Using a Shared Representation to Generate Action and Social Language for a Virtual Dialogue Environment

Swati Gupta, Marilyn A. Walker, Daniela M. Romano

University of Sheffield, UK

Department of Computer Science, University of Sheffield, Regent Court, 211 Portobello Street, Sheffield, UK, S14DP

{s.gupta, m.walker, d.romano}@dcs.shef.ac.uk

Abstract

This paper describes a conversational system POLLy (POliteness in Language Learning) which uses a common planning representation to generate actions to be performed by embodied agents in a virtual environment, and to generate spoken utterances for dialogues about the steps involved in completing the task. In order to generate socially appropriate dialogue, Brown and Levinson's theory of politeness is used to constrain the dialogue generation process. We report the results of a cross-cultural user experiment to test differences between user perceptions of the role of the social variables of power and social distance on the appropriate linguistic form of an utterance. Our results suggest that, contrary to Brown and Levinson's theory, different users have different perceptions of these variables, and weight them differently in social interaction.

Introduction

Research in Embodied Conversational Agents (ECAs) has explored embedding ECAs in domain-specific Virtual Environments (VE) where users interact with them using different modalities, including Spoken Language. There are a wide variety of applications (Wahlster et al, 1993; Bersot et al, 1998; Traum et al, 2003; Kopp et al, 2003; Takenobu et al, 2003; op den Akker, 2000; Predinger & Ishizuka, 2001; Paiva et al, 2005; Eichner et al, 2007). However, in order to support dialogic interaction in such environments, an important technical challenge is the synchronization of the ECA Spoken Interaction module with the ECA non-verbal actions in the VE.

In this paper, we propose an approach that uses a common high level representation which is broken down to simpler levels to generate the agents' verbal interaction and the agents' non-verbal actions synchronously for task-oriented applications that involve performing some actions to achieve a goal while talking about the actions using natural language. This approach is implemented in our Conversational System POLLy which produces utterances with a socially appropriate level of politeness as discussed in more detail below. We have implemented POLLy in a VE for the domain of teaching English as a second language (ESL). Figure 1 shows a user interacting with POLLy.
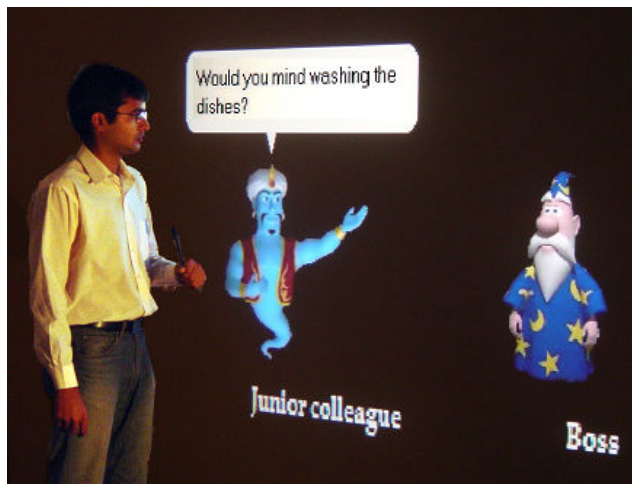

Figure 1: A user interacting with the Agents in RAVE

In previous work, Wahlster et al (1993) describe their system WIP which generates alternate multimodal presentations of the same content. Their basic underlying principle is that various constituents of a multimodal presentation should have a common representation of what is to be conveyed. While explaining a process, WIP generates and realizes plans for communicating domain plans provided by the back-end system. The elements of this plan are communicative acts that verbalize and visualize the physical acts specified in a given domain plan. Bersot et al (1998) present a conversational agent called Ulysses embedded in a collaborative VE which accepts spoken input from the user and enables him or her to navigate within the VE. They use a 'reference resolver' which maps the entities mentioned in utterances to geometric objects in the VE and to actions. Max, a VR based conversational agent by Kopp et al (2003) allows multimodal conversational abilities including natural language interaction for task-oriented dialogues in virtual construction tasks. It builds on a database of utterance templates formulated in an XML representation language which contains the verbal part, augmented with accompanying gestures and their cross-modal affiliation. In order to deal with the vagueness of language in specifying spatial relations in virtual space, the K_2 system (Takenobu et al 2003) proposed a bilateral symbolic and numeric
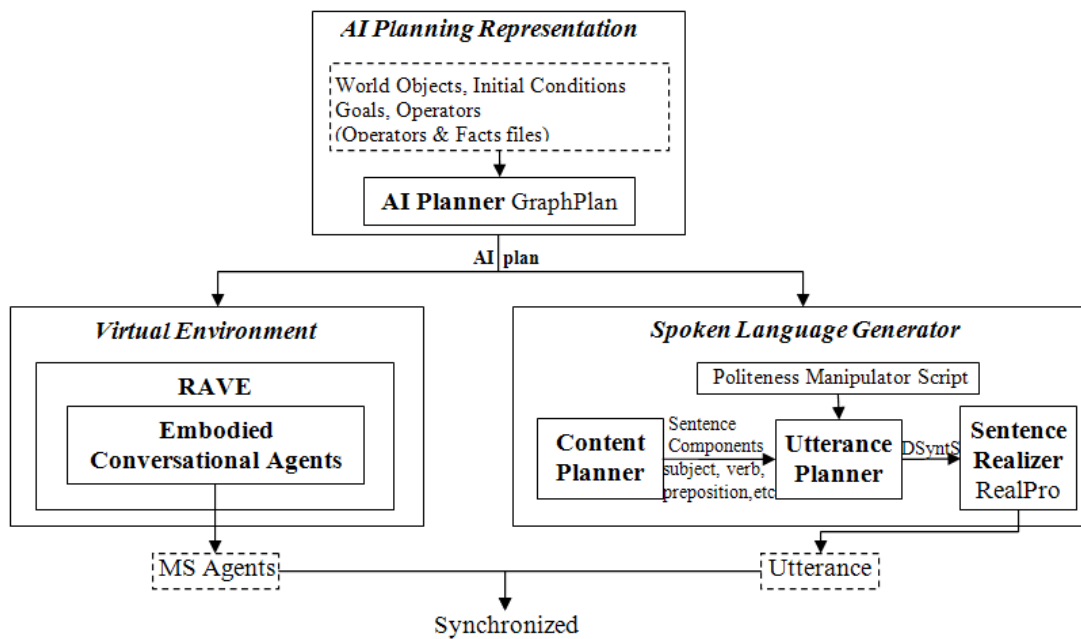
Figure 2: POLLY's Architecture

representation of locations, to bridge the gap between language processing (a symbolic system), and animation generation (a continuous system). K₂ extracts a user's goal from the utterance and translates it into animation data. The FearNot! demonstrator by Paiva et al (2005) provides training to kids against bullying via virtual drama in which one virtual character plays the role of a bully and the other plays the role of victim, who asks the child for advice. The advice modifies the character's emotional state and affects its behaviour in the next episode. FearNot!'s spoken interaction is template-based. In case of interaction between the child and the character, the incoming text is matched against a set of language templates. The information about the character's action is defined in a collection which contains the utterance to be spoken as well as the animation. Eichner et al (2007) describe an application in which life-like characters present MP3 players in a virtual showroom. An XML scripting language is used to define the content of the presentation as well as the animations of the agents.

Since these ECAs function in scenarios where they interact with the world, other agents, and the user, they must be 'socially intelligent' (Dautenhahn, 2000) and exhibit social skills. Our work is based on the hypothesis that the relevant social skills include the ability to communicate appropriately, according to the social situation, by building on theories about the norms of human social behaviour. We believe that an integral part of such skills is the correct use of politeness (Brown & Levinson, 1987; Walker et al 1997). For instance, note the difference in the effect of requesting the hearer to clean the floor by saying 'You must clean the spill on the floor now!' and 'I know I'm asking you for a big favour but could you kindly clean the spill on the floor?'

According to Brown and Levinson (1987) (henceforth B&L), choices of these different forms are driven by sociological norms among human speakers. Walker et al (1997) were the first to propose and implement B&L's theory in ECAs to provide interesting variations of character and personality in an interactive narrative application. Since then B&L's theory has been used in many conversational applications e.g. animated presentation teams (André et al 2000; Rehm & André, 2007), real estate sales (Cassell & Bickmore, 2003), and tutorials (Johnson et al, 2004; Johnson et al, 2005; Porayska-Pomsta 2003; Wang et al 2003). Rehm & André, 2007) show that gestures are used consistently with verbal politeness strategies and specific gestures can be used to mitigate face threats.

Work in literary analysis has also argued for the utility of B&L's theory, e.g. Culpeper (1996) argues that a notion of 'impoliteness' in dramatic narratives creates conflict by portraying verbal events that are inappropriate in real life. Thus impoliteness often serves as a key to move the plot forward in terms of its consequences.

We start with a brief explanation of the architecture and theoretical basis of POLLy, a description of our virtual environment and some example dialogues. We then describe a user experiment to study the human perception of politeness in two different situations; one where a stranger is talking to the subject and another in which a junior colleague is talking to the subject, in a collaborative task domain of cooking where subjects imagine that they are making a recipe together with another person. We show that (1) people perceive social distance and power relationships differently and (2) contrary to B&L's theory, it appears that these two variables should be weighted differently when calculating the face threat of a given

situation and hence the politeness strategy required. Lastly, we give our conclusions.

POLLy's Architecture

POLLy uses a shared representation for generating actions to be performed by the ECAs in the virtual domain on one hand and on the other, for generating dialogues to communicate about the actions to be performed. It consists of three components: A Virtual Environment (VE), a Spoken Language Generation (SLG) system and a Shared AI Planning Representation for VE and SLG as illustrated in Figure 2. A classic STRIPS-style planner called GraphPlan (Blum & Furst, 1997) produces, given a goal e.g. cook pasta, a plan of the steps involved in doing so (Gupta et al., 2007). POLLy then allocates this plan to the Embodied Conversational Agents (ECA) in the VE as a shared collaborative plan to achieve the cooking task with goals to communicate about the plan via speech acts (SAs), needed to accomplish the plan collaboratively, such as Requests, Offers, Informs, Acceptances and Rejections (Grosz,1990; Sidner, 1994; Walker, 1996). It also allocates this plan to the SLG component (Gupta et al., 2007) which generates variations of the dialogue based on B&L's theory of politeness that realizes this collaborative plan, as in (Andre et al, 2000; Walker et al, 1997). Some example dialogues are shown in the next section. In the VE, the human English language learner is able to interact with the Embodied Conversational Agent and plays the part of one of the agents in order to practice politeness real-time.

Brown and Levinson's theory

B&L's theory states that speakers in conversation attempt to realize their speech acts (SAs) to avoid threats to one another's face, which consists of two components. Positive face is the desire that at least some of the speaker's and hearer's goals and desires are shared by other speakers. Negative face is the want of a person that his action be unimpeded by others. Utterances that threaten the conversants' face are called Face Threatening Acts (FTAs). B&L predict a universal of language usage that the choice of linguistic form can be determined by the predicted Threat θ as a sum of 3 variables:

1. P: power that the hearer has over the speaker;
2. D: social distance between speaker & hearer;
3. R: a ranking of imposition of the speech act.

Linguistic strategy choice is made according to the value of the Threat θ. We follow Walker et al.'s (1997) four part classification of strategy choice. The Direct strategy is used when θ is low and executes the SA in the most direct, clear and unambiguous way. It is usually carried out either in urgent situations like "Please Help!", or where the face threat is small as in informing the hearer "I have chopped the vegetables" or if the speaker has power over the hearer, "Did you finish your homework today?". The Approval strategy (Positive Politeness) is used for the next level of

threat θ - this strategy is oriented towards the need for the hearer to maintain a positive self-image. Positive politeness is primarily based on how the speaker approaches the hearer, by treating him as a friend, a person whose wants and personality traits are liked, and by using friendly markers "Friend, would you please close the door?" or exaggerating "Amazing, you are the best cook in the world!" The Autonomy Strategy (Negative Politeness) is used for great face threats, when the speaker may be imposing on the hearer, intruding on their space or violating their freedom of action. These face threats can be mitigated by using hedges, "I wonder if you would mind closing the door for me," or by minimizing imposition, "I just want to ask you if you could close the door." The Indirect Strategy (Off Record) is the politest strategy and is therefore used when θ is greatest. It depends on speaking in an indirect way, with more than one attributable intention so that the speaker removes himself from any imposition. For example, using metaphor and irony, rhetorical questions, understatement, hints etc. "Its cold in here," which implies a request to close the door, or being vague like "Perhaps someone should clean the table." Some strategies produced by POLLy that were used in our experiment are given in Table 2.

Virtual Environment

We rendered POLLy with Microsoft Agent Characters (Microsoft, 1998) in our Virtual Environment RAVE at Sheffield University which consists of a 3-dimensional visualisation of computer-generated scenes onto a 10ft x 8ft screen and a complete 3D surround sound system driven by a dedicated computer. Figure 1 showed a user interacting with POLLy. The Microsoft Agent package provides libraries to program control using various developing environments like the .NET framework and visual studio and includes a voice recognizer and a text-to-speech engine. It also provides controls to embed predefined animations which make the characters' behaviour look more interesting and believable (Cassell & Thórisson, 1999). However, B&L theorize only linguistic politeness and not non verbal behaviour but as mentioned earlier, Rehm and Andre (2007) showed how politeness is expressed by means gestures and that gestures are used to strengthen the effect of verbal acts of politeness. We have embedded animations like gesturing in a direction, looking towards the other agents, blinking, tilting the head, extending arms to the side, raising eyebrows, looking up and down etc while the agents speak and listen to the utterances and holding the hand to the ear, extending the ear, turning the head left or right etc when the agents don't understand what the user says or the user doesn't speak anything. These animations not only make the characters look more believable but since they are chosen such that they don't portray any negative feelings, they conform to the linguistic politeness strategies as employed by the agents. The Agents share the AI plan to collaborate on it together to achieve the cooking task. Goals to communicate about the plan are also allocated to the agents

as speech acts (SAs) such as Requests, Offers, Informs, Acceptances and Rejections, needed to accomplish the plan collaboratively. While interacting with the system, the user sees one agent on the screen and plays the part of the second agent in one case and in another case sees two agents on the screen and plays the part of third. Since Microsoft Agents are 2D, they are not rendered 3D but a life size image of the characters is visible to the users on the screen to make them appear believable. When we extend this to a real-time immersive Virtual Reality environment, a Virtual Kitchen in this case, the ECAs will actually perform the task of cooking a recipe together in the virtual kitchen while conversing about the steps involved in doing so, as laid out by the AI plan.

Example Dialogues

Here are some example dialogues that illustrate the difference in the politeness strategies used in discourse contexts of varying power. Two Microsoft Agents, Merlin and Genie are involved in a conversation while cooking and cleaning in a kitchen together. Consider the difference in the degree of politeness in each situation.

The structure of the two dialogues shown here might seem somewhat different, this is just in order to create variation. We have selected exactly the same speech acts for both Merlin and Genie, which are only ordered differently in the two cases, and different tasks that carry equal amount of face threat for both. The ordering of speech acts and the difference in tasks in the two cases hence does not account for politeness.

Conversation 1: This is an example conversation in which Merlin is Genie's boss.

Agent (Speech act: Politeness strategy): Utterance
Merlin: Would you mind washing the dishes? (Approval: RequestAct)
Genie: Sure, I'll wash the dishes. (Direct: AcceptRequest)
Genie: I'm wondering whether I should boil the pasta.(Autonomy:Offer)
Merlin: Yeah you can. (Direct: AcceptOffer)
Merlin: You've burnt the vegetables. (Direct: Inform)
Genie: Yeah. (Direct: AcceptInform)
Genie: The oven is dirty. (Indirect: RequestAct)
Merlin: I'm sorry I cannot clean the oven. (Approval: RejectRequest)
Genie: Ok. (Direct: AcceptReject)
Genie: If you don't mind, can I clean the burner? (Autonomy: Offer)
Merlin: No, forget it. (Direct: RejectOffer)
Merlin: I must wipe the counter now. (Direct: Offer)
Genie: Alright, if you insist. (Autonomy: AcceptOffer)
Genie: Do you know that you've broken the baking dish mate? (Approval: Inform)
Merlin: Yes. (Direct: AcceptInform)
Merlin: I'll cook the vegetables. (Direct: Offer)
Genie: No please don't bother yourself. (Autonomy: RejectOffer)
Genie: The vegetables are healthy. (Indirect: RequestAct)
Merlin: Alright, I'll add the vegetables to the pasta. (Direct: AcceptRequest)
Merlin: Could you please add the cheese to the pasta for me? (Approval: RequestAct)
Genie: Please don't mind but I can't do that. (Autonomy: RejectRequest)
Merlin: Alright. (Direct: AcceptReject)

Conversation 2: In this case Genie and Merlin are colleagues at an equal level in hierarchy

Agent (Speech act: Politeness strategy): Utterance
Merlin: Could you possibly clean the oven for me?(Approval:RequestAct)
Genie: Sure. (Direct:AcceptRequest)
Genie: I'll clean the burner. (Direct:Offer)
Merlin: Ok. (Direct:AcceptOffer)
Merlin: You've burnt the vegetables. (Direct:Inform)
Genie: Yeah. (Direct:AcceptInform)
Genie: Would you mind washing the dishes? (Approval:RequestAct)
Merlin: I'm sorry but I can't wash the dishes. (Approval:RejectRequest)
Genie: Alright. (Direct:AcceptReject)
Genie: I must boil the pasta. (Direct:Offer)
Merlin: No, thanks. (Direct:RejectOffer)
Merlin: I can wipe the counter. (Direct:Offer)
Genie: Yeah you can. (Direct:AcceptOffer)
Genie: You've broken the baking dish. (Direct:Inform)
Merlin: Yes. (Direct:AcceptInform)
Merlin: I'll cook the vegetables. (Direct:Offer)
Genie: No, forget it. (Direct:RejectOffer)
Merlin: Could you please add the vegetables to the pasta? (Approval:RequestAct)
Genie: Please don't mind but I can't do that. (Approval:RejectRequest)
Merlin: Ok. (Direct:AcceptReject)
Genie: Will you please wipe the table mate? (Approval:RequestAct)
Merlin: Sure. (Direct:AcceptRequest)

Experiment

We conducted an experiment to study the perception of politeness by subjects in different discourse contexts, with 34 subjects from two different cultural backgrounds: 17 were Indians and 17 British, most of them being students of mixed gender with an age between 20 to 35 years. Subjects were administered a web-based questionnaire and presented with a series of tasks. They were asked to rate the various utterances as though they *had been said to them by their partner* in the collaborative task of cooking a recipe together. The subjects had to score how polite they perceived their partner to be, on a five point Likert-like scale: Excessively Overpolite, Very Polite, Just Right, Mildly Rude or Excessively Rude.

All of the tasks were selected to have relatively high R (ranking of imposition) as per B&L's theory. Requests were to 'chop the onions', 'wash the dishes', 'take out the rubbish' and 'clean the spill on the floor.' The events for the propositional content of the Inform SAs were "You have burnt the pasta" and "The milk is spoilt", "You have broken the dish" and "The oven is not working". The subjects rated a total of 84 sentences spread across these eight different tasks as shown in Table 1. There was also a text box for subjects to write optional comments.

Hypothesis. The subjects were told that the discourse situation was cooking in a cooking class with their Junior Colleague from work in one case and with a Stranger in

			B&L Strategies					
Speech Act	**Situation**	**Tasks**	**Direct**	**Approval**	**Autonomy**	**Indirect**	**Total**	
Request	**Friend**	chop onions	4	4	4	4	16	
		clean spill on floor	4	4	4	4	16	64
	Stranger	wash dishes	4	4	4	4	16	
		take out rubbish	4	4	4	4	16	
Inform	**Friend**	oven not working	1	2	2	0	5	
		burnt the pasta	1	2	2	0	5	20
	Stranger	milk is spoilt	1	2	2	0	5	
		broken the dish	1	2	2	0	5	

Table 1: Distribution of the dialogue utterances in the experiment

another. The evaluating user who is the hearer is the **Boss** in one case and a **Stranger** in the other. This was in order to evaluate the weight of B&L's P variable representing the Power that the hearer has over the speaker vs. the D variable representing social distance. On the continuum of social distance, while a friend lies on one end, a stranger lies on the other, and on the continuum of power, a junior colleague lies on one end and a boss on the other. It is unclear though whether a stranger and a boss contribute the same amount to the face threat θ, according to B&L's theory, and whether a friend and a junior colleague make a similar contribution at the lower end of the scale. Our hypothesis was that power should have relatively more weight than social distance because it seems plausible that when the hearer has power, the speaker would take more care in selecting a politeness form due to fear that if the

B&L	**Request Speech Act**		**Inform Speech Act**	
	Strategy Forms	*Strategy Names*	*Strategy Forms*	*Strategy Names*
Direct	Do X.	RD1Imperative	X	ID1DirectAssert
	Please Do X.	RD2ImperativePlz	-	-
	You must do X.	RD3ImperativeInsist	-	-
	You could do X.	RD4AsModAbility	-	-
Approval	Could you please do X mate?	RAp1QModAbility	Do you know that X?	IAp1QKnowledge
	If you don't mind you can do X.	RAp2AsModAbility	Do you know that X mate?	IAp2QueryKNowledgeAddress
	Would it be possible for you to do X?	RAp3AsPossible	-	-
	Let's clean the spill on the floor.	RAp4AsSH	-	-
Autonomy	Could I just ask you if you could possibly do X?	RAu1QModAbilityMinImp	It seems that X.	IAu2AsAppear
	I'm really sorry to bother you but could you please do X?	RAu2ApologizeQModAbility	I am wondering if you know that X.	IAu1AsConfuse
	I'm wondering whether it would be possible for you to do X.	RAu3AsConfusePossibility	-	-
	It would be a big favour if you could please do X.	RAu4FavourAsModAbility	-	-
Indirect	X has not been done yet	RI1AsNegation	-	-
	X should have been done.	RI2AsModRight	-	-
	Someone should have done X.	RI3AsModRightAbSub	-	-
	No one has done X yet	RI4AsNegationAbsSub	-	-
	Where X is a task request. For example 'You could chop the onions,' or 'Would it be possible for you to clean the spill on the floor?'	*These strategies are applied to the various tasks requests X.*	*Where X is an inform event, like 'Do you know that the milk is spoilt mate?' or 'It seems that that you have burnt the pasta.'*	*These strategies are applied to the various inform events X.*

Table 2: The individual B&L strategies used for Request and Inform speech act

hearer does not like what was said or gets upset, he/she may exercise his/her power and the speaker may have to bear the consequences. On the other hand, interactions with a stranger are controlled by social norms, but the speaker would have no such fear. Thus we can afford to be impolite to strangers, but not to our boss! The speech acts tested were: **Request** and **Inform**. The ranking of imposition R for speech acts has Requests with higher R than Inform, so θ should be greater for requests, implying the use of a more polite B&L strategy. For the Request speech act, each subject judged 32 example utterances and for Inform speech act, 10 example utterances. The distribution of these utterances used in the experiment is given in Table 1. No Indirect strategies were used for Inform SAs because those given by B&L of hints, being vague, jokes, tautologies are not implemented in our system. The B&L strategies for Requests and Informs are in Table 2.

Results and Observations

We calculated an ANOVA with B&L category, situation (friend/stranger), speech act, syntactic form, politeness formula and the nationality of subjects as the independent variables and the ratings of the perception of politeness by the subjects as the dependent variable. Some mean values are given in Tables 3 and 4.

Quantitative Observation. It was observed that though the overall measure of politeness was almost the same in both the cases (df=1, F=0.92, p=0.3), when the user was a boss and when the user was a stranger, Indians rated the overall sentences in case the speaker was a junior colleague as slightly less polite as compared to strangers (df=1, F=2.8, p<0.1) as shown in Table 3 which shows that they expect more politeness from a junior colleague as compared to a stranger. The politeness measures by British subjects were same in both cases (df=1, F=0.2, p=0.7). However, upon a closer examination of the quantitative and the qualitative results at the level of the individual ratings of the subjects, we observed that the interpretation was very subjective. Out of the 17 British subjects, 6 marked the utterances by a stranger as overall more polite, 6 marked the utterances by a junior colleague as more polite and 5 rated both as equal. Out of the 17 Indian subjects, 10 rated the utterances by a stranger to be overall more polite, 6 rated utterances by junior colleague as more polite and 1 rated both as equal. Also, the average rating at the individual level for both the cases varied considerably, suggesting that individual appraisals of the social situation people are very subjective.

Nationality	Junior Colleague	Stranger
Indian	2.7	2.8
British	2.8	2.8
Overall	2.7	2.8

Table 3: Nationality vs. context

B&L Strategies	Junior Colleague	Stranger
Direct	2.0	2.0
Approval	2.8	3.0
Autonomy	3.7	3.6
Indirect	2.0	2.0
Inform SA		
B&L Strategies	**Junior Colleague**	**Stranger**
Direct	2.3	2.4
Approval	2.4	2.7
Autonomy	2.9	3.0
Indirect	-	-
Request SA		
B&L Strategies	**Junior Colleague**	**Stranger**
Direct	2.0	2.0
Approval	3.0	3.0
Autonomy	4.0	4.0
Indirect	2.0	2.0

Table 4: Effect of B&L strategies

Qualitative Observation. Analyzing the quantitative nature of the results, we observed that people gave wide comments about different aspects. For power vs. social distance relationship, one person said that "there is a difference in the two situations where two strangers are talking and two people in a power relationship are talking. When a junior colleague is talking to a senior, he will always be just right or very polite or may even be excessively polite in some cases. But in case of strangers, you may be rude or even very rude sometimes… it depends a lot on the mood in case of strangers. You may be rude to a stranger if you just had a fight with someone but no matter how angry you may be, you will never be rude to your boss, at the most you will be just right, if not very polite." Another said that "while interacting with a junior colleague, I will expect a level of respect from him and hence I will feel that he is being rude even if he is making general statements. On the other hand when I am interacting with a stranger, I will not have any expectation from him and would not feel that he is being rude if he is speaking in the same way as my junior colleague. It will work the same the other way round, when my junior colleague says something politely, it is what is I am expecting and it will seem to be normal for me, whereas the same statement from the stranger will seem to be excessively polite or very polite." One person brought in the social distance to the power relationship and said "I would expect some sort of respect from both a junior colleague as well from a stranger. However, the addressing statements from either of them in a same situation should be different considering the fact that one of them is known to me and the other person is a stranger."

Subjects also said that things like mood, the basic nature of a person also come into play and tone is another very important factor which can make the otherwise seemingly polite sentence impolite and seemingly impolite sentence polite. Also, excessively overpolite sentences may be interpreted as sarcastic and therefore rude.

Conclusion

Our results suggest that in B&L's equation ($\theta = P + D + R$) the weight of the P and D variables is not the same. The weights appear to be subjectively determined, with a higher weight for P in some cases and higher for D in others and in some cases it can also be equal. Where there is large social distance and power, the politeness required is high and on the other hand where the hearer has a lesser social distance with the speaker despite having power, the amount of politeness expected is reduced. B&L implicitly state the equality of these two variables whereas we see now that not only do these variables have different weight, they are also not independent. The amount of power calculated also depends upon the social distance. Watts et al. (Watts, Ide & Ehlich, 1992) state that the amount of ranking of imposition of an FTA is determined by the Social Distance and Power variables. For instance a request act carried out before a close friend would differ from that before a person with a very high power. Considering this, the weighting of the Rx also becomes subjective in a similar manner. (Walker et al., 1997) state that Rx should be a function of both the speech act type and the type of action in the domain, i.e., the context in which an utterance has been used should also be considered while calculating the Rx. Obviously a request act speech act for passing salt would be lesser of an imposition than asking for twenty pounds.

We also report observations regarding the effect of the different B&L's strategies, the linguistic form of realization of the politeness strategies and the speech act type. We observed that (1) politeness perceptions of POLLy's output are generally consistent with B&L's predictions for choice of form for discourse situation, i.e. utterances to strangers or a superior person need to be very polite, preferably Autonomy oriented, (2) our indirect strategies which should be the politest forms, are seen as the rudest. The overall politeness ratings from least polite to most were Indirect, Direct, Approval and then Autonomy; and (3) English and Indian speakers of English have different perceptions of politeness. These results are consistent with those reported in (Gupta et al, 2007), a similar experiment to study the intra variable difference of D, the social distance variable. There we compared the difference in the perception of politeness in two situations, where in one situation a friend was speaking to the user and in another, a stranger was speaking to the user.

Here we tested an offline version of POLLy to calibrate our language generation before incorporating it in our virtual world. In future work, we plan to evaluate complete dialogue interactions in the VE, where the user is immersed in our virtual world, and where the agents actions are carried out in that world. We hope to also be able to show that users like the immersive system and that their capabilities to choose the socially correct form of politeness in a new language improves as a result of their experience in this world. We also plan to explore the use of more expressive agents who are capable of producing socially appropriate gestures and facial expressions, such as the Greta agent developed by Pelachaud et al (Poggi et al, 2005) and used in an evaluation of gesture and politeness as reported in Rehm and Andre (2007).

References

André, E., Rist, T., Mulken, S.v., Klesen, M., & Baldes, S. 2000. *The automated design of believable dialogues for animated presentation teams.* In Embodied Conversational Agents (pp. 220–255). Cambridge, MA, USA: MIT Press.

André, E., Rist, T., Mulken, S.v., Klesen, M., & Baldes, S. 2000. *The automated design of believable dialogues for animated presentation teams.* In Embodied Conversational Agents (pp. 220–255). Cambridge, MA, USA: MIT Press.

Bersot, O., El-Guedj, P.O., God´ereaux, C. and Nugues. P. 1998. *A conversational agent to help navigation & collaboration in virtual worlds.* Virtual Reality,3(1):71–82.

Blum, A., Furst, M. 1997 *Fast Planning Through Planning Graph Analysis.* Artificial Intelligence 90, 281–300.

Cassell, J. and Thórisson, K.R. 1999. *The Power of a Nod and a Glance: Envelope vs. Emotional Feedback in Animated Conversational Agents.* Applied Artificial Intelligence 13: 519-538.

Cassell, J. Bickmore, Timothy W. Negotiated Collusion. 2003. *Modeling Social Language and its Relationship Effects in Intelligent Agents.* User Model. User-Adapt.Interact. 13(1-2):89-132.

Culpeper, J. 1996. *(Im)politeness in dramatic dialogue.* Exploring the Language of Drama: From text to context. Routledge, London.

Dautenhahn, K. 2000. *Socially Intelligent Agents: The Human in the Loop* (Papers from the 2000 AAAI Fall Symposium). The AAAI Press, Technical Report.

Eichner, T., Prendinger, H., André, E. and Ishizuka, M. 2007. *Attentive presentation agents.* Proc. 7th International Conference on Intelligent Virtual Agents (IVA-07), Springer LNCS 4722. pp 283-295.

Grosz, B.J., Sidner, C.L. 1990. *Plans for discourse.* In: Cohen, P.R., Morgan, J.L., Pollack, M.E. (eds.) Intentions in Communication, MIT Press, Cambridge, MA.

Gupta, S., Walker, M.A., Romano, D.M. 2007. *How Rude are You?: Evaluating Politeness and Affect in Interaction.* Affective Computing & Intelligent Interaction (ACII-2007).

Johnson, L.W. and Rizzo, P. and Bosma, W.E. and Ghijsen, M. and van Welbergen, H. 2004. *Generating socially appropriate tutorial dialog.* In: ISCA Workshop on Affective Dialogue Systems. pp. 254-264.

Johnson, L., Mayer, R., André, E., & Rehm, M. 2005. *Cross-cultural evaluation of politeness in tactics for pedagogical agents.* Proc. of the 12th Int. Conf. on Artificial Intelligence in Education.

Kopp, S., Jung, B., Lessmann, N. and Wachsmuth, I. 2003. *Max – A multimodal assistant in virtual reality construction.* KI Zeitschift (German Magazine of Artificial Intelligence), Special Issue on Embodied Conversational Agents, vol.4, pp.11–17.

Microsoft. 1998. *Developing for Microsoft Agent.* Microsoft Press.

op den Akker, H.J.A. and Nijholt, A. 2000. *Dialogues for Embodied Agents in Virtual Environments.* In: Natural Language Processing - NLP 2000, 2nd Int. Conf. pp. 358-369. LNAI 1835.

Paiva, A., Dias, J., & Aylett, R.S. 2005. *Learning by feeling: evoking empathy with synthetic characters.* Applied Artificial Intelligence: 19 (3-4), 235-266.

Poggi, I., Pelachaud, C., de Rosis, F., Carofiglio, V., De Carolis, B. 2005. *GRETA. A Believable Embodied Conversational Agent.* in O. Stock and M. Zancarano, eds, Multimodal Intelligent Information Presentation, Kluwer.

Prendinger, Helmut and Ishizuka, Mitsuru. 2001. *Let's talk! Socially intelligent agents for language conversation training.* IEEE Transactions on xSystems, Man, and Cybernetics - Part A: Systems and Humans, Vol. 31, No. 5, pp 465-471.

Porayska-Pomsta, K. 2003. *Influence of Situational Context on Language Production: Modelling Teachers' Corrective Responses.* PhD Thesis. School of Informatics, University of Edinburgh.

Rehm, M. and Andre, E. 2007. *Informing the Design of Agents by Corpus Analysis.* Conversational Informatics, Edited by T. Nishida.

Sidner, C.L. 1994. *An artificial discourse language for collaborative negotiation.* In: Proc. 12th National Conf. on AI, pp. 814–819.

Takenobu, T., Tomofumi, K., Suguru, S., Manabu, O. 2003. *Bridging the Gap between Language and Action.* IVA 2003, LNAI 2792, pp. 127-135.

Traum, D., Rickel, J., Gratch, J., Marsella, S. 2003. *Negotiation over Tasks in Hybrid Human-Agent Teams for Simulation-Based Training.* Proceedings of the 2nd Int. Joint Conf. on Autonomous Agents and Multiagent Systems.

Wahlster, W., Andre, E., Finkler, W., Profitlich, J. J. and Rist, T. 1993. *Plan-based Integration of Natural Language and Graphics Generation',* Artificial Intelligence, 63(1-2):387--428.

Walker, M.A. 1996. *The effect of resource limits and task complexity on collaborative planning in dialogue.* Artificial Intelligence Journal 85, 1–2.

Walker, M., Cahn, J. and Whittaker, S. J. 1997. *Improving linguistic style: Social and affective bases for agent personality.* In Proc. Autonomous Agents'97. 96–105. ACM Press.

Wang, N., Johnson, W.L., Rizzo, P., Shaw,E., & Mayer, R. 2005. *Experimental evaluation of polite interaction tactics for pedagogical agents.* Proceedings of IUI '05. ACM Press.

Watts, Richard J. Ide, S. and Ehlich, K. 1992. Introduction, in Watts, R, Ide, S. and Ehlich, K. (eds.), *Politeness in Language: Studies in History, Theory and Practice.* Berlin: Mouton de Gruyter, pp.1-17.