A Computational Model of Social Perlocutions

David Pautler and Alex Quilici
University of Hawaii at Manoa
Department of Electrical Engineering
2540 Dole St. Holmes 483
Honolulu, HI 96822

Abstract

The view that communication is a form of action serving a variety of specific functions has had a tremendous impact on the philosophy of language and on computational linguistics. Yet, this mode of analysis has been applied to only a narrow range of exchanges (e.g. those whose primary purpose is transferring information or coordinating tasks) while exchanges meant to manage interpersonal relationships, maintain "face", or simply to convey thanks, sympathy, and so on have been largely ignored. We present a model of such "social perlocutions" that integrates previous work in natural language generation, social psychology, and communication studies. This model has been implemented in a system that generates socially appropriate e-mail in response to user-specified communicative goals.

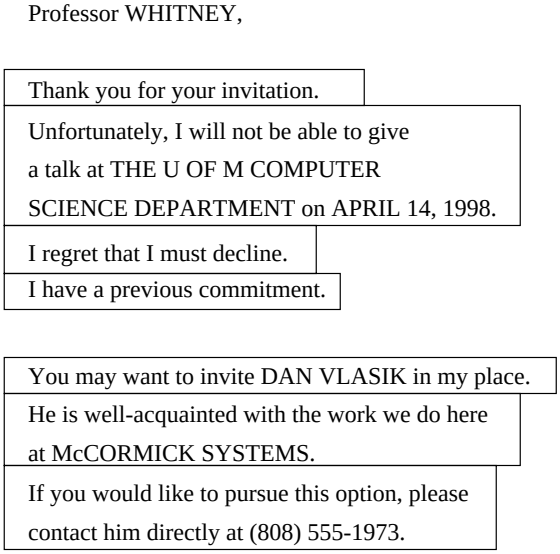

Figure 1: A LetterGen Output Sample

1 Introduction

The importance of viewing utterances as not simply statements of fact but also as real actions (speech acts) with consequences has long been well understood (Searle, 1969; Austin 1975; Grice 1975). As a result, it is important to study not just the formal aspects of language forms but also how speakers use different forms to serve different functions. For example, one function of the act of informing another person is to make the person aware of a state of affairs; similarly, one function of promising is to secure the return of a favor.

Unfortunately, the study of speech acts has been largely limited to the collection and classification of act types and the conditions for appropriate use of each type (Searle 1969; Wierzbicka 1987). The range of functions, or *perlocutionary effects*, served by different act types has been largely ignored. In particular, there has been little or no work on the impact that speech acts can have on social attitudes and behavior. Yet, without an account of how communication can affect social situations, it is impossible to construct systems that are capable of generating socially appropriate text.

This paper provides a computational model of *social perlocutions*, and it describes how this model has been used to construct an automated system, *LetterGen*, for generating socially appropriate e-mail messages and letters. This system takes general communicative and social goals from the user, such as demanding action or expressing congratulations, queries the user about subgoals and pertinent background information, and generates the text of an appropriate message by planning individual speech acts.

As an example, Figure 1 shows a message generated by LetterGen in response to an input goal to decline an invitation politely. In this example, the writer was invited by the addressee to travel and give a talk, but the writer had a previous commitment and must decline. However, the writer knows some-

one who could give the talk in his place. The system planned a set of speech acts and realized each as a clause or phrase using a text template library. These acts include (1) thanking, (2) declining-request, (3) apologizing, (4) making-excuse, (5) advising, (6) assuring, and (7) requesting.

Most of the text in the letter is devoted to addressing the writer's social goals of being polite and helpful. In contrast, a letter writer concerned only with informing the addresee that he was not participating would likely say little other than "I won't be giving a talk at your event", a socially inappropriate response.

2 Previous Research

Our work builds on results from three disparate areas: natural language generation (NLG), communication studies, and social psychology.

The NLG community has focused on a small subset of the five generally accepted categories of speech acts (Levinson, 1983):

1. Representatives—statements given as true depictions of the world (e.g., asserting, concluding).
2. Directives—statements attempting to persuade the hearer to do something (e.g., ordering, advising, warning).
3. Commissives—statements that commit the speaker to a course of action (e.g., promising, accepting a request, taking a side).
4. Expressives—statements expressing a psychological state (e.g., apologizing, congratulating, condoling).
5. Declarations—statements effecting an immediate change in the institutional state of affairs (e.g., christening, firing from employment).

In particular, research in NLG has been limited to one type of representative (i.e., informing) and one type of directive (i.e., requesting), and it has further focused on informing's potential to *convince the hearer* of some fact and requesting's potential to *persuade the hearer* to do some action (Allen et al., 1994; Appelt, 1985; Bruce, 1975; Cohen and Perrault, 1979; Hovy, 1988; Perrault and Allen, 1979). As a result, it has largely ignored speech acts in other categories, such as promising, advising, and crediting, as well as their potential perlocutionary effects of creating affinity between speaker and hearer, securing future favors for the speaker, and so on.

In contrast, research in communication studies has explored strategies for persuading, creating affinity, comforting, and many other interpersonal goals (Daly and Wiemann, 1994; Marcu, 1997). For example, the strategies for persuading include not only requesting, but also exchange, ingratiation, and sanctions. However, these efforts have not analyzed these strategies in terms of speech act types and perlocutionary effects so that these strategies might be realized in computational form.

Finally, research in social psychology has looked at how personality traits affect interpersonal interaction. For example, Kiesler (1983) formulated general rules for describing how one person expressing one trait (e.g., merciful) can lead to another person expressing a symmetric and complementary trait (e.g., appreciative). Such interaction dyads are directly mappable to the speaker/hearer dyad of speech act theory, and the vocabulary of trait terms and predictive rules suggest one way of lending organization to the great variety of perlocutionary effects. Yet, social psychologists have not mapped their general trait terms to the classes of speech acts that might express these traits.

What's been lacking is an attempt to integrate the lessons learned from these different research efforts to provide an initial model of social perlocutions; that is, a model that describes how specific speech act types have the potential to produce specific effects in a hearer corresponding to a speaker's social goals, and that is specified formally enough to be used as part of text generation systems.

3 Our Model

There are two key questions to address in forming a computational model of social perlocutions:

- What are the possible socially-relevant effects of speech acts?
- What are the relationships between different effects?

3.1 Social Perlocutionary Effects

We have developed a taxonomy of social perlocutionary effects of speech acts. These effects are defined in terms of mental attitudes of the hearer, following the assumption in speech act theory that all perlocutionary effects follow from the hearer's recognition of the speaker's communicative intent. The taxonomy is:

1. Beliefs about speaker's precise communicative content and communicative intent.

2. Beliefs about the speaker's intent to benefit or harm the hearer.

3. Beliefs about the hearer's or speaker's responsibilities (ascribed or undertaken).

4. Beliefs about (or, impressions of) the speaker's personality traits.

5. The hearer's emotions.

6. The relationship between the hearer and the speaker.

7. The hearer's goals.

We developed this taxonomy by reviewing the communications studies and social psychology literature, as well as by analyzing a corpus of letters and e-mail messages for their speech acts and most prominent social effects. Prior research on speech acts has largely ignored several of these categories, especially the effects on personality impressions, emotions, and the speaker-hearer relationship.

3.2 Relationship Between Social Effects

This taxonomy is important because there appear to be significant restrictions on the relationships between these different classes of effects.

Figure 2 shows how these different types of effects are related. The arrows represent potential causal links between effects. These links are *potential* because there are specific conditions associated with specific effects that dictate whether one effect will cause another.

Essentially, the effects start with the hearer's recognition and acceptance of a message's content and culminates in changes to hearer goals and the relationship between the hearer and the speaker. That is, a speech act directly results in beliefs about the content and intent of utterances and these beliefs indirectly result in changes to goals, emotions, and interpersonal relationships. Specficially, these belief can lead to indirect changes in the hearer's belief about the speaker's intent to benefit or harm the hearer, as well as changes to the hearer's responsibilities that involve the speaker. In turn, changes in belief about whether the speaker intends to benefit or harm the header can lead to changes in the hearer's goals, the hearer's emotions, and the hearer's impressions of the speaker's personality traits. Finally,

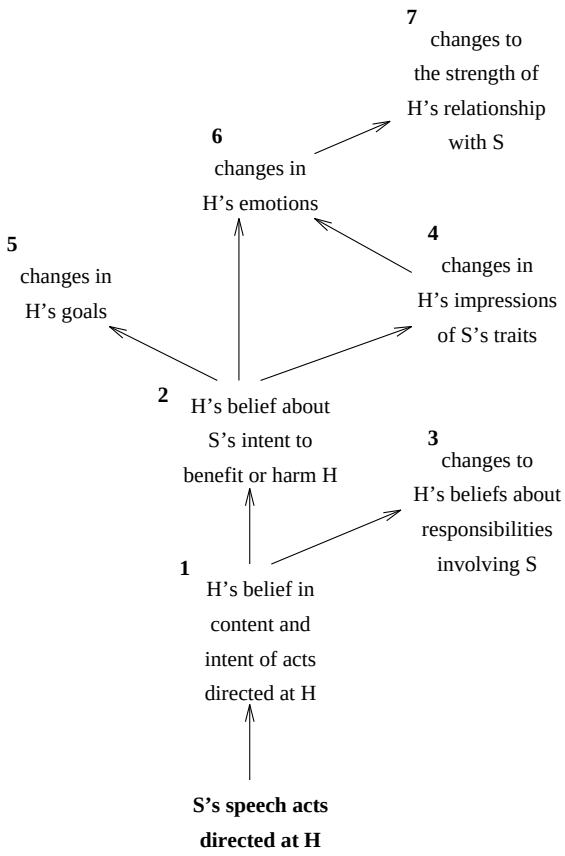

Figure 2: The Relationships Between Social Effects

changes to the hearer's emotions can lead to changes in the hearer's relationship with the speaker.

Our hypothesis is that Figure 2 provides a framework into all speech acts with social effects can be mapped. To test this hypothesis, we analyzed in detail the relationship between the effects of 40 different types of speech acts, and we successfully placed each into this framework (Pautler, 1999). These speech acts were typical of the letters and messages we collected, and they were representative of four of the five main categories of speech acts.[1]

Figure 3 is an example, showing these effects for *apologizing*.[2] Although not shown in Figure 3, the causal relationships between these effects have conditions attached to them. In Figure 3, for example, a condition on an apology leading to the hearer believing the speaker feels regret is that the hearer believes the speaker is sincere and there is an act for which

[1] We did not represent declarations because we chose to focus on acts used in casual, interpersonal interactions rather than acts that were institutionally framed.

[2] We do not claim that the model applies to groups other than adult Westerners. See Barnlund (1989) for comparisons on the use of different speech acts by Americans and Japanese.

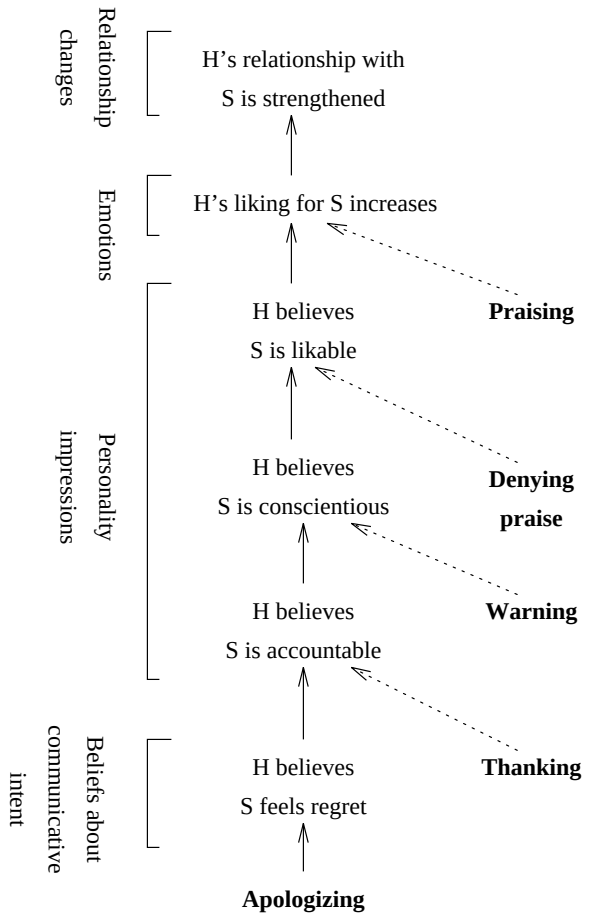

Figure 3: The Effects Of Apologizing

an apology is appropriate.

We draw our terminology for describing specific personality traits (e.g., likeable, conscientious) and emotions (e.g., gratitude, liking) from existing taxonomies (Kiesler, 1983; Ortony et al., 1988).

Figure 3 shows effects with arrows leading to them from other speech acts, such as praising, warning, thanking, and so on. These speech acts are there to illustrate that speech acts are related through a web of interlocking effects. That is, the causal relationships between speech acts and effects is many-to-many: a single act can have many different effects and any single effect can be brought about by many different acts. For example, expressing a demand can bring about compliance, anger, or both, and similarly, anger can be caused by a variety of other acts, such as issuing a threat. In Figure 3, both praising and apologizing are examples of acts that can increase the hearer's liking for the speaker, and both apologizing and thanking can lead the hearer to believe the speaker is accountable.

This large web of relationships between the effects of social speech acts leads to the question: How can we efficiently generate the speech acts we need to achieve an appropriate emotional response in the hearer?

4 A Model Of Letter Generation

To illustrate the power of our model of social perlocution, we have applied it to the task of e-mail generation in a system called LetterGen. The system's primary task is to take a high-level communicative goal (e.g., inform a colleague that one can't attend a meeting) and suggest a set of speech acts to achieve that goal. However, once it has made this suggestion, the system then interact with the user to determine which speech acts will appear in the final message and to acquire any additional background information needed to instantiate sentence text templates associated with each speech act.

In addition to the user's explicit input goal, the system works with a set of "standard" user goals. These goals fall into three classes:

1. Cost avoidance—avoiding undesired aspects of a current or incipient situation, such as unwanted social perceptions of oneself.

2. Status-quo maintenance—selection of an act because one of its effects would reinforce a desired aspect of the current situation (e..g, offering to help another person because it would reinforce one's self-image as a generous person).

3. Trait-based habit—performing of an act as a timeworn expression of a personality trait.

These goals can be thought as a stereotypical model of the user (Chin, 1989). These goals are achieved opportunistically during the process of determining speech acts for the explicitly provided user goal.

4.1 A Graph-Based Representation Of Speech Act Relationships

LetterGen essentially represents the perlocutionary effects of speech acts as a large graph. Figure 4 illustrates a portion of this representation that relates the speech acts of declining, thanking, and apologizing. The nodes of the graph represent various effects, and the unlabled edges represent a causal relationships between two effects. There are also constraints on when edges can be traversed (although hey are

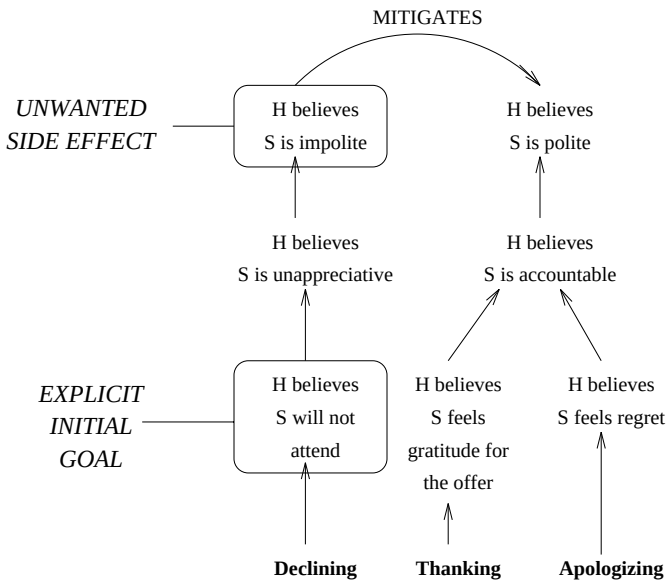

Figure 4: A representation for Declining, Thanking, and Apologizing.

not shown in this figure). Finally, there are *mitigates* links between nodes when two effects are contradictory.

A reasonable view of LetterGen's approach is that there is a *script* associated with each speech act that captures the causal chain of effects that potentially follow from it. While these effects could presumably be determined by reasoning from first principles, these scripts can be viewed as *standard methods* of achieving communicative goals, and they are essentially equivalent to the communicative strategies proposed by others (Marco, 1997).

4.2 Determining Appropriate Speech Acts

LetterGen's algorithm for producing a response involves 5 steps:

1. Metch the user's goal to one of the nodes (effects) in the graph.

2. From the matching effect, traverse graph links "downward" toward the speech act, checking the conditions on each link.

3. For every path that reaches an act by satisfying all conditions along the path, add the act to the new message by instantiating the act's text template.

4. Detect undesirable side effects of each added speech act by traversing all links back "upward" as far as possible.

5. If an effect is indexed by a mitigates link, follow the link to the mitigating effect in the other chain. Continue with steps 2 and 3.

As an example, consider the user's communicative goal to make the hearer believe that the speaker will not attend. Lettergen traverses the graph downwards to locate the speech act Declining. After determining this speech act, LetterGen then traverses the graph upward, moving through its effects, verifying that none of them conflict with known speaker goals. In this case, one of the effects of Declining conflicts with the speaker's goal that the hearer believes the speaker is polite. At this point, LetterGen generates a new goal to mitigate that effect, and recursively uses its algorithm to locate speech acts to achieve that goal. With the failed goal of being perceived as polite, LetterGen's downward traversal locates Thanking and Apologizing as appropriate speech acts to mitigate that failure.

4.3 An Alternative To Planning

This approach can be viewed as a form of reactive planning. LetterGen can be viewed as having a simple goal (communicate a particular belief to the hearer), forming a plan (finding a set of speech acts that communicate this belief), analyzing the effects of the plan (looking for user goals that are violated by these speech acts), and opportunistically pursuing new goals (to mitigate these violations).

LetterGen differs significantly from most other efforts in planning speech acts. These efforts explicitly represent speech acts and their effects as plan operators and attempt to synthesize sequences of operators. Unfortunately, as others have pointed out (Cohen and Levesque, 1980; 1990), plan operators are not a good representation when acts have long chains of effects. That's because each chain that results from a given act must be conflated to a flat list of effects, or each effect must be re-envisioned as an act, with one operator for each effect and appropriate preconditions so the operators can form the appropriate chain.

LetterGen's approach is most similar to the alternative to planning for speech-modeling proposed by Cohen and Levesque (1980, 1990). Their approach uses a set of inference rules and act type definitions and is explicitly designed to capture sequences of this type,

```
          c1      c2            ci
A(d) ---> E1 ---> E2 ---...---> Ei
```

where A(d) is an act that communicates propositional content d (definitional content for some act type), which induces effect E1 under conditions c1, which induces effect E2 under conditions c2, and so on.

This rule formalism is directly mappable to the conditionalized causal relations used in our social perlocutions model, with two exceptions. One is that we capture the rules with an annotated graph structure that makes the connectivity among rules explicit (scripts). The other provide a specialized graph-traversal algorithm that takes advantage of key properties of the graph, which allows us to substitute efficient graph traversal for generalized planning.

5 Implementation

The current implementation contains a very detailed model of speech act effects, containing over 400 effects and constraints. It is able to generate a dozen different types of messages, including initiating or terminating a friendship, applying or resigning from a job, congratulating or consoling someone, accepting or declining an invitation, encouraging or discouraging someone from doing an act, thanking someone, and apologizing to someone. Each of these different message types includes an organizational template that places generated acts in an appropriate order for the task.

An important part of LetterGen is its interaction with the user. Given a selected message type, LetterGen suggests at least three speech acts for the user to choose from. For example, the thanking message type (i.e., make them believe you feel gratitude) can be instantiated crediting (distributing credit), offering (to repay), as well as an overt expression of gratitude (i.e., thanking). For each act chosen by the user, the system queries the user for the background information needed to instantiate an appropriate text template.

6 Limitations and Future Work

The model currently has three major limitations. First, it does not cover all aspects of social interactions. For example, it does not have conditions or effects involving the relative status of the speaker and hearer, or specialized roles they might play (e.g., judge, employer, and so on). Second, the conditions on exactly when effects occur need to be elaborated significantly. Finally, there are socially-related speech acts we have not yet represented (e.g., expressing sadness, joy, and so on).

The primary implementation limitation involves the background information required to determine whether various conditions hold. Currently, the implementation does not query the user for all the background information it could take advantage of. The reason is that too many queries makes the program loses its appeal as a work-saving device. A related limitation is that its model of the speaker's goals is static, rather than dynamic (e.g., the speaker is always assumed to have a goal of being polite). We are addressing both of these problems by exploring techniques for forming a detailed user profile and applying across a large set of generated letters. The other important limitation is that its organizational and text templates are not particularly flexible (e.g., they demand a specific speech act order and they realize each speech act as a single sentence). One way to address this problem is to take the set of speech acts that LetterGen wants to generate as a goal and to plan exactly how they will be realized (Hovy, 1993; Moore and Paris, 1994; Hobbs, 1982).

One interesting area for future exploration is the problem of applying the model to letter understanding as well as generation. This problem is potentially difficult, as there are a variety of social reasons why a particular speech act might have appeared. For example, the thanking act might have been included in the example of Figure 1 in order to lessen the social debt the invitee owes to the inviter, or to avoid insulting the inviter through curtness, or to make the invitee feel that he is a polite person, or simply out of habit.

7 Conclusions

This paper has presented a computational model of the social perlocutionary effects of speech acts. Our model extends previous formal models of speech acts to take into account effects involving emotions, impressions, and the interpersonal relationship between the speaker and the hearer. In doing so, we have integrated earlier results from natural language generation on speech acts, from communication studies on communication strategies, and from social psychology on how interactions affect personality traits.

We have used this model to construct a prototype program that generates letters that meet social goals. This task is a key aspect of any general-purpose, intelligent, personal assistant that is in-

volved in mediating interpersonal interaction.

References

[1] J. F. Allen, L. K. Schubert, G. Ferguson, P. Heeman, C. H. Hwang, T. Kato, M. Li ght, N. G. Martin, B. W. Miller, M. Poesio, and D. R. Traum. 1994. *The TRAINS project: A case study in building a conversational planning agen t.* Technical Report 532, Computer Science Department, University of Rochester, Rochester, NY.

[2] D. Appelt. 1985. *Planning English sentences.* New York: Cambridge Universit y Press.

[3] J.L. Austin. 1975. *How to do things with words.* Cambridge, MA:Harvard University Press.

[4] D. Barnlund. 1989. *Communicate styles of Japanese and Americans: Images and Realities.* Belmont, CA: Wadsworth Publishing.

[5] B. Bruce. 1975. *Belief systems and language understanding.* Technical Repor t 2973. Cambridge, MA: Bolt, Beranek, and Newman.

[6] D. Chin. 1989. KNOME: Modeling What The User Knows In UC. In, A. Kobsa and W. Wahlster (eds): *User Modeling in Dialog Systems.* Berlin, Heidelberg: Spring-Verlag, pages 74-107.

[7] P. R. Cohen and H. J. Levesque. 1990. *Rational interaction as a basis for communication.* In Cohen, Morgan, and Pollack 1990.

[8] P. R. Cohen and C. R. Perrault. 1979. Elements of a plan-based theory of speech acts. *Cognitive Science* 3. 177-212.

[9] P. R. Cohen, J. Morgan, and M. E. Pollack (eds). 1990. *Intentions in communication.* Cambridge, MA: MIT Press.

[10] J. A. Daly and J. M. Wiemann (eds). 1994. *Strategic interpersonal communication.* Hillsdale, NJ: Lawrence Erlbaum Associates.

[11] H. P. Grice. 1975. Logic and conversation. In *Syntax and semantics III: Speech acts*, (eds) P. Cole and J. L. Morgan. New York: Academic Press.

[12] J. R. Hobbs. 1982. Towards an understanding of coherence in discourse. In *S trategies for natural language processing*, (eds.) W. Lehnert and M. Ringle. Hil lsdale, NJ: Lawrence Erlbaum Associates.

[13] E. H. Hovy. 1988. *Generating natural language under pragmatic constraints.* Hillsdale, NJ: Lawrence Erlbaum Associates.

[14] E. H. Hovy. 1993. Automated discourse generation using discourse structure relat ions. *Artificial Intelligence* 63. 341-385.

[15] D. Kiesler. 1983. The 1982 interpersonal circle: A taxonomy for complementarity in human transactions. *Psychological Review* (90) 3. 185-214.

[16] J. D. Moore and C. L. Paris. 1994. Planning text for advisory dialogues: Capturing intentional and rhetorical information. *Computational Linguistics* (19) 4. 651-694.

[17] A. Ortony, G. Clore, and A. Collins. 1988. *The cognitive structure of emotions.* New York: Cambridge University Press.

[18] D. Pautler, 1999. *A Computer Model of Strategic Aspects of Interpersonal Communication.*, Forthcoming Phd Dissertation.

[19] C. R. Perrault and J. Allen. 1980. A plan-based analysis of indirect speech acts . *American Journal of Computational Linguistics* (6) 3-4. 167-182.

[20] J. Searle. 1969. *Speech acts.* New York: Cambridge University Press.

[21] A. Wierzbicka. 1987. *English speech act verbs.* Sydney: Academic Press.