number of reactive social behaviors, including those reflecting personal space [13] and the F-formation system [15], in a general steering framework inspired by [29]. This complete management of position and orientation is the foundation of the Impulsion Engine used in the work presented here. All these models did not take into account the expression of attitudes while exhibiting the behaviors of the agents.

Turn-taking models for ECAs. Ter Maat and colleagues [33] found that the perception of an ECA's personality (in a Wizard of Oz dyadic setting) varies when changing its turn-taking strategy. In [25], they proposed a model to manage the turn-taking between a user and a spoken dialog system using data-driven knowledge. In [18], the authors proposed to add to the ECA Max a turn-taking system based on states (Wanting the turn or Yielding the turn for instance) and possible transitions between them. A similar work on ECA's turn-taking is the model YTTM by Thórisson [34]. While it started as a turn-taking model for agents in dyadic interactions, it has been extended to multi-party interactions. It is based on Sacks' turn-taking model [30] and Duncan's findings on behavior for turn-taking management [9]. Another work focusing on these behaviors is the Spark architecture presented in [35]. This architecture supports the automatic generation of animations for avatars depending on the conversational activity of their user (typing on a keyboard). Some of these systems were designed for face-to-face interaction only [33,25,18] or did not consider the expression of attitudes [34].

ECAs expressing interpersonal attitudes. Different models enabling ECAs to exhibit social attitudes through their verbal and non-verbal behavior have been proposed. For instance, in [10], the system Demeanour supported the design of virtual characters within a group with different social attitudes (expressed through posture and gaze) following Argyle's Status and Affiliation model. In [28], the authors proposed a toolbox for manipulating the decision making of agents in a group based on different theories on social relations. In [17] they designed the nonverbal behavior of ECAs depending on their conversational role and their social relations. In [3], the authors conducted a study where users evaluated the perception of an ECA's attitude (friendly, hostile, extraversion) in the first seconds of an encounter with an ECA exhibiting different behaviors (smiles, gaze and proxemics). In [4], they evaluated the perception of attitudes

(friendly, hostile) conveyed in both the nonverbal behavior of the agent while speaking and the content of the speech. In [6], they explored how sequences of nonverbal signals (a head nod followed by a smile for instance) can convey interpersonal attitudes in a dyadic interaction.

These systems have been mainly designed for face-to-face interaction and scripted scenarios. In our system, we propose a model for the automatic generation of nonverbal behavior, we do not focus onverbal content at the moment, and it takes into account multi-party interaction in a small group formation.

3 Computational Model

We propose a computational model for the generation of nonverbal behavior supporting simulated group conversations (F-formation, turn-taking and conversational behaviors) and conveying interpersonal attitudes. The components of our model are the following: a *turn-taking component*, a *group behavior component* and a *conversational behavior component*. The expression of interpersonal attitudes is obtained, within each component, by modulating the produced nonverbal behavior as function of the attitude that every agent intends to express towards all the other group members. Given the different roles that participants can assume in a group conversation [11], as a first step, we focus on speaker and listener roles. The turn-taking mechanism is based on Clark's model [7] and builds on top of previous agent multi-party interaction research [34]. Our model triggers nonverbal behavior specific to each agent depending on its role in the conversation and its social attitudes towards the members of the group. This computation is done in **real-time** and in a **continuous** fashion. Moreover, our model deals with conflicting behaviors emerging from different social norms (interpersonal attitude, group formation and conversational behavior). For instance, one agent may need to orient itself towards its addressee and keep an orientation towards the group at the same time while gazing at the most dominant agent in the group.

3.1 Turn-taking Component

We based the design of this component on Clark's work [7]. Clark described a turn as an emergent behavior from a joint action of speaking and listening. Therefore our model generates a value for the desire of speaking that will trigger an utterance depending on the current speakers. This means that more than one agent at time can speak. We hypothesize that the attitudes expressed towards each other members of the group affect this desire to speak. We assumed that each agent has the intention to communicate and we do not consider the content of speech but only the willingness to actually utter sounds (sequence of words). However, we are aware of the importance of the speech content in an interaction so our model is designed in a way it could receive inputs from a dialog manager (speech utterances and desire to speak) in future works. An agent successfully takes the floor on the basis of the interpersonal attitudes it wants to

express towards the others. We modeled the turns as a state machine similarly to Thórisson [34]. These states are based on Goffman's ratified conversational roles [11] and are the following (as depicted in Figure 1): *own-the-speech* when the agent is the only one speaking, *compete-for-the-speech* when the agent is not the only one speaking, *end-of-speech* transitional state when the agent is willing to stop speaking, *interrupted* transitional state when the agent is forced to stop speaking, *addressed-listener* when the agent is listening and directly addressed to, *unaddressed-listener* when the agent is listening and not directly addressed to and *want-to-speak* when the agent is willing to speak and let the other know it by its behavior. The states and the transitions between them are depicted Figure 1.
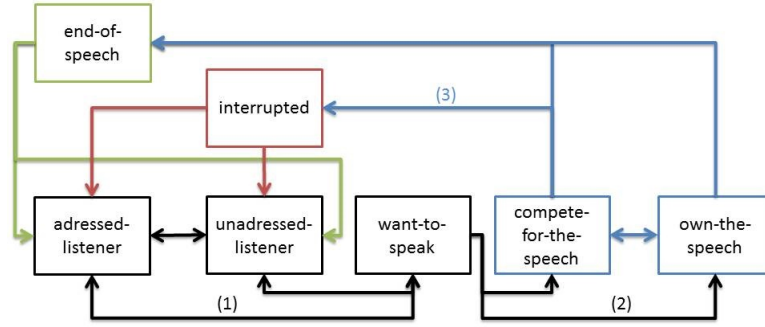
Fig. 1. The states of the turn-taking component

An agent starts in the *unaddressed-listener* state. Every 250ms, the system check if a transition from the current state activates. It is the minimum time required for a human to react to a stimuli [16]. For a transition to activate, it needs specific input values. These inputs are: the current list of speakers and who they are talking to, what attitudes the agents express to the others and the time passed since the last time the agent spoke. We chose to represent the attitudes on a two-axis space where each axis (Status and Affiliation) goes from -1 to 1. The minimum value of the Status axis represents a submissive attitude (respectively a hostile attitude on the Affiliation axis) and the maximum value represents a dominant attitude (respectively a friendly attitude). Regarding the transitions from and to *unaddressed-listener* and *addressed-listener*, they activate if another agent is addressing the agent or not. From *unaddressed-listener* or *addressed-listener* to *want-to-speak*, the function is using the time of the last utterance ($Last$) and the attitudes of the agent ($Stat$ and Aff as respectively the mean Status expressed and the mean Affiliation expressed). According to [5] and [12] the higher the Status and the Affiliation, the more a person is willing to speak and the quicker he/she wants to take the floor again. Whereas in a submissive attitude the desire to speak drops, an hostile attitude does not

have that effect. Since in our simulation, all agents have the intent to speak, even if they are submissive, for instance, they keep their desire of speaking. We empirically defined a maximum delay ($Delay$) before having the desire of speaking again. Therefore our transitional function $f_{li,wa}$ (from a listener state to *want-to-speak* state), where Now is the current time, is the following:

$$f_{li,wa}(Last, Stat, Aff) = \begin{cases} 1 & if(Now \geq Last+ \\ & Delay * (1 - \frac{1}{2}(Stat + norm(Aff))) \\ 0 & else \end{cases} \qquad (1)$$

The function $norm()$ is the normalization from $[-1, 1]$ to $[0, 1]$. From *want-to-speak* to *own-the-speech* or *compete-for-the-speech*, there are two strategies (i.e. transitions). First, the agent tries to see if there is another agent willing to give it the turn. If this does not succeed, the second strategy is trying to get into the conversation by interrupting or overlapping. However, in order to do so, the agent has to feel that it is compatible with its attitudes towards the others. The inputs of this transition are the currently speaking agents and the attitudes towards them. Let $Stat_{sp}$ and Aff_{sp} be respectively the mean Status and the mean Affiliation expressed towards the current speakers. A person expressing dominance interrupts others more easily [12]. Expressing friendliness and hostility result in possible overlapping [12,21]. Therefore our function $f_{wa,sp}$ (from *want-to-speak* to a speaker state) is the following:

$$f_{wa,sp}(Stat_{sp}, Aff_{sp}) = \begin{cases} 1 & if(Stat_{sp} + |Aff_{sp}| > 0) \\ 0 & else \end{cases} \qquad (2)$$

The transition from and to *compete-for-the-speech* and *own-the-speech* is activated if at least an other agent is talking at the same time. Within these states, the model selects the addressee by choosing randomly among the other agents, with a preference for those towards whom it expresses the most friendliness. From *compete-for-the-speech* to *interrupted*, it is similar to the $f_{wa,sp}$ function but this time the agent is not trying to interrupt but it is trying not to be interrupted. Therefore our function $f_{sp,in}$ is the following:

$$f_{sp,in}(Stat_{sp}, Aff_{sp}) = \begin{cases} 1 & if(Stat_{sp} + |Aff_{sp}| \leq 0) \\ 0 & else \end{cases} \qquad (3)$$

From *compete-for-the-speech* or *own-the-speech* to *end-of-speech*, this transition is activated automatically at the end of the agent's speech. From *interrupted* or *end-of-speech* to *unaddressed-hearer* or *addressed-hearer*, the transition is activated after producing the behavior associated to these states (see Section 3.3). Within each state, the agent produces different behaviors depending on the interpersonal attitudes it expresses. As the attitudes are also used to determine the flow among the internal states, this model differs from previous existing models as it allows an agent to convey interpersonal attitudes through its turn-taking strategies.

3.2   Group Behavior Component

Our model alters the behavior of the agents supporting the group formation so their different interpersonal attitudes are reflected accordingly. The model comes as a value producer for the parameters of a **Group Behavior** component that keeps the coherence of the formation based on the works of Kendon's F-Formation [15] and Scheflen's territoriality [32]. These parameters are preferred interpersonal distances between each member of the group, preferred target for gaze and preferred target for body orientation. The desired interpersonal distance ranged from 0.46m to 2.1m. These are the boundaries of personal and social areas as defined by Hall's proxemics theory [13], and these areas represent the space where social interaction takes place. We then compute a factor of interpolation α in this interval based on the expressed dominance and liking. A difference of status (either Dominance or Submissiveness) leads to a higher distance whereas friendliness leads to a smaller distance [20][8]. Let $Stat_i$ and Aff_i be respectively the expressed Status and Affiliation towards an agent Ag_i, the function $d(Ag_i)$ to compute the desired interpersonal distance with an agent Ag_i is:

$$d(Ag_i) = 0.46 + \alpha(2.1 - 0.46) \text{ with } \alpha = \frac{(|Stat_i| + (\frac{1-Aff_i}{2}))}{2} \qquad (4)$$

For the preferred gaze target, we compute a potentiality to be gazed at for each other participant and we select the participant with the higher potentiality. We do a similar process for the preferred body target. In order to compute the potentiality to be gazed at, we sum the result of a trigonometric interpolation on both the dominance expressed and the liking expressed whereas we consider only the dominance for the potentiality of the body orientation [20]. The function $p_g(Ag_i)$ to compute the potentiality to gaze at an agent Ag_i is:

$$p_g(Ag_i) = 1 + \frac{sin(\frac{1}{2} + 2 \times Aff_i) + sin(\frac{1}{2} + 2 \times (-Stat_i))}{2} \qquad (5)$$

And the function $p_b(Ag_i)$ to compute the potentiality to orient the body towards an agent Ag_i is:

$$p_b(Ag_i) = \frac{1 - Stat_i}{2} \qquad (6)$$

3.3   Conversational Behavior Component

In the initial state (*unaddressed-listener*), the agent is idle and the *Group Behavior* component is the only one producing an output, thus handling the agent's group behavior (i.e. interpersonal distance, body orientation and gaze). In the other states, the *Conversational Behavior* component might request resources (i.e. body joint to animate) from the *Group Behavior* one (preferred gaze target and preferred body orientation) to achieve the behavior needed by the turn-taking component and will also produce additional behavior such as gestures and facial expressions. To realize these additional behaviors, we extended the

model introduced in [27]. This model has been learnt on data from a crowdsourcing experiment where participants configured the behavior of a virtual agent for different attitudes. The proposed model works as a SAIBA Behavior Planner. Upon the reception of an utterance and an attitude, the model generates the nonverbal behavior associated with that intention and attitude (gestures, facial expressions, head orientations and gaze behaviors). For instance, it can produce three different facial expressions (a smiling, frowning or neutral face) blended with the lips animation produced while speaking. It also produces beat gestures to accompany the speech with different amplitude (small, normal or wide) and strength (weak, normal or strong). When entering the *want-to-speak* state, the model outputs as preferred gaze target the current speaker to indicate the desire to take the floor [9]. In the states *own-the-speech* or *compete-for-the-speech*, the *Conversational Behavior* component receives from the *Turn-Taking* component an utterance to produce (a target to address and a sentence to say) and it retrieves from the *Group Behavior* component the details of the group. From these parameters, it will generates the nonverbal behavior corresponding to the interpersonal attitudes expressed thanks to the *Behavior Planner* from [27] and selects as preferred body orientation the addressee. However, when in the state *compete-for-the-speech*, the preferred gaze target is the agent towards whom the most submissiveness is expressed. In the *own-the-speech* state, the preferred gaze target is the addressee [20]. In the *end-of-speech* state, the preferred gaze target is the addressee [9]. In the state *interrupted*, the previous behavior of the agent is interrupted and both the preferred gaze target and preferred body orientation are the other speaker towards whom the highest submissiveness is expressed [12].

4   Implementation

The implementation of our model has been realized as a Unity3D application that uses two additional frameworks, the VIB/Greta agent platform [22] and the Impulsion AI library [23]. Once the application is started, a scene that takes place in a public square starts in which the user can navigate. Within the public square, the user can find a group of agents having a conversation, simulated by our model. The two additional frameworks handles the production of nonverbal behaviors for each agent. The VIB agent platform computes gestures and facial expressions accompanying communicative intentions while the Impulsion AI library handles group cohesion behaviors, body orientation, gaze direction and spacial distance between group members. The combined integration of these two frameworks has been presented in [26]. These frameworks have been extended in order to take into account the interpersonal attitudes in the production of behaviors. Our implemented turn-taking model, written in Unity3D scripts, encapsulates these two frameworks within the application and sends them the current attitudes and the utterances of each agent. In response, VIB and Impulsion produce the corresponding nonverbal behaviors.

5   Model Evaluation

We questioned whether the attitude expressed by our agents would emerge and be perceived by user when observing the behavior of the agents during simulated conversation. It was impractical to test all possible group configurations. Therefore we fixed at 4 the number of agents in order to represent a typical small gathering featured in the applications mentioned earlier in Section 1. A group of four agents expressing attitudes on two dimensions (i.e. status and affiliation) towards all other members yields an exponential number of possibilities. Considering the levels of attitude on a discrete scale (e.g. submissive vs. neutral vs. dominant) there are 3^{24} possible configuration among 4 agents. Therefore, we simplified the design by splitting the study in two trials focusing on each separate dimension, respectively named **Status Trial** and **Affiliation Trial**.

Secondly, based on the Interpersonal Complementarity Theory (IC) [31], we studied the attitudes expressed by two participants as described in the following section.

5.1   Experimental Design

Fig. 2. A screenshot of the group interaction as seen by users in our evaluation study. The two central agents were respectively identified as the Left Agent (the male in this image) and the Right Agent (the female).

The IC theory claims that to obtain a positive outcome in an interaction, people behaviors need to reflect a similar attitude on the Affiliation axis and/or an opposite attitude on the Status axis. Inspired by this theory, for each trial, we asked participants to watch videos taken from our implemented system featuring 4 agents (2 males and 2 females, all with a distinct appearance) in a group conversation. The user is an observer and the agents are positioned in a

way that two appear at the immediate sides of the frame and two are located at the center of the frame (as depicted in Figure 2). The participants rated how they perceived the attitudes of the two central agents while they were expressing similar and opposite attitudes towards each other according to the IC theory. We called these two agents **Left Agent** and **Right Agent**. The main research questions were the following: will the participants recognize the attitudes (validating the model we proposed for a subset of the possible configurations)? Is there any difference in the perceived attitudes when showing complementarity and anti-complementarity situations to participants? Does our model generate believable conversational group simulation and do users find it more believable in complementarity situations?

Stimuli. We describe the videos that were presented to the participants. We were aware of possible appearance and positioning biases (e.g. clothing and gender). Since we could not fully factor in these elements in our design, we considered them as blocking factors. The four characters could be arranged in 6 different possible orders while keeping a circular group formation, we named this blocking factor *GroupArrangement*. Given an arrangement, the left and right agents (those exhibiting the stimuli, positioned at the center of the frame) could be swapped, we named this factor *PairPosition*. While the left and right agent (at the center of the frame) were the two exhibiting the stimuli, the other two side agents were still actively participating in the simulated conversations expressing and receiving a neutral attitude in both dimensions. According to our blocking factors, we generated 48 video stimuli for each trial. In both trials, the independent variables (IVs) were the following: Expressed Left Agent Status **ExpStatusL** (respectively **ExpAffL** in the *Affiliation* trial) and Expressed Right Agent Status **ExpStatusR** (respectively **ExpAffR** in the *Affiliation* trial). Both variables had two levels, *Submissive* and *Dominant* (respectively *Hostile* and *Friendly*).

Measurements. For each video stimuli, we asked the participants to answer a questionnaire in order to measure their perceived attitudes of the two agents. This questionnaire was designed by including adjectives classified around the interpersonal circumplex chosen from [36]. There were 4 questions on how accurate a sentence using one the 4 adjectives (2 with positive valence and 2 with negative) described the attitude of an agent towards the other one. For the *Status* Trial, the adjectives were: controlling, insecure, dominating and looking for reassurance. For the *Affiliation* Trial, the adjectives were: warm, detached, likeable and hostile. One of the question was, for instance, *Left Agent expresses warmth towards the Right Agent*. We also asked 3 questions about the group: its believability, the engagement of its participants and its social richness. All answers were on 5 points Likert scales (anchors *Completely disagree* and *Completely agree*). In sum, the dependent variables (DVs) in the *Status* trial (and respectively the *Affiliation* trial) were the Measured Left Agent Status **MeasureStatusL** (respectively **MeasureAffL**) and the Measured Right Agent Status **MeasureStatusR** (respectively **MeasureAffR**). We aggregated the answers from positive

and negative items to produce a single normalized value for each DVs in the range [0,1]. In both trials, the variables related to the questions about the group were: the **Group Believability**, the **Group Engagement** and the **Group Social Richness**.

Hypotheses

- **H1 (Left Agent)**: The value of MeasureStatusL is higher when ExpStatusL is at Dominant level as opposed to Submissive.
- **H2 (Right Agent)**: The value of MeasureStatusR is higher when ExpStatusR is at Dominant level as opposed to Submissive.
- **H3 (IC Theory)**: With respect to Interpersonal Complementarity theory, participants should better recognize the attitudes when ExpStatusL and ExpStatusR show opposite values
- **H4 (Group)**: The values for Group Believability, Group Engagement and Group Social Richness should be rated higher in complementarity configuration than anti-complementarity configuration.

We call H1.s, H2.s, H3.s and H4.s the hypotheses in the Status Trial and H1.a, H2.a, H3.a and H4.a the hypotheses in the Affiliation Trial.

Procedure and participants. In both trials, each participant was assigned to 4 videos in a fully counterbalanced manner according to our blocking factors. We ran the study on the web. A participant was first presented with a consent page and a questionnaire to retrieve demographic information (nationality, age and gender). Then, we showed a tutorial video with a sample question. And then we presented in a fully randomized order the 4 videos, each in a different page with questions at the bottom, in a within-subjects design. Finally, a debriefing page was shown. We recruited a total of 144 participants via mailing lists, 72 in each trial. In the Status Trial, 58.34% of the participants were between 18 and 30 years old and 50% were female, 48.61% were male and 1.39% did not say. In the Affiliation Trial, 66.66% were between 18 and 30, 56.94% were female while 43.06% were male. We had participants from several cultural backgrounds but most of them were from France (47.22% in the Status Trial and 52.68% in the Affiliation Trial).

5.2   Results

Status Trial. In order to test H1.s, H2.s and H3.s, we ran a 2x2 repeated measures MANOVA (Doubly Multivariate Analysis of Variance) on MeasureStatusL and MeasureStatusR with within-subjects factors ExpStatusL and ExpStatusR. We found an overall main effect of ExpStatusL ($WilksLambda = 0.50$, $F(2,70) = 35.7$, $p < 0.001$) and ExpStatusR ($WilksLambda = 0.57$, $F(2,70) = 25.7$, $p < 0.001$). No significant interaction effects were found. Since the sphericity assumption was not violated, we performed a follow-up analysis that looked

at univariate repeated measures ANOVAs for the 2 DVs. For MeasureStatusL, the ANOVA confirmed a significant main effect of ExpStatusL ($F(1, 71) = 63$, $p = .0$). In particular, Left Agent was rated as more dominant when ExpStatusL was at the Dominant level (M=.56, SE=.02) as opposed to the Submissive level (M=.43, SE=.01). No other interaction effects were found therefore **H1.s** is **supported**. For MeasureStatusR, the ANOVA confirmed a significant main effect of ExpStatusR ($F(1, 71) = 48$, $p = .0$). In particular, Right Agent was rated as more dominant when ExpStatusR was at the Dominant level (M=.60, SE=.01) as opposed to the Submissive level (M=.46, SE=.01). No other interaction effects were found therefore **H2.s** is **supported**. The interaction of the two IVs had no effects on both measures, **H3.s** is **rejected**. We ran a further MANOVA analysis with two additional between-subjects factors Group Arrangement and Subject Gender. We did not find significant interaction effects (all $p > .38$). As for the 3 group measures, we ran 3 similar univariate repeated measures ANOVAs. Except for an effect of ExpStatusR on Group Believability ($p = 0.03$) with a small effect size ($\eta_p^2 = .06$), no other significant main effects and interactions were found (all $p > .13$).

Affiliation Trial. We ran a similar 2x2 repeated measures MANOVA on MeasureAffL and MeasureAffR with within-subjects factors ExpAffL and ExpAffR. We found an overall main effect of ExpAffL ($WilksLambda = 0.42$, $F(2, 70) = 48.4$, $p < 0.001$) and ExpAffR ($WilksLambda = 0.55$, $F(2, 70) = 28$, $p < 0.001$). No significant interaction effects were found. We also performed a follow-up analysis that looked at univariate repeated measures ANOVAs for the 2 DVs. For MeasureAffL, the ANOVA confirmed a significant main effect of ExpAffL ($F(1, 71) = 93$, $p = .0$). In particular, Left Agent was rated as more friendly when ExpStatusL was at the Friendly level (M=.61, SE=.02) as opposed to the Hostile level (M=.37, SE=.01). No other interaction effects were found therefore **H1.a** is **supported**. For MeasureAffR, the ANOVA confirmed a significant main effect of ExpAffR ($F(1, 71) = 54$, $p = .0$). In particular, Right Agent was rated as more friendly when ExpAffR was at the Friendly level (M=.58, SE=.02) as opposed to the Hostile level (M=.41, SE=.01). No other interaction effects were found therefore **H2.a** is **supported**. The interaction of the two IVs had no effects on both measures, **H3.a** is **rejected**. We also ran a further MANOVA analysis with two additional between-subjects factors Group Arrangement and Subject Gender. We did not find significant interaction effects (all $p > .17$). As for the 3 group measures, we ran 3 similar univariate repeated measures ANOVAs. Except for an effect of ExpAffR on Group Engagement ($p = 0.01$) with a small effect size ($\eta_p^2 = .09$), no other significant main effects and interactions were found (all $p > .12$).

6 Discussion and Conclusion

We presented a computational model for generating agents' nonverbal behavior in a conversational group. This nonverbal behavior supports the expres-

sion of interpersonal attitudes within group formation management, conversational behavior and turn-taking strategies adopted by the agents. We also designed an evaluation protocol that we used to conduct a two trials study aimed at testing the capacity of this model to produce believable attitudes in anti-complementarity and complementarity situations. Results showed that agents' attitudes were properly recognized (H1 and H2 supported in both trials). We didn't find any interaction effect between the expressed attitudes as the IC theory suggests (H3 rejected in both trials). The reason might be that since we did not consider the content of the speech, it was maybe easier for participants to clearly distinguish each attitude (and not to consider them in interaction). The expressed Attitudes (both in the Status and Affiliation trials) of the two central agents (i.e. the left and right agents for which we manipulated the attitudes expressed) had a main effect on the respective measured Attitudes. Similarly, we obtained means for the Group Dependent Variables (Believability, Engagement, and Social Richness) all > 3.472 (outcomes were in the range 1-5) but we didn't find any significant differences when looking at complementarity or anti-complementarity situations (H4 rejected). Finally, the blocking factors (in particular the group arrangement) and the user's gender have been considered as between-subjects factors but they had no effects. Each agent was able to express its attitude regardless of the other's attitude, the group arrangement and the gender of the user. Some limitations should be considered. The model should be extended with additional nonverbal behaviors (e.g. supporting backchannels) and the generation of verbal content (also reflecting interpersonal attitudes). Regarding the evaluation, we have considered only a subset of all the possible configurations, limiting the manipulated attitudes to the two (central) characters. However, the intended attitudes that our model aimed at expressing, emerged from the overall group behavior exhibited by the agents. Furthermore, we introduced an evaluation protocol that other researchers could adopt when running similar studies on group behavior. In the short term, we intend to make the user an active participant in the group conversation, allowing him/her to interact with the agents and have the agents expressing their attitudes towards the user.

7   Acknowledgments

References

1. Argyle, M.: Bodily Communication. University paperbacks, Methuen (1988)

2. Beebe, S.A., Masterson, J.T.: Communication in small groups: principles and practices. Boston: Pearson Education, Inc (2009)
3. Cafaro, A., Vilhjálmsson, H.H., Bickmore, T., Heylen, D., Jóhannsdóttir, K.R., Valgarðsson, G.S.: First impressions: Users' judgments of virtual agents' personality and interpersonal attitude in first encounters. In: Intelligent Virtual Agents. pp. 67–80. Springer (2012)
4. Callejas, Z., Ravenet, B., Ochs, M., Pelachaud, C.: A computational model of social attitudes for a virtual recruiter. In: Autonomous Agent and Multiagent Systems (2014)
5. Cappella, J.N., Siegman, A.W., Feldstein, S.: Controlling the floor in conversation. Multichannel integrations of nonverbal behavior pp. 69–103 (1985)
6. Chollet, M., Ochs, M., Pelachaud, C.: From non-verbal signals sequence mining to bayesian networks for interpersonal attitudes expression. In: Intelligent Virtual Agents. pp. 120–133. Springer (2014)
7. Clark, H.H.: Using language. Cambridge University Press (1996)
8. Cristani, M., Paggetti, G., Vinciarelli, A., Bazzani, L., Menegaz, G., Murino, V.: Towards computational proxemics: Inferring social relations from interpersonal distances. In: Privacy, security, risk and trust. pp. 290–297. IEEE (2011)
9. Duncan, S.: Some signals and rules for taking speaking turns in conversations. Journal of personality and social psychology 23(2), 283 (1972)
10. Gillies, M., Crabtree, I.B., Ballin, D.: Customisation and context for expressive behaviour in the broadband world. BT Technology Journal 22(2), 7–17 (2004)
11. Goffman, E.: Forms of talk. University of Pennsylvania Press (1981)
12. Goldberg, J.A.: Interrupting the discourse on interruptions: An analysis in terms of relationally neutral, power-and rapport-oriented acts. Journal of Pragmatics 14(6), 883–903 (1990)
13. Hall, E.T.: The hidden dimension, vol. 1990. Anchor Books New York (1969)
14. Johnson, W.L., Marsella, S., Vilhjalmsson, H.: The darwars tactical language training system. In: Proceedings of I/ITSEC (2004)
15. Kendon, A.: Conducting interaction: Patterns of behavior in focused encounters, vol. 7. CUP Archive (1990)
16. Kosinski, R.J.: A literature review on reaction time. Clemson University 10 (2008)
17. Lee, J., Marsella, S.: Modeling side participants and bystanders: The importance of being a laugh track. In: Vilhjálmsson, H.H., Kopp, S., Marsella, S., Thórisson, K.R. (eds.) Intelligent Virtual Agents, Lecture Notes in Computer Science, vol. 6895, pp. 240–247. Springer Berlin Heidelberg (2011)
18. Leßmann, N., Kranstedt, A., Wachsmuth, I.: Towards a cognitively motivated processing of turn-taking signals for the embodied conversational agent max. pp. 57–64. Proceedings Workshop Embodied Conversational Agents: Balanced Perception and Action, IEEE Computer Society (2004)
19. Maxis: http://www.thesims.com (Nov 2014), `http://www.thesims.com`
20. Mehrabian, A.: Significance of posture and position in the communication of attitude and status relationships. Psychological Bulletin 71(5), 359 (1969)
21. O'Connell, D.C., Kowal, S., Kaltenbacher, E.: Turn-taking: A critical analysis of the research tradition. Journal of psycholinguistic research 19(6), 345–373 (1990)
22. Pecune, F., Cafaro, A., Chollet, M., Philippe, P., Pelachaud, C.: Suggestions for extending saiba with the vib platform. In: Proceedings of the Workshop on Architectures and Standards for Intelligent Virtual Agents at IVA 2014 (2014)
23. Pedica, C., Vilhjálmsson, H.H., Lárusdóttir, M.: Avatars in conversation: the importance of simulating territorial behavior. In: Intelligent Virtual Agents. pp. 336–342. Springer (2010)

24. Prada, R., Paiva, A.: Believable groups of synthetic characters. In: Proceedings of the Fourth International Joint Conference on Autonomous Agents and Multiagent Systems. pp. 37–43. AAMAS '05, ACM, New York, NY, USA (2005)
25. Raux, A., Eskenazi, M.: A finite-state turn-taking model for spoken dialog systems. In: Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics. pp. 629–637. Association for Computational Linguistics (2009)
26. Ravenet, B., Cafaro, A., Ochs, M., Pelachaud, C.: Interpersonal attitude of a speaking agent in simulated group conversations. In: Bickmore, T., Marsella, S., Sidner, C. (eds.) Intelligent Virtual Agents, Lecture Notes in Computer Science, vol. 8637, pp. 345–349. Springer International Publishing (2014), `http://dx.doi.org/10.1007/978-3-319-09767-1_45`
27. Ravenet, B., Ochs, M., Pelachaud, C.: From a user-created corpus of virtual agent's non-verbal behaviour to a computational model of interpersonal attitudes. In: Proceedings of Intelligent Virtual Agent (IVA) 2013 (2013)
28. Rehm, M., Endrass, B.: Rapid prototyping of social group dynamics in multiagent systems. AI and Society 24, 13–23 (2009)
29. Reynolds, C.: Steering behaviors for autonomous characters. In: Proceedings of the Game Developers Conference. pp. 763–782. Miller Freeman Game Groups, San Francisco, CA (1999)
30. Sacks, H., Schegloff, E.A., Jefferson, G.: A simplest systematics for the organization of turn-taking for conversation. Language pp. 696–735 (1974)
31. Sadler, P., Woody, E.: Interpersonal complementarity. Handbook of interpersonal psychology: Theory, research, assessment, and therapeutic interventions p. 123 (2010)
32. Scheflen, A.E., Ashcraft, N.: Human territories: How we behave in space-time. (1976)
33. Ter Maat, M., Truong, K.P., Heylen, D.: How turn-taking strategies influence users'impressions of an agent. In: Intelligent Virtual Agents. pp. 441–453. Springer (2010)
34. Thórisson, K.R., Gislason, O., Jonsdottir, G.R., Thorisson, H.T.: A multiparty multimodal architecture for realtime turntaking. In: Intelligent Virtual Agents. pp. 350–356. Springer (2010)
35. Vilhjálmsson, H.H.: Animating conversation in online games. In: Entertainment Computing–ICEC 2004, pp. 139–150. Springer (2004)
36. Wiggins, J.S., Trapnell, P., Phillips, N.: Psychometric and geometric characteristics of the revised interpersonal adjective scales (ias-r). Multivariate Behavioral Research 23(4), 517–530 (1988)