

# Influences on Attribute Selection in Redescriptions: A Corpus Study

Pamela W. Jordan (jordan@isp.pitt.edu)

Intelligent Systems Program

University of Pittsburgh

Pittsburgh PA 15260

## Abstract

We report on the results of the first stage of a corpus analysis that tests five hypotheses about how domain and discourse goals and the functions of repetition can influence the content of redescrptions in dialogue. We found a positive correlation between the attributes expressed in redescrptions and contexts in which three of these types of goals are predicted. These results provide us with guidance on the selection strategies we will test in the next stage of analysis.

## Introduction

In an extended discourse, speakers often redescribe objects that were introduced earlier in order to say something more about the object or the event in which it participates. The main goal when redescribing an entity is to re-voke the appropriate discourse entity. However, a goal-directed view of sentence generation suggests that speakers can attempt to satisfy multiple goals with each utterance [Appelt, 1985] and that a single linguistic form can opportunistically contribute to the satisfaction of multiple goals [Stone and Webber, 1998]. The possibility that goals besides identification could influence the content of a nominal expression<sup>1</sup> have not been fully addressed in computational work on generating nominal expressions.

The many-one mapping of goals to linguistic forms is more generally referred to as *overloading intentions* [Pollack, 1991]. Overloading can involve tradeoff across linguistic levels [Di Eugenio and Webber, 1996, Stone and Webber, 1998]. For example, an intention which is achieved by complicating a form at the semantic level may allow the speaker to simplify at the syntactic level by omitting important information [Stone and Webber, 1998].

Although we have learned that overloading is natural and perhaps even necessary, we have no well supported account of what degree of overloading is reasonable and what forms can more readily address multiple goals in dialogue. Without such an account, we have no principled way to deploy overloading in the automatic generation of natural language. Without well supported constraints on overloading, we are liable to create overloads in unnatural ways which will actually impede effective communication.

To investigate whether overloading applies to redescrptions, we examined 166 non-pronominal redescrptions found in 13 dialogues of the COCONUT

corpus [Di Eugenio et al., 2000]. This corpus contains computer-mediated dialogues in which two people collaborate on a simple design task, buying furniture for two rooms of a house. The participants' main goal is to negotiate the purchases; the items of highest priority are a sofa for the living room and a table and four chairs for the dining room. The participants also have specific secondary goals which further complicate the problem solving task. Participants are instructed to try to meet as many of these goals as possible, and are motivated to do so by associating points with satisfied goals. The secondary goals are: (1) Match colors within a room, (2) Buy as much furniture as you can, (3) Spend all your money. Each participant is given a separate budget and inventory of furniture and must decide what to make mutually known. Every furniture item in the inventory is described by five attributes; type, color, price, owner, and quantity.

In this article, we report on the results of the first stage of a corpus analysis that tests five hypotheses about how domain and discourse goals and the functions of repetition can influence the content of redescrptions in dialogue. We found a positive correlation between the attributes expressed in redescrptions and contexts in which three of these types of goals are predicted. The positive correlations will guide us in implementing and testing selection strategies for generating nominal expressions.

## Hypotheses about influences on redescrptions

Our hypotheses reflect non-identification goals that could influence the choice of attributes for a redescription. These goals are derived from work on the functions of repetition at the utterance or propositional level [Walker, 1993, Johnstone, 1994] and from observations about task intentions and constraint changes (e.g. matching colors) that were not directly communicated by the dialogue participants.

Our first hypothesis is based on the observation that in the COCONUT corpus, people often adjusted task constraints with no explicit discussion (38%).<sup>2</sup> Similarly to [Walker, 1993] for the propositional level, we suggest that the hearer is inferring changes from the redundancies in the redescription. So the repeated property could

<sup>1</sup>Identification being satisfied by more than a nominal expression also deserves consideration.

<sup>2</sup>Constraints are found in planning and scheduling tasks as well as design tasks [Di Eugenio et al., 2000, Jordan, 2000].

both help uniquely identify the intended discourse entity and enable the hearer to infer constraint changes.

To illustrate this possibility for COCONUT, assume that there is an initial constraint setting to match colors, that the speaker just discussed using a red table and prior to that introduced tables of various colors, four \$100 red chairs, and four \$75 green chairs. Finally, assume that she has decided to drop the color match constraint and suggest the cheaper \$75 chairs. When she communicates her suggestion, we hypothesize that she will prefer saying *green chairs* or *\$75 green chairs* over the more economical *\$75 chairs*. By choosing “green,” she adequately identifies the target chairs while also enabling the hearer to infer that she intends to drop the color match constraint. She has eliminated having to explicitly communicate the information [Walker, 1993] and reduced the risk of the hearer missing the inference [Carletta, 1992].

DOMAIN CONSTRAINT CHANGES HYPOTHESIS: Properties related to constraint changes are expressed in a context where the change must be inferred by the hearer.

Intentional relations indicate that an utterance, for example, elaborates or motivates other utterances, helping to bind together utterances to form a coherent discourse [Mann and Thompson, 1987]. Since the relations between utterances can influence the content and form of utterances [McKeown, 1985, Moser and Moore, 1995], we suggest that they could influence redescriptions as well. For example, in a context where all the object information is mutually known, a table needs to be chosen and red chairs have already been selected, (1b) can be considered motivation for the choice made in (1a).

- (1) a. Let’s use my table.  
b. It is red.

We know that it is beneficial under certain cognitive resource limitations to make motivations explicit even when the motivation is mutually known [Walker, 1993]. When we replace (1) with (2), the redescription may subsume the motivation and do the same thing.

- (2) Let’s use my red table.

PERSUASION HYPOTHESIS: Property values that are pivotal for deliberation are expressed in the context of goals to communicate a proposed action.

If a speaker repeats an utterance and provides no new information, this can show that a stage of the interaction is complete [Whittaker and Stenton, 1988, Jordan and Di Eugenio, 1997]. Repeating properties for a recently evoked item could show that the current stage has just been completed while doing so for an

older item could indicate that a higher level subproblem has been completed. In (3), S’s second utterance appears to end a stage in the interaction, in this case the end of the agreement process for a *select sofa* action [Di Eugenio et al., 2000].

- (3) S: ...I have a \$300 yellow sofa...

G: My sofa’s are more expensive so buy *your \$300 yellow sofa*. Also...

S: ... I will go ahead and buy *the \$300 yellow sofa*.

COMMITMENT HYPOTHESIS: In the context of a commitment to a proposal, all the properties expressed in the proposal will be repeated.

The second case of indicating that a higher level subproblem has been completed is a summary. Given the goals for the COCONUT design problem, the participants need to agree on the furniture items selected, whether colors should match in a room and whether they have selected as many different items as they can. In addition they have the hard constraint of not overspending. We suggest that a speaker would need to review all of these decisions in order to summarize the currently agreed upon solution state. Since the furniture items all have attributes related to each of these decisions, we hypothesize that a speaker will economize his summarization by including all the attributes that relate to decisions when listing the agreed upon items in the current solution. For example, in (4), the participants have previously decided on the dining room items and are completing their selections for the living room. Note that when G requests a summary of the living room, D includes all the attributes that relate to decision making for the task. D also decides to review all the items that they had previously decided upon for the dining room.

- (4) G: I got the rug. What do you have in the living room and what are the prices of the items

D: the green sofa in the living room 350. dining room—> 3 yellow chairs 75 each, 1 high-table yellow, 1 yellow rug

SUMMARIZATION HYPOTHESIS: In the context of a previously completed problem or subproblem, all decision relevant, mutually known properties for an item will be repeated.

Finally, a speaker might also repeat an utterance to show that it was understood [Clark and Schaefer, 1989, Brennan, 1990, Walker, 1992, Walker, 1993]. In the COCONUT corpus, the hearer sometimes repeats the description in the turn immediately following. For example, in (3), G repeats S’s description of the sofa, although the sofa was introduced by S. We claim that this type of redescription could help verify that the property information was correctly understood.

VERIFICATION HYPOTHESIS: In the context of a newly introduced entity, all the properties expressed will be repeated by the hearer in his/her next turn.

### Analyzing the Corpus

To identify the contexts and attribute usages we described in the hypotheses we used annotated features<sup>3</sup> and other easily extractable features of the corpus (e.g. the utterance speaker and the proximity of a redescription to its last mention). We used two types of corpus annotation features to support our study: (1) discourse entity level annotations that capture (a) the definitions and updates for discourse entities as a dialogue progresses and (b) the properties selected to redescribe discourse entities, and (2) utterance level annotations that capture (a) the problem solving state in terms of goals and constraint changes, and (b) discourse features such as commitments and offers. All of the features we used were found to have good intercoder reliability [Di Eugenio et al., 1998, Jordan, 1999]. The annotation features are described in detail in [Jordan, 2000].

Of these feature, the discourse features are the most complicated. They are based on elements of the agreement process described in [Di Eugenio et al., 2000]. The high-level definitions for these features are:

- propose: The speaker offers an item in a context where he already knows the hearer’s alternatives.
- partner decidable option: The speaker describes or offers an item but does not know the hearer’s alternatives.
- unconditional commit: The speaker indicates his commitment to using an item.
- unendorsed option: The speaker offers an item in a context where he knows the hearer’s alternatives and indicates an alternative is better.

Next we will describe how we used the annotation features to identify the contexts and redescrptions indicated in our hypotheses.

### Results of Corpus Analysis

We used chi-square and the Fisher exact tests<sup>4</sup> to check for correlations between factors. Although these tests assume independence, we feel we can violate this assumption given that the dependencies between redescrptions aren’t necessarily direct and obvious.<sup>5</sup> In all of the contingency tables, the counts are restricted to utterances

<sup>3</sup>The annotators for the corpus only knew the high-level goal of the study. Because of this, their annotation decisions were not influenced by the hypotheses we are studying here.

<sup>4</sup>We use the Fisher exact test when  $N < 20$  and an expected frequency is  $\leq 5$ .

<sup>5</sup>On average, the dialogues have 42 utterances, 25 discourse entities, and 6 utterances between redescrptions.

Changes	Related Properties
Room Color Limit	color
Price Upper Limit	price
Price Evaluator	price
Property Limit	color, price

Table 1: Associated Properties and Changes

that contain redescrptions. Finally the counts were all done automatically using software that interpreted the annotation features since the contextual factors generally involved multiple annotation features.

**Domain Constraint Changes Hypothesis** For this hypothesis we test whether there is a difference in property usage when a constraint change is communicated implicitly or explicitly. COCONUT is annotated with features indicating (1) whether a constraint change was communicated and whether this was accomplished implicitly or explicitly (2) which properties were included in the redescrptions. We examined each utterance for every constraint change that is generally possible for the domain when populating the cells of the contingency table.

We only count properties that relate to constraints. For example, we only look at the usage of the color property for the color match constraint or price for placing price limits. In Table 1, we list each of the constraint types that we examined and the property that we expected would be useful for inferring that change.<sup>6</sup> Our expectations derive from the instructions given to the COCONUT dialogue participants.

	Property Used	Property not Used
Implicit change	9	0
Explicit change	2	11

Table 2: Contingencies for Domain Constraint Changes Hypothesis

Table 2 shows that in the context of an implicit constraint change, properties related to the change are more likely to be used in the description than when the change is explicit (Fisher Exact Test,  $p < 0.0002$ ).

**Persuasion Hypothesis** For the Persuasion hypothesis, we wish to test whether expressing a property in a redescription is related to whether the expressed property makes the redescrbed item more desirable as a solution for a goal than the alternatives. For example, the cost of the item being redescrbed might be lower than any of the alternatives that have been discussed so far.

<sup>6</sup>The relevant property for the property limit constraint is indicated in the annotation for the constraint change.

A persuasion context exists when a proposal is to be made and alternate solutions exist and there is a contrast between the colors or prices that make the proposed item clearly a better choice. Given the analysis of the agreement process in [Di Eugenio et al., 2000], we first look for either a propose utterance, or an unconditional commitment utterance where the previous state is an unendorsed option, a partner decidable option or a list of options in which the speaker intentions are unclear.

For each of the unconditional commitment cases, we present examples. First, in (5), A’s partner decidable option is followed by B’s unconditional commitment.

(5) A: I have a blue sofa for \$200.

B: I have a yellow sofa for \$250. Let’s go with your \$200 sofa.

In (6), B does not endorse the option he presents but A overrides his objection with an unconditional commitment to it.

(6) A: We have \$100 left. I still have that \$50 blue chair.

B: I have a rug for \$100, but it is yellow.

A: We don’t need to match. Let’s get your \$100 rug.

Finally, in (7), A lists all of the items he has available. From the perspective of the agreement structure, lists such as this have no high-level task goals associated with them. However, the items do become part of the dialogue participants shared knowledge allowing all the items to be considered during problem solving so that they can become alternative options for the goals they are implicitly associated with. Because B is a position to deliberate<sup>7</sup>, his second utterance is annotated as an unconditional commitment. In this case there are two possibilities for what sofa to select, a persuasion context arises.

(7) A: I only have 2 red tables for \$200, 1 green table for \$350 and 4 \$50 blue chairs. I don’t have any rugs or lamps but I have 1 yellow sofa for \$200.

B: I have yellow rug for \$75 and a blue sofa for \$200. Let’s buy your yellow sofa and my rug.

Once we have identified possible persuasion contexts, we need to check for contrasts with alternatives. The alternatives are approximated by accumulating a list of the items evoked for each action. After a propose or unconditional commitment, all the items in the list for an action get flushed before starting over with the proposed item.

<sup>7</sup>This deliberation requirement for unconditional commitment is related to the problem solving architecture and is justified in [Di Eugenio et al., 2000].

Contrast	Related Property
Matches room but not alternatives	color
Cheaper than alternatives	price
More expensive than alternatives (near end of problem)	price

Table 3: Associated Properties and Contrasts

Next we check for contrasts. The contrast possibilities are shown in table 3 and arise from the COCONUT problem description. We were unable to accurately model the goal of buying as much as possible with the annotations available. For color we compare the color of the proposed item to those items already selected for the room and the alternate items. If the proposed item matches items already selected for the room while none of the alternates do, then a persuasion context exists. For prices there are two possibilities that depend on whether or not the end of the problem solving effort is nearing. An item may be a better choice when either (1) the price of the proposed item is greater than that of each alternate (i.e. it may be helping to spend out the budget) or (2) the price of the proposed item is less than that of each alternate (i.e. the cheaper item may be preferred since it leaves some money for other purchases).

Table 4 shows support for the persuasion hypothesis ( $\chi^2 = 5, p < .05, df = 1$ ).

	Property Not Used	Property Used
no contrast	18	9
contrast	13	24

Table 4: Contingencies for Persuasion Hypothesis

**Commitment Hypothesis** Here we test whether in the context of a commitment to a proposed action all the properties expressed in the proposal are more likely to be repeated. A commitment context exists when either (1) there is a previous proposal or unconditional commitment for the action involving the entity in the immediately previous turn and no other items must have been discussed for the action in the interim or (2) a speaker unconditionally commits again after doing so in his previous turn.

When determining repeated properties, we discount the type and owner properties. The type property is excluded because it involves pronominalization and zero anaphora; issues we are not addressing in this research. We exclude the owner property because its only function is identification in this domain.

Table 5 indicates that in contexts where a commitment is predicted, all mutually known properties are more likely to be included in redescrptions (Fisher Ex-

act Test,  $p < .0171$ ).

	Not Repeat Properties	Repeat Properties
No Commitment	7	8
Commitment	2	20

Table 5: Contingencies for Commitment Hypothesis

**Summarization Hypothesis** Here we test if the previous completion of a problem or subproblem correlates with expressing all the decision related, mutually known properties in a redescription. First, we must isolate redescriptions that occur after an agreement has been reached for the action.

A summarization context exists when an agreement has been reached for the action without the action being readdressed between the agreement and the current turn. The achievement of an agreement state is approximated when either (1) a propose or partner decidable option was the last state for the action and it happened more than two turns ago or (2) an unconditional commit was the last state and it happened two or more turns ago. In the first case, the agreement must be inferred and in the other the agreement is more explicit.

For the agreement state under condition (1), we require more than two turns to intervene because we want to allow for the cases where the partner left the decision pending by moving on to a dependent action (e.g. a final table decision may be left pending until the chair options are explored). We are estimating that if the action is not revisited after three turns, then it was not put on hold pending work on another action and that the partner agreed by moving on to another independent action.<sup>8</sup> This test for agreement takes into consideration that the initiation of the relevant next contribution shows evidence of understanding [Clark and Schaefer, 1987] and possibly joint commitment. For condition (2), we require that there be an intervening turn so that the partner is able to show that he has moved on to some other problem.

As with the commitment hypothesis, the type and owner properties are excluded when determining whether mutually known properties are repeated.

Table 6 indicates there is no correlation between a summarization context as we have characterized it and whether all the mutually known properties that relate to decisions get repeated ( $\chi^2 = 1.49$ ,  $df = 1$ , NS).

**Verification Hypothesis** With this hypothesis we test whether the repetition of all the properties presented in a previous description correlate with a con-

<sup>8</sup>In the initial version of the annotation scheme, there was a feature for indicating dependent actions but it was dropped because of poor intercoder reliability.

	All Mutual Properties Used	Not All Mutual Properties Used
Not End of Agreement Process	54	117
End of Agreement Process	8	8

Table 6: Contingencies for Summarization Hypothesis

text in which the entity was just introduced. In this case we collect all the properties that were presented in the turn where the item was first described and check whether this mention of the item was in the immediately previous turn or further back in the dialogue. As with the commitment and summarization hypotheses, the type and owner properties are excluded when determining whether properties are repeated. Table 7 shows no correlation between the verification context and the choice of attributes ( $\chi^2 = .06$ ,  $df = 1$ , NS).

	Properties Not All Repeated	Properties All Repeated
initial not in previous turn	1	0
initial in previous turn	44	2

Table 7: Contingencies for Verification Hypothesis

## Conclusion

Our analysis of the COCONUT corpus, shows positive correlations between the content of redescriptions and three of the contexts in which the repetition and domain and discourse goals we considered are expected. In particular, the contexts in which constraint changes, reasons for proposing, and commitment to proposals are predicted, positively correlated with the attributes expressed in the redescriptions of discourse entities. Finally, we found no support for the hypotheses that the properties expressed in a redescription correlate with verification or summarization contexts. In the case of the verification context, it is possible that the non-interruptibility of the COCONUT communications setting makes this sort of repetition function unnecessary. In the case of the summarization context, our ability to accurately detect this context may have been hampered by the estimates we had to make about the current state of the problem solving. Furthermore, there may be additional influences that depend on the reason for the summarization or the point at which it occurs during problem solving.

In future work we will test an attribute selection algorithm that embodies these hypotheses and compare

it against human performance and baseline algorithms that only consider the identification goal for redescrptions (e.g. IDAS [Dale and Reiter, 1995]).

## References

- [Appelt, 1985] Appelt, D. E. (1985). Planning English referring expressions. *Artificial Intelligence*, 26(1):1–33.
- [Brennan, 1990] Brennan, S. E. (1990). *Seeking and Providing Evidence for Mutual Understanding*. PhD thesis, Stanford University Psychology Dept. Unpublished Manuscript.
- [Carletta, 1992] Carletta, J. C. (1992). *Risk Taking and Recovery in Task-Oriented Dialogue*. PhD thesis, Edinburgh University.
- [Clark and Schaefer, 1987] Clark, H. H. and Schaefer, E. F. (1987). Collaborating on contributions to conversations. *Language and Cognitive Processes*, 2:19–41.
- [Clark and Schaefer, 1989] Clark, H. H. and Schaefer, E. F. (1989). Contributing to discourse. *Cognitive Science*, 13:259–294.
- [Dale and Reiter, 1995] Dale, R. and Reiter, E. (1995). Computational interpretations of the Gricean maxims in the generation of referring expressions. *Cognitive Science*, 19(2):233–263.
- [Di Eugenio et al., 1998] Di Eugenio, B., Jordan, P. W., Moore, J. D., and Thomason, R. H. (1998). An empirical investigation of collaborative dialogues. In *ACL-COLING98, Proceedings of the Thirty-sixth Conference of the Association for Computational Linguistics*, Montreal, Canada.
- [Di Eugenio et al., 2000] Di Eugenio, B., Jordan, P. W., Thomason, R. H., and Moore, J. D. (2000). The agreement process: An empirical investigation of human-human computer-mediated collaborative dialogues. *To Appear in International Journal of Human-Computer Studies*.
- [Di Eugenio and Webber, 1996] Di Eugenio, B. and Webber, B. (1996). Pragmatic overloading in natural language instructions. *International Journal of Expert Systems, Special Issue on Knowledge Representation and Reasoning for Natural Language Processing*, 9(1):53–84.
- [Johnstone, 1994] Johnstone, B. (1994). Repetition in discourse: A dialogue. In Johnstone, B., editor, *Repetition in Discourse: Interdisciplinary Perspectives, Volume 1*, volume XLVII of *Advances in Discourse Processes*, chapter 1. Ablex.
- [Jordan, 1999] Jordan, P. W. (1999). An empirical study of the communicative goals impacting nominal expressions. In *Proceedings of the ESSLLI workshop on The Generation of Nominal Expressions*.
- [Jordan, 2000] Jordan, P. W. (2000). *Intentional Influences on Object Redescrptions in Dialogue: Evidence from an Empirical Study*. PhD thesis, Intelligent Systems Program, University of Pittsburgh.
- [Jordan and Di Eugenio, 1997] Jordan, P. W. and Di Eugenio, B. (1997). Control and initiative in collaborative problem solving dialogues. In *Computational Models for Mixed Initiative Interaction. Papers from the 1997 AAAI Spring Symposium. Technical Report SS-97-04*, pages 81–84. The AAAI Press.
- [Mann and Thompson, 1987] Mann, W. and Thompson, S. (1987). Rhetorical Structure Theory: A Framework for the Analysis of Texts. Technical Report RS-87-190, USC/Information Sciences Institute.
- [McKeown, 1985] McKeown, K. R. (1985). *Text Generation. Using Discourse Strategies and Focus Constraints to Generate Natural Language Text*. Cambridge University Press.
- [Moser and Moore, 1995] Moser, M. and Moore, J. D. (1995). Investigating cue placement and selection in tutorial discourse. In *Proceedings of 33rd Annual Meeting of the Association for Computational Linguistics*, pages 130–135.
- [Pollack, 1991] Pollack, M. E. (1991). Overloading intentions for efficient practical reasoning. *Noûs*, 25:513–536.
- [Stone and Webber, 1998] Stone, M. and Webber, B. (1998). Textual economy through close coupling of syntax and semantics. In *Proceedings of 1998 International Workshop on Natural Language Generation*, Niagra-on-the-Lake, Canada.
- [Walker, 1992] Walker, M. A. (1992). Redundancy in collaborative dialogue. In *Fourteenth International Conference on Computational Linguistics*, pages 345–351.
- [Walker, 1993] Walker, M. A. (1993). *Informational Redundancy and Resource Bounds in Dialogue*. PhD thesis, University of Pennsylvania.
- [Whittaker and Stenton, 1988] Whittaker, S. and Stenton, P. (1988). Cues and control in expert client dialogues. In *Proc. 26th Annual Meeting of the ACL, Association of Computational Linguistics*, pages 123–130.

## Acknowledgments

This research was supported by the National Science Foundation under Grant No. IRI-9314961.