Reactive Redundancy and Listener Comprehension in Direction-Giving

Rachel E. Baker
Department of Linguistics
Northwestern University
Evanston, IL 60208
r-baker2@northwestern.edu

Alastair J. Gill, Justine Cassell
Center for Technology and Social Behavior
Northwestern University
Evanston, IL 60208
{alastair,justine}@northwestern.edu

Abstract

We explore the role of redundancy, both in anticipation of and in response to listener confusion, in task-oriented dialogue. We find that direction-givers provide redundant utterances in response to both verbal and non-verbal signals of listener confusion. We also examine the effects of prior acquaintance and visibility upon redundancy. As expected, givers use more redundant utterances overall, and more redundant utterances in response to listener questions, when communicating with strangers. We discuss our findings in relation to theories of redundancy, the balance of speaker and listener effort, and potential applications.

1 Introduction

Our everyday conversations represent a carefully negotiated balance between the perceived needs of the speaker and the listener. These opposing forces affect every aspect of language from phonetics to pragmatics. A careful balance between these two forces allows speakers to produce language that is both efficient and effective at communicating a message (Lindblom, 1990; Horn, 1993). Of course, the same balance is not appropriate for every situation. When accuracy is critical to the message, or when the speaker perceives the listener to have difficulty understanding, the speaker is more likely to prioritize clarity over efficiency, resulting in more explicit communication. In contrast, during casual conversation or when speed is a factor, the speaker may choose a more reduced, efficient, communication style (Lindblom, 1990; Horton and Keysar, 1996). A number of scholars have pointed out that speakers seem to use the information available to themselves rather than that available to the listener to guide certain linguistic decisions, such as clarity of pronunciation and choice of syntactic structure (Bard et al., 2000; Branigan et al., 2003). However, these studies examine utterance form, while our study examines content, which is more influenced by audience design (Branigan et al., 2003). In every utterance, a speaker either reduces the likelihood of listener misunderstanding by being more explicit, or reduces their own effort by providing a minimal amount of information. Regardless of whether speakers pro-actively monitor the information needs of listeners, they do need to respond when listeners say or do something to indicate confusion. Developing a better understanding of the factors that affect how and when speakers respond to signs of listener confusion is important at both theoretical and applied levels: first, it can better explain the variation in discourse strategies used in different communicative situations; second, it can help in the design of dialogue systems (Kopp et al., 2008; Theune et al., 2007).

In this study, we examine what types of listener behavior increase the likelihood that a speaker will produce a redundant utterance. We also examine how communicative context affects the amount redundancy a speaker produces overall (Walker, 1992, 1996) and a speaker's use of redundancy in response to listener confusion. In contrast to previous work, we study *reactive* redundancy, or redundancy produced in response to signs of listener confusion. We investigate two factors that may influence a speaker's tendency to produce redundant utterances and to respond to listener confusion with redundancy: the relationship between the interlocutors and their visual contact.

In the following section, we review relevant literature and present our hypotheses; we then describe the direction-giving experiment which we used to examine redundancy in task-oriented dialogue, and present our results; we discuss our results in light of the literature and conclude by noting potential applications and future work.

Proceedings of the 9th SIGdial Workshop on Discourse and Dialogue, pages 37–45,
Columbus, June 2008.

2 Related Work and Predictions

2.1 Redundancy

Grice's (1975) second Maxim of Quantity: 'Do not make your contribution more informative than is required' has led to the general impression that redundancy (providing discourse-old information) is avoided in language (Stalnaker, 1978), with this mirrored by work in natural language generation (Dalianis, 1999). However, Walker (1992, 1996) points out that such conclusions relating to redundancy are often based on flawed assumptions. For example, they assume that agents have unlimited working memory and the ability to automatically generate all the inferences entailed by every utterance, that utterance production should be minimized, and that assertions by Agent A are accepted by default by Agent B (Walker, 1996: 183).

In fact, redundancy can serve many desirable purposes in communication. Redundancy has been shown to increase text cohesion and readability (Horning, 1991) as well as provide evidence of understanding and grounding, make a proposition salient, and make inferences explicit (Walker, 1996). A computer simulation of a cooperative task dialogue between two agents suggested that the use of certain types of redundant utterances improved the performance of the pair (Walker, 1996).

Fussell and Krauss (1989a) point out that there are two methods that speakers can use to tailor their message for the listener. The first method involves predicting what information it is necessary to communicate, using knowledge of the listener's interests and background. The second method involves modifying the message in response to listener feedback. Walker's model only captures the use of redundancy in the service of the first method. We will refer to this type of redundancy as *proactive* redundancy, whereby a speaker provides redundant information without waiting for the listener to express a need for it. The advantages of providing redundant information proactively include being able to integrate the redundant information with the new information, and avoiding conflict by removing the necessity for the listener to express a lack of understanding (Brown and Levinson, 1987).

We hypothesize that speakers also use redundancy *reactively*, after the listener signals a lack of understanding, either verbally or non-verbally. This is redundancy in service of Fussell and Krauss' second method of message-tailoring. The advantages of providing redundant information reactively include increasing the efficiency of the exchange by only providing redundant information that the listener communicates a need for, and reducing the burden on the speaker of having to decide when to include redundant information.

One important distinction between proactive and reactive redundancy is the grounding status of the redundant information. Reactive redundancy is likely to provide information that has not been accepted by the listener, and is therefore not part of the common ground (Clark and Schaeffer, 1989), even though it is discourse-old. In contrast, proactive redundancy is likely to provide information from the interlocutors' common ground. Indeed, Walker (1996) describes Attitude redundant utterances as providing evidence of grounding. Walker's other types of proactive redundancy (Consequence and Attention) make inferences based on grounded utterances explicit and make elements of the common ground salient again.

Reactive redundancy is one type of repair, like expansions and replacements, which can be used in response to non-understanding or misunderstanding (Hirst et al., 1994). The type of miscommunication may influence a speaker's choice of repair strategy, with reactive redundancy being an appropriate response to mishearing or misremembering.

However, producing redundant information, even when the listener signals a need for it, incurs a cost. Including redundant information increases the length of the dialogue and the speaker's effort, and decreases the amount of new information provided within a certain length of time. In these cases the speaker must decide how much redundant information to provide and when to provide it.

2.2 Signals of Confusion

Listeners can express a need for information to be repeated or restated in a number of ways, both verbally and non-verbally. Brinton et al. (1988) used questions and statements of confusion ("I didn't understand") as signs of communication breakdowns. Morrow et al. (1993) describe inaccurate and partial repetitions of instructions as elements of miscommunication. This prior work leads us to examine questions, utterances signaling non-understanding (e.g. "I don't remember what's next"), incorrect repetitions (e.g. "take the third right" after the direction-giver said "take the second right") and abandoned utterances (e.g. "Then I'll turn…") as possible signs of listener confusion. We predict redundancy after such statements because they all indicate that a piece of information has not been understood.

We also examine eye-gaze as a non-verbal marker of listener comprehension. Goodwin (1981) described gaze towards the speaker as a sign of listener attention. However, Nakano et al. (2003) found that speakers seemed to interpret a listener gazing at them rather than at a map as a sign of listener misunderstanding. Therefore, shifting eye-gaze away from the speaker can signal that a listener is losing attention, perhaps due to confusion, while shifting gaze towards the speaker can signal misunderstanding. In this study there is no map, and listeners who can see the speaker spend most of the conversation gazing at the speaker. Still, due to the opposing findings in the literature, we analyze eye-gaze shifts both towards and away from the speaker as potential signs of listener confusion.

2.3 Relationship and Communication

Speakers are more explicit when communicating with strangers or people with whom they share less common ground. This explicitness can take the form of highly informative self-introductions on the phone (Hornstein, 1985), longer descriptions of abstract figures (Fussell and Krauss, 1989b), and explicit references to utterance topics (Svedsen and Evjemo, 2003). These studies indicate that speakers attempt to make up for the small amount of common ground they share with strangers by including more information in the discourse itself.

Another difference between friends and non-friends is that acquaintances tend to be more formal, more concerned with self presentation, less negative, and less likely to disagree than friends (Schlenker, 1984; Tickle-Degnen and Rosenthal, 1990; Planalp and Benson, 1992). Therefore, we expect that in an initial interaction, a speaker will try to appear competent and avoid conflict.

As noted above, speakers talking to strangers are more explicit, leading us to predict more redundancy overall. They are also more likely to try to impress their interlocutor and avoid conflict, leading to more reactive redundancy in response to confusion when the pair are strangers.

2.4 Visibility and Communication

Visibility also has a number of effects on communication. One of the most basic is that when interlocutors cannot see each other they cannot use non-verbal signals to communicate, so they must rely on verbal communication. For example, the use of eye-gaze as a sign of listener attention (Argyle and Cook, 1976; Goodwin, 1981) is only possible when interlocutors can see each other. When they cannot see each other, they must indicate attention verbally or do without this information.

Visibility affects both the form and the outcomes of a conversation. When interlocutors cannot see each other, conversations are longer and contain more, shorter, utterances than when they can (Nakano et al., 2003). Interlocutors in an investment game who could not see each other also did not establish trust to the same extent as those who met face-to-face (Bos et al., 2002).

Because speakers who cannot see each other have fewer channels of communication available to them, their interaction can be more difficult than a face-to-face interaction. We predict that this will lead them to use more redundancy and more reactive redundancy in an effort to be clear.

2.5 Hypotheses

In order to study how responsive speakers are to signs of listener confusion, we must first determine what signs speakers respond to. In this study we examine a number of verbal and non-verbal signs speakers may use to gauge listener confusion. In particular, we expect that speakers will provide redundancy in response to both verbal signs like questions, statements of non-understanding, incorrect statements, and abandoned utterances, and non-verbal signs like eye-gaze changes. We expect that speakers will strike a different balance between efficiency (minimizing speaker effort) and clarity (minimizing listener effort) depending on the relationship between the speaker and listener, and the physical context of the interaction. We expect speakers to use redundancy strategies focused on minimizing speaker effort when addressing friends and people they can see. Such strategies involve less redundancy (and therefore less speaking), and less reactive redundancy (requiring less listener monitoring). Conversely, we expect to find redundancy strategies maximizing clarity when speakers address strangers and people they cannot see. Such strategies involve more redundancy overall (providing the listener with more information in general) as well as more reactive redundancy (which provides the listener with the specific information they may require).

Hypothesis 1 - Redundancy and Non-Understanding

(a) Verbal cues - Direction-givers will provide redundancy when the receiver verbally expresses a lack of understanding by asking a question, abandoning an utterance, making an incorrect statement or explicitly expressing non-understanding.

(b) Non-verbal cues - Givers will provide redundancy when the receiver non-verbally expresses a lack of understanding by shifting eye-gaze.

Hypothesis 2 - Redundancy and Relationship Givers will prioritize clarity over efficiency in their redundancy use when speaking to strangers, providing (a) more redundancy and (b) more reactive redundancy than when speaking to friends.

Hypothesis 3 - Redundancy and Visual Contact Givers will prioritize clarity over efficiency in their redundancy use when they cannot see their partner, providing (a) more redundancy and (b) more reactive redundancy than when they can see them.

3 Methods

3.1 Participants

Twenty-four university students participated, resulting in twelve dyads. All were paid $10 for their participation and received $5 gift certificates if they successfully completed the task. In each dyad the direction-giver was familiar with the building in which the experiment took place, and the direction-receiver was unfamiliar with it. Half the dyads were pairs of friends and half were strangers.

3.2 Procedure

The task consisted of three consecutive direction-giving sessions, as described in Cassell et al. (2007). At the start of each session, the experimenter led the direction-giver to a point in the building, and back to the experiment room. Half of the dyads sat facing each other during the direction-giving (the Vision condition) and half sat back-to-back with a screen between them (the No-vision condition). The direction-giver then explained the route to the direction-receiver. There were no time limits or restrictions on what could be said, but the dyads could not use maps or props. When the dyad decided that direction-giving was complete, they signaled the experimenter, who the receiver led to the goal, following the directions.

The direction-giving sessions were videotaped. Participants' speech was transcribed and coded for possible redundancy triggers and redundant utterances using the coding scheme described below. The time-aligned codings for the giver and receiver were aligned with each other using scripts that calculated which of the receiver's utterances or actions directly preceded which of the giver's utterances. The scripts classify a receiver's utterance or action as 'preceding' a giver's utterance if its start precedes the start of the giver's utterance and its end is not more than two seconds before the start of the giver's utterance. The two-second limit was used to avoid positing connections between a giver's utterance and receiver utterances that came long before it.

3.3 Data Coding

Each dialogue was divided into clauses, defined as units that include a subject and predicate and express a proposition. Each clause was coded using a modified version of DAMSL (Core and Allen, 1997). Direction-givers' and receivers' speech was coded differently because we only studied redundancy produced by the giver. We coded the receiver's speech for signs of confusion. We describe the labels we used in more detail below.

Each direction-giver's clauses were coded for Statements and Info-requests. The Info-request tag marks questions and other requests for information. In a Statement, a speaker makes a claim about the world. The class of Statements was broken down into Non-redundant, in which the speaker is trying to change or add to the hearer's beliefs, and Redundant, which contain only information that has already been stated or entailed.

Each direction-receiver's clauses were coded for Statements, Info-requests, Signal non-understandings (S.N.U.), and Abandoned utterances. The receiver's Statements were classified as either Correct or Incorrect. If an utterance explicitly expressed non-understanding of an earlier utterance it was coded as Signal non-understanding. This label was only used for direct statements of non-understanding, such as "I didn't follow that," and not for signals of non-understanding covered by other labels such as Info-requests and Incorrect Statements. Utterances that were abandoned (the speaker stops the utterance and it provides no content to the dialogue) were coded as Abandoned. Receiver utterances that were not coded as Info-requests, Incorrect Statements, Signal-non-understandings, or Abandoned, were coded as No-trigger. No-trigger utterances included correct statements and statements about task management.

4 Results

We found that a large proportion of giver utterances were redundant, ranging from 17% to 38% with a mean of 25%. Examples of redundancy from our recordings are listed in the Appendix.

We first analyzed the data using a hierarchical loglinear analysis with the variables: visual condition (Vision, No-vision), relationship (Friends, Strangers), receiver-utterance (Info-request, Incorrect statement, Signal non-understanding, Abandoned, No-trigger), and giver-utterance (Redundant, Non-redundant). The overall model is significant ($\chi^2_{(39,5294)}$=13254.157,p<.001), justifying chi-square comparisons of individual factors within the model. We report tests of partial association and chi-square tests to indicate where significant differences lie between groups.

4.1 Redundancy and Non-Understanding

Verbal Signals of Non-Understanding

We tested part (a) of Hypothesis 1 by running a test of partial associations (adjusted for all effects in the model) and an unpartialled chi-square (ignoring variables not included in the effect being tested). These showed a significant association between receiver-utterance and giver-utterance type (Partial $\chi^2_{(4,5294)}$=117.7, p<.001; $\chi^2_{(4,5294)}$=121.2,p<.001).

Chi-square tests comparing giver-utterances following predicted redundancy triggers to giver-utterances after No-trigger receiver utterances, indicate that Info-requests, Incorrect statements and Abandoned utterances all significantly increase the likelihood that the giver will produce a redundant utterance ($\chi^2_{(1,4907)}$=57.3,p<.001; $\chi^2_{(1,4562)}$=28.4, p<.001; $\chi^2_{(1,4651)}$=49.1,p<.001, respectively). Explicit Signal-non-understandings do not have significant effects on the likelihood of a redundant-utterance ($\chi^2_{(1,4539)}$=.3,p=.619). Figure 1 shows the percentages of giver utterances that were redundant following various receiver dialogue acts.

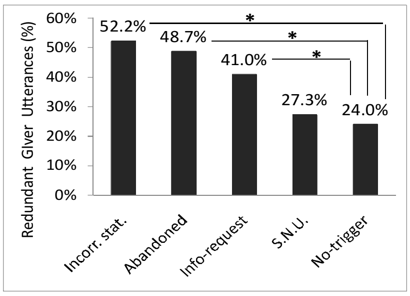

Figure 1. Percent of redundant giver utterances following various receiver dialogue acts.

Non-Verbal Signals of Non-Understanding

We tested part (b) of Hypothesis 1 with a separate hierarchical loglinear analysis examining only the dyads in the Vision condition for the effects of: relationship, receiver-utterance, giver-utterance, and receiver-gaze (Gaze-to, Gaze-away, and No-gaze-change). The first- and second-order effects are significant ($\chi^2_{(59,2815)}$=9582.4, p<.001).

A test of partial associations and a chi-square test indicate a significant association between giver-utterance and receiver-gaze (Partial $\chi^2_{(2,2815)}$= 22.7, p<.001; $\chi^2_{(2,2815)}$=24.7,p<.001). Chi-square tests comparing receiver gaze changes to non-changes show that redundant utterances are significantly more likely after a gaze change toward the giver ($\chi^2_{(1,2433)}$=21.5,p<.001) and after a gaze change away from the giver ($\chi^2_{(1,2475)}$=6.5,p<.05) than after no gaze change. A chi-square test comparing gaze change toward the giver to gaze change away from the giver shows that the difference between them is not significant ($\chi^2_{(1,722)}$=2.7, p=.098). These effects are shown in Figure 2.

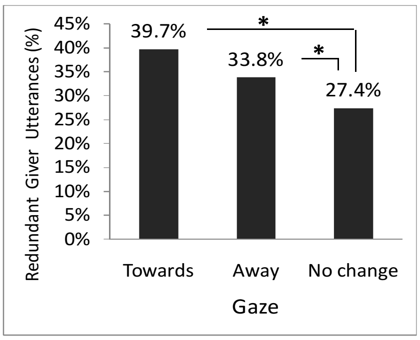

Figure 2. Percent of redundant giver utterances following receiver eye-gaze changes toward and away from the giver, and following no gaze change

4.2 Redundancy and Relationship

Part (a) of Hypothesis 2 was confirmed by the significant association between relationship and giver-utterance (Partial $\chi^2_{(1,5294)}$=13.3, p<.001; $\chi^2_{(1,5294)}$=6, p<.05) in our original analysis. A larger percentage of giver utterances are redundant in the Strangers condition (27.8%) than in the Friends condition (24.8%).

To examine part (b) of Hypothesis 2 we ran a hierarchical loglinear analysis after collapsing all receiver-utterances into question/non-question categories. This reveals a significant partial association among giver-utterance, receiver-utterance, and relationship (Partial $\chi^2_{(1,5294)}$=7.5, p<.01). A chi-square test comparing utterances after questions in the Friends and Strangers conditions shows that redundant utterances are significantly more likely after questions in the Strangers condition than the Friends condition ($\chi^2_{(1,412)}$= 14.6, p<.0005), as shown in Figure 3.

Three-way interactions among giver-utterance, receiver-utterance and relationship are not significant in any of the other analyses.

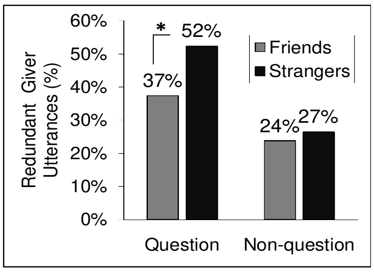

Figure 3. Percent of redundant giver utterances following questions and non-questions, by relationship.

4.3 Redundancy and Visual Contact

There is a trend-level association between visual condition and giver-utterance type (Partial $\chi^2_{(1,5294)}$ =4.6,p<.05; $\chi^2_{(1,5294)}$=3.3,p=.071). Contrary to Hypothesis 3, a larger percentage of utterances are redundant in the Vision condition (27.7%) than in the No-vision condition (25.5%). No significant association was found among giver-utterance, receiver-utterance, and visual condition, even when collapsed into question/non-question categories.

5 Discussion

This study set out to discover what verbal and non-verbal behaviors increase the likelihood of redundant utterances in direction-givers' speech. We also examined whether the interlocutors' relationship or visual contact influence whether speakers provide redundant utterances in anticipation of and in response to listener confusion. We found that givers used a large proportion of redundant utterances, (around 25% of utterances). Walker (1996) found that about 12% of utterances were redundant in a corpus of recordings from a call-in financial radio show. The higher proportion of redundant utterances in our study is predicted by Walker's (1996) model, in which a task's tolerance for comprehension errors influences whether redundant utterances are produced. In a radio advice show, a misunderstanding may be more easily recovered from than in direction-giving, in which one wrong turn could make it impossible to reach the goal.

In addition to revealing the impact of task tolerance to error on redundancy, this study sheds light on other circumstances that influence redundancy use. Givers produced reactive redundancy in response to the verbal triggers: Info-requests, Abandoned utterances, and Incorrect statements. However, even these triggers were not always followed by redundancy. In fact, only around 50% of the utterances following these triggers were redundant. Such a low response rate is surprising until we consider the diversity of utterances covered by these labels. For instance, some Info-requests seek new information (e.g. "What's at the top of the stairs?"), and some receiver utterances are abandoned because the giver interrupts with new information. Our study lays the groundwork for future examinations of speaker responses to listener confusion, which can refine these broad categories. We must also consider the variability in responses to listener confusion. We found that givers are more likely to provide redundant utterances in response to questions when speaking to strangers, but this is only one of many factors that could affect levels of responsiveness, including speaker personality, time pressure, and task difficulty.

The non-significant effect of Signals non-understandings on redundancy is surprising. This may be due to the small number of examples of this category in our recordings. We found only 44 instances of Signal non-understandings, in contrast to, for example, 156 Abandoned utterances.

The non-verbal cue gaze change also increased the likelihood of a redundant utterance. Interestingly, gaze changes both to and away from the giver triggered redundancy. This is consistent with both Nakano et al.'s (2003) finding that gazing at the speaker signals listener misunderstanding and Goodwin's (1981) finding that gazing away from the speaker indicates a lack of listener attention.

It is interesting that 24% of giver utterances following No-trigger receiver utterances were redundant. These probably include both redundant utterances triggered by signs of listener confusion that we did not code for, and proactive redundancy. Proactive redundancy can appear within the first description of some directions (see the No-trigger example in the Appendix) and when the whole set of directions is repeated as a memory aid.

The relationship between the interlocutors does affect the amount of redundancy speakers produce overall and in response to listener signs of confusion. Strangers used more redundant utterances than friends and provided more redundant utterances after questions. This supports our hypothesis that direction-givers speaking to strangers will prioritize clarity over efficiency. The more consistent use of reactive redundancy in the Strangers condition may be due to speakers' tendency to avoid confrontation with strangers. When responding to questions from friends, direction-givers may provide some new information because they know that their friend will feel comfortable asking another question if their answer is unclear. However, when answering questions from a stranger, the giver may wish to avoid the embarrassment of further confusion by repeating more discourse-old information.

However, contrary to our predictions, we did not find more redundancy or more reactive redundancy in the No-vision condition than the Vision condition. In fact, we found numerically more redundancy in the Vision condition. Given the low level of significance, we do not discuss this in detail, however we suggest that this could be due to the fact that there are more ways of signaling non-understanding available to the receivers in the Vision condition (both verbal and non-verbal). Therefore, even if givers do not increase their rates of reactive redundancy in the Vision condition, they could provide more reactive redundancy (and more redundancy overall) because they are receiving more cues to react to. Not all situations leading to communication difficulties encourage more redundancy or more reactive redundancy, but the increased explicitness and positivity typical of conversation between strangers do encourage it.

6 Conclusion

This study explored the use of redundancy in task-oriented dialogue, specifically the effects of listener behavior and communicative context on the amount of redundancy produced. We found that direction-givers provided redundant utterances in response to verbal and non-verbal signs of listener confusion. As predicted, givers were more likely to prioritize clarity over efficiency in their redundancy use (using more redundancy overall and more redundancy in response to questions) when speaking to strangers than friends. Contrary to our predictions, givers did not provide more redundant utterances when they could not see their listener.

Direction-giving, due to its high memory load and the need for the receiver to understand the giver almost completely, is a type of discourse that may encourage more redundancy than other types. Indeed, we note that our data have a much greater proportion of redundancies than discussions taken from radio talk shows (Walker, 1996). Future work should examine the nature of proactive and reactive redundancy in more varied discourse contexts, such as negotiation, teaching, and play. It should also explore the effects of memory load on redundancy by varying task complexity, which may be easier with a more controlled task like the Maptask. Researchers could study the relationship between saliency and redundancy by studying correlations between a segment's salience and its likelihood of being used in a redundant utterance.

Our findings can be used to improve the communicative efficacy of natural language generation systems like those used in Embodied Conversational Agents (ECAs; Kopp et al., 2008). For example, like strangers, direction-giving ECAs could use increased overall and reactive redundancy to compensate for the lack of shared common ground with the human user of the system. Analyses of the syntactic structures of different types of redundant utterances will be important for incorporating these results into generation systems.

Acknowledgments

We thank Paul Tepper, Gregory Ward, Darren Gergle, Alex Podbelski, and our anonymous reviewers for their helpful advice and hard work. We are grateful for generous funding from Motorola and NSF HCC 0705901.

References

M. Argyle and M. Cook. 1976. *Gaze and Mutual Gaze*. Cambridge University Press, New York.

E. Bard, A. Anderson, C. Sotillo, M. Aylett, G. Doherty-Sneddon and A. Newlands. 2000. Controlling the intelligibility of referring expressions in dialogue. *J. Memory and Language*, 42(1):1-22.

N. Bos, J. Olson, D. Gergle, G. Olson, and Z. Wright. 2002. Effects of four computer-mediated communication channels on trust development. In *Proceedings of SIGCHI 2002,* pages 135-140, Minneapolis, MN.

H. P. Branigan, J. F. McLean, and H. Reeve. 2003. Something old, something new: Addressee knowledge and the given-new contract. In *Proceedings of the 25th Annual Conference of the Cognitive Science Society,* pages 180-185, Boston, MA.

B. Brighton, M. Fujiki, and E. Sonnenberg. 1988. Responses to requests for clarification by linguistically normal and language-impaired children in conversation. *J. Speech and Hearing Disorders*, 53:383-391.

P. Brown and S. C. Levinson. 1987. *Politeness: Some Universals in Language Usage.* Cambridge University Press, Cambridge, UK.

J. Cassell, A. J. Gill, and P. Tepper. 2007. Coordination in conversation and rapport. In *Proceedings of the Workshop on Embodied Language Processing at ACL*, pages 41-50, Prague.

H. H. Clark and E. F. Schaeffer. 1989. Contributing to discourse. *Cognitive Science,*13:259-294.

M. G. Core and J. F. Allen. 1997. Coding dialogs with the DAMSL annotation scheme. In *Proceedings of AAAI Fall Symposium on Communicative Action in Humans and Machines*, pages 28-35, Boston, MA.

H. Dalianis. 1999. Aggregation in natural language generation. *J. Computational Intelligence*, 15(4):384-414.

S. R. Fussell and R. M. Krauss. 1989 a. The effects of intended audience on message production and comprehension: Reference in a common ground framework. *European J. Social Psychology,* 25:203-219.

S. R. Fussell and R. M. Kraus. 1989 b. Understanding friends and strangers: The effects of audience design on message comprehension. *European J. Social Psychology*, 19:445-454.

C. Goodwin. 1981. *Conversational Organization: Interaction between Speakers and Hearers.* Academic Press, New York.

H. P. Grice. 1975. Logic and conversation. In P. Cole and J. Morgan, editors, *Syntax and Semantics III – Speech Acts*. Academic Press,New York, pages 41-58.

G. Hirst, S. McRoy, P. Heeman, P. Edmonds, and D. Horton. 1994. Repairing conversational misunderstandings and non-understandings. *Speech Communication*, 15: 213-230.

L. Horn. 1993. Economy and redundancy in a dualistic model of natural language. *SKY 1993: Yearbook of the Linguistic Association of Finland*: 33-72.

A. Horning. 1991. Readable writing: The role of cohesion and redundancy. *J. Advanced Composition*, *11*:135-145.

G. Hornstein. 1985. Intimacy in conversational style as a function of the degree of closeness between members of a dyad. *J. Personality and Social Psychology,* 49(3):671-681.

W. Horton and B. Keysar. 1996. When do speakers take into account common ground? *Cognition,* 59:91–117.

S. Kopp, P. Tepper, K. Ferriman, K. Striegnitz and J. Cassell. 2008. Trading spaces: How humans and humanoids use speech and gesture to give directions. In T. Nishida, editor, *Conversational Informatics.* John Wiley & Sons, New York, pages 133-160.

B. Lindlom. 1990. Explaining phonetic variation: A sketch of the H and H theory. In W. Hardcastle & A. Marchal, editors, *Speech Production and Speech Modeling*. Kluwer, Dordrecht, pages 403-439.

D. Morrow, A. Lee, and M. Rodvold. 1993. Analysis of problems in routine controller-pilot communication. *International J. Aviation Psychology*. 3(4): 285-302.

Y. Nakano, G. Reinstein, T. Stocky, and J. Cassell. 2003. Towards a model of face-to-face grounding. In *Proceedings of ACL 2003*, pages 553-561, Sapporo, Japan.

S. Planalp and A. Benson. 1992. Friends' and acquaintances' conversations I: Perceived differences. *J. Social and Personal Relationships,* 9:483-506.

B. Schlenker. 1984. Identities, identifications, and relationships. In V. Derlega, editor, *Communication, Intimacy and Close Relationships.* Academic Press, New York, pages 71-104.

R. Stalnaker. 1978. Assertion. In P. Cole, editor, *Syntax and Semantics, Volume 9: Pragmatics*. Academic Press, New York, pages 315-332.

G. Svendsen and B. Evjemo. 2003. Implicit referring as an indication of familiarity in face-to-face and phone conversations. In *Proceedings of INTERACT '03:* pages 920-923, Zurich.

M. Theune, D. Hofs and M. van Kessel. 2007. The Virtual guide: A direction giving embodied conversational agent. In *Proceedings of Interspeech 2007*, pages 2197-2200, Antwerp, Belgium.

L. Tickle-Degnen and R. Rosenthal. 1990. The nature of rapport and its nonverbal correlates. *Psychological Inquiry*, 1(4):285-293.

M. A. Walker. 1992. Redundancy in collaborative dialogue. In *Proceedings of the 14th International Conference on Computational Linguistics*, pages 345-351, Nantes, France.

M. A. Walker. 1996. The effect of resource limits and task complexity on collaborative planning in dialogue. *Artificial Intelligence Journal*, 85:181-243.

Appendix: Examples from Dialogues

In the following examples, utterances in italics are the triggers produced by the receiver, and underlined utterances are redundant. Commas indicate pauses. Receiver utterances in square brackets overlap with the portion of the preceding giver utterance in brackets.

Question Example

Giver (G): as soon as you come outta the door, uhh on the second floor you'll [see like a window] in front of you
Receiver (R): [mmhm]
G: [and then], you'll wanna take a left
R: [hm]
…
G: if you look to your left you'll see the exit sign, uhh with for the stairwell
R: ok so then I go to this second floor
G: mmhm
R: *and then do I go right?*
G: no
R: *or left?*
G: you go left [once you come outta] the second floor
R: [you go left]

Incorrect Statement Example

G: and you're gonna go towards the computer, and pass the computer, and there will be, copy machines on your right after you pass the computer
R: mhmm
G: so after you, walk, just past the copy machines you're gonna want to take a hard left, almost like a U-turn
…
G: once you turn to the right at after the first stairs you'll you'll see a computer
R: oh a computer right ok *and then I'm gonna take a really hard left like a U-turn*
G: right well you go past the computer and then you'll see copying machines
R: oh ok
G: and then but, the copy machines are like maybe three five feet after the computer
R: ok
G: and then that's when you take the hard left

Abandoned Example

G: and then you're gonna hear some kids and people talking and stuff, you're gonna be heading toward the clinic
R: oh okay
G: okay, the clinic you're is gonna come up on your right, [there's gonna] be, kind of, semi circular blue couches
R: [okay], uhhuh
G: down there, the stapler, is on the floor, right next to a pillar, [um] so basically you're gonna like, you're gonna kind of, turn right to look into the clinic
R: [okay], okay
G: and then, the stapler's kinda just over there to the left, on the floor by one of the pillars
…
G: and you're gonna hear people talking and there's gonna [be kids]
R: *[okay] so and then the, pillar its' like gonna be one of the pillars on the, right by like I guess it's on the*
G: basically, basically um you walk into, the clinic, and there's blue, couches
R: mmhm
G: and then it's just a little bit over to the left
R: oh okay
G: on the floor

No-Trigger Example

G: open the door, and you're gonna see a set of stairs
R: okay
G: go down those stairs, to the second floor
R: mmhm
G so you're gonna be on the third floor, you're gonna then you're gonna take the stairs down to the second floor
R: okay