UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Not saying what's on your mind : how speakers avoid grounding references in privileged information

Permalink

<https://escholarship.org/uc/item/5kj3g7wr>

Author

Wardlow Lane, Liane Christine

Publication Date

2007

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Not saying what's on your mind:

How speakers avoid grounding references in privileged information

A Dissertation submitted in partial satisfaction of the requirements for the degree

Doctor of Philosophy

in

Psychology

by

Liane Christine Wardlow Lane

Committee in charge:

Professor Victor S. Ferreira, Chair
Professor Jeff Elman
Professor Tamar Gollan
Professor Gail Heyman
Professor Craig C.R.M. McKenzie

2007

The Dissertation of Liane Christine Wardlow Lane is approved, and it is acceptable in quality and form for publication on microfilm:

Chair

University of California, San Diego

2007

DEDICATION

I dedicate this dissertation to my family, friends and colleagues. I could never have completed this work without your guidance and encouragement. And to John, Claire and Molly – thank you for reminding me to be happy.

TABLE OF CONTENTS

Signature page.....	iii
Dedication.....	iv
Table of contents.....	v
List of tables.....	vii
List of figures.....	viii
Acknowledgements.....	x
Curriculum Vita.....	xi
Abstract.....	xv
Chapter 1: Introduction.....	1
Chapter 2: Don't talk about pink elephants! Speakers' control over leaking private information during language production	15
Methods.....	21
Results.....	23
Discussion.....	25
Chapter 3: Speaker-external versus speaker-internal forces on utterance form: Do cognitive demands override threats to referential success?.....	33
Experiment 1	
Methods.....	43
Results.....	46
Discussion.....	48
Experiment 2	
Methods.....	51
Results.....	53
Discussion.....	54
Experiment 3	
Methods.....	57
Results.....	58
Discussion.....	59

Experiment 4	
Methods.....	61
Results.....	62
Discussion.....	63
Chapter 4: Be wary of paying people to keep secrets: Examining the use of incentives to conceal confidential information.....	84
Experiment 1	
Methods.....	93
Results.....	97
Discussion.....	100
Experiment 2	
Methods.....	102
Results.....	104
Discussion.....	107
Chapter 5: General Discussion.....	127

LIST OF TABLES

Chapter 4, Table 1: Experiment Framing. The original Wardlow Lane et al (2006) task was converted to the more naturalistic and practiced context of buying and selling items.....	114
Chapter 4, Table 2: Point Structure for Experiments 1 and 2.	115
Chapter 4, Table 3: Experiment 1. Number of contrasting trials on which 0, 1/2 or 1 point was earned according to participant role and pay-type.	116
Chapter 4, Table 4: Experiment 1. Average points earned by sellers and buyers on each commission trial according to whether a modifier was or was not used. Sellers earned more points than buyers when modifiers were not used. Sellers earned fewer points than buyers when modifiers were used.....	117
Chapter 4, Table 5: Experiment 2. Number of contrasting trials on which 0, 1/2 or 1 point was earned according to participant role, pay-type and incentive-type.....	118
Chapter 4, Table 6: Experiment 2. Average points earned by sellers and buyers on each commission trial according to whether a modifier was or was not used. Sellers earned more points than buyers on trials that did not include a modifier. Sellers earned fewer points than buyers on trials that did include a modifier across both incentive-type conditions	119

LIST OF FIGURES

Chapter 2, Figure 1: Example experimental set-up.....	29
Chapter 2, Figure 2: Example stimuli	30
Chapter 2, Figure 3: Percentage of target descriptions including foil-contrasting modifiers as a function of whether the foil and target were the same (contrasting trials) or different shapes (non-contrasting trials), and whether speakers were given instructions to conceal the hidden shape (conceal trials) or not (baseline trials). Error bars illustrate 95% confidence interval of the interaction by speakers.....	31
Chapter 3, Figure 1: Example experimental set up for Experiments 1 & 2.....	74
Chapter 3, Figure 2: Example experimental set up for the Privileged-Ground-Target condition (where some filler targets are in privileged ground) and the Privileged-ignorable target condition (where all targets are in common ground).....	75
Chapter 3, Figure 3: Experiment 1. Percentages of target descriptions including foil-contrasting modifiers as a function of whether the foil and target were the same (contrasting trials) or different shapes (non-contrasting trials), and whether speakers were required to name some privileged ground targets on filler trials. Error bars illustrate 95% confidence interval of the interaction by speakers.....	76
Chapter 3, Figure 4: Experiment 2. Percentages of target descriptions including foil-contrasting modifiers as a function of whether the foil and target were the same (contrasting trials) or different shapes (non-contrasting trials), and whether speakers were asked to identify privileged objects that had pair mates in common ground (privileged-pair trials) or that were unique to the set (privileged-unique trials). Error bars illustrate 95% confidence interval of the interaction by speakers.	77
Chapter 3, Figure 5: Example experimental set up for Experiments 3 and 4. Note that the binder was angled such that the instructions were only visible to the speaker.....	78
Chapter 3, Figure 6: Experiment 3. Percentages of target descriptions including foil-contrasting modifiers as a function of whether the foil and target were the same (contrasting trials) or different shapes (non-contrasting trials), and whether speakers were shown the target by a point (low salience trials) or through the use of a picture of the privileged object (high salience	

trials). Error bars illustrate 95% confidence interval of the interaction by speakers.....	79
Chapter 3, Figure 7: Experiment 4. Percentage of target descriptions including foil-contrasting modifiers as a function of whether the foil and target were the same (contrasting trials) or different shapes (non-contrasting trials), and whether speakers were shown the target by a point (low salience trials) or through the use of a picture of the privileged object (high salience trials). Error bars illustrate 95% confidence interval of the interaction by speakers.....	80
Chapter 4, Figure 1: Example experimental set-up for Wardlow Lane et al (2006).	121
Chapter 4, Figure 2: Example experimental stimuli for Experiments 1 and 2.....	122
Chapter 4, Figure 3: Experiment 1. Percentage of target descriptions including foil-contrasting modifiers as a function of whether the foil and target were the same (contrasting trials) or different products (non-contrasting trials), and whether sellers were given instructions to conceal the confidential product (commission trials) or not (salary trials). Error bars illustrate 95% confidence interval of the interaction.....	123
Chapter 4, Figure 4: Experiment 2. Percentage of target descriptions as a function of whether trials included foil-contrasting modifiers (contrasting trials) or not (non-contrasting) and whether sellers were given instructions to conceal the confidential product (commission trials) or not (salary trials) and whether they were given a lottery payment incentive or a course credit incentive. Error bars illustrate 95% confidence interval of the interaction.....	124

ACKNOWLEDGEMENTS

Chapter 2, in full, is a reprint of the material as it appears in Wardlow Lane, L., Groisman, M. & Ferreira, V.S. (2006), Don't talk about pink elephants! Speakers' control over leaking private information during language production. *Psychological Science*, 17(1), 273- – 277. Reprinted with permission from Association for Psychological Science. The dissertation author was the primary investigator and author of this paper.

Chapter 3, in full, has been submitted for publication of the material as it appears in 2007, Wardlow Lane, L. & Ferreira, V.S. Speaker-external versus speaker internal forces on utterance form: Do cognitive demands override threats to referential success? The dissertation author was the primary investigator and author of this paper.

Chapter 4, in full, has been submitted for publication of the material as it appears in 2007, Wardlow Lane, L. & Liersch, M.J. Be wary of paying people to keep secrets: Examining the use of incentives to conceal confidential information. The dissertation author was the primary investigator and author of this paper.

CURRICULUM VITAE

EDUCATION

Ph.D., Psychology	University of California, San Diego	2007
M.S., Education	University of Southern California	2000
B.A., Linguistics	University of California, Santa Cruz	1995

AWARDS & HONORS

- National Institute of Health, Pre-Doctoral Training Grant, **Center for Research in Language, University of California, San Diego, 2006**
- **Conference Travel Award** – University of California, San Diego, Dept of Psychology, 2005
- **National Institute of Health, Pre-Doctoral Training Grant**, Center for Research in Language, University of California, San Diego, 2003
- University of Southern California’s Center for Excellence in Teaching, **Outstanding Teaching Assistant Award**, 2000
- Nominated to become the Teaching Assistant for USC president
- **Graduate Fellowship**, USC College of Letters, 1997
- **Scholarship**, Citizen’s Foundation, University of California, Santa Cruz, 1993

PUBLICATIONS

Wardlow Lane, L., Groisman, M. & Ferreira, V.S. (2006). Don’t talk about pink elephants! Speakers’ control over leaking private information during language production. *Psychological Science*, 17(4), 273-277.

Wardlow Lane, L. , Slevc, L.R. & Ferreira, V.S. (2007). Experimental methodology for eliciting stem exchange errors and syntactic category shifts. MIT Working Papers in Linguistics.

Slevc, L.R., **Wardlow Lane, L.** & Ferreira, V.S. (2007). Pronoun production: Word or world knowledge? MIT Working Papers in Linguistics.

IN PROGRESS

Wardlow Lane, L. & Ferreira, V.S. (Under Revision). Intrinsic versus extrinsic forces on the shapes utterances take: Do cognitive demands override threats to referential success?

Wardlow Lane, L. & Ferreira, V.S. (Under Revision). A frame driven account of syntactic flexibility in sentence production.

Wardlow Lane, L. & Liersch, M. (Submitted). Be wary of paying people to keep secrets: Examining incentives to conceal privileged information.

PRESENTATIONS

Wardlow Lane, L. & Ferreira, V.S. (2007). Intrinsic versus extrinsic forces on the shapes utterances take: Do cognitive demands override threats to referential success? The 20th Annual CUNY Sentence Processing Conference. Paper presented in La Jolla, CA.

Wardlow Lane, L., & Ferreira, V.S. (2006). How lexically based is syntactic production? The 12th Annual Architectures and Mechanisms for Language Processing Conference. Poster presented in Nijmegen, Netherlands.

Wardlow Lane, L., Groisman, M. & Ferreira, V.S. (2006). Don't talk about pink elephants! Speakers' control over leaking private information during language production. The 19th Annual CUNY Sentence Processing Conference. Paper presented in New York, NY.

Wardlow Lane, L., Groisman, M. & Ferreira, V.S. (2006, April). Don't talk about pink elephants! Speakers' control over leaking private information during language production. Center for Research in Language, University of California, San Diego.

Wardlow Lane, L. & Ferreira, V.S. (2005). A Frame-Based Model of Sentence Production. The State of the Art in Speech Error Research: A Workshop at the 2005 LSA Linguistic Institute. Poster Presented in Cambridge, MA.

Slevc, L.R., Wardlow Lane, L. & Ferreira, V.S. (2005). When pronouns are attracted to the wrong gender. The State of the Art in Speech Error Research: A Workshop at the 2005 LSA Linguistic Institute. Poster Presented in Cambridge, MA.

Wardlow Lane, L., Groisman, M. & Ferreira, V.S. (2005). Don't talk about pink elephants! Speakers' control over leaking private information during language production. The 11th Annual Architectures and Mechanisms for Language Processing Conference. Poster presented in Ghent, Belgium.

Wardlow, L. & Ferreira, V.S. (2005, January). A Frame-Driven Mechanism for Sentence Production. Center for Research in Language, University of California, San Diego.

Wardlow, L. & Ferreira, V.S. (2005, January). A Frame-Driven Mechanism for Sentence Production. Cognitive Psychology Discussion Group, University of California, San Diego.

Wardlow Lane, L. & Ferreira, V.S. (2004). A Frame-Based Model of Sentence Production. The 18th Annual CUNY Sentence Processing Conference. Poster Presented in Tucson, Az.

Wardlow, L. & Ferreira, V.S. (2003). Finding Common Ground: How do speakers block privileged information? The 9th Annual Architectures and Mechanisms for Language Processing. Poster presented in Glasgow, Scotland.

Wardlow, L. & Ferreira, V.S. (2003, May). Finding Common Ground: How good are speakers at blocking privileged information? Cognitive Psychology Discussion Group, University of California, San Diego.

TEACHING EXPERIENCE

Instructor , Writing and Critical Thinking, University of Southern California	1997-2000
---	-----------

Instructional Coordinator , Expository Writing University of Southern California	1999-2000
--	-----------

Teaching Assistant University of California, San Diego	2002-2006
--	-----------

Developmental Psychology Instructor: Dr. Monica Sweet	Winter 2006
--	-------------

Cognitive Psychology Instructor: Dr. David Peterzell	Fall 2005
---	-----------

Impulse Non-contrasting Disorders Professor: Dr. George Koob	Spring 2005
---	-------------

Psychology of Language Professor: Dr. David Swinney	Winter 2004
--	-------------

Infant Cognitive Development Professor: Dr. Leslie Carver	Fall 2004
--	-----------

Drugs, Addiction and Mental Disorders

Professor: Dr. George Koob

Fall 2003

Psychology of Language

Professor: Dr. David Swinney

Winter 2003

RELEVANT WORK EXPERIENCE

Curriculum Producer – Research & Development

Vivendi Universal, Houghton Mifflin (Educational Publishing)

2001-2002

Producer – Research & Development

Vivendi Universal, Knowledge Adventure (Educational Software)

2000-2001

Research & Development Associate

Neurosmith (Educational Toy Development)

1998-2000

PROFESSIONAL AFFILIATIONS

American Psychological Society

PROFESSIONAL SERVICE

Ad Hoc reviewer, *Language and Cognitive Processes*

Graduate Student Association representative, UCSD

Graduate Student Representative, Distinguished Teaching Committee, UCSD

ABSTRACT OF THE DISSERTATION

Not saying what's on your mind:

How speakers avoid grounding references in privileged information

by

Liane Christine Wardlow Lane

Doctor of Philosophy in Psychology

University of California, San Diego, 2007

Professor Victor S. Ferreira, Chair

This dissertation consists of three studies that examine whether speakers can adjust their speech when their knowledge differs from their addressee's. In the following work, information that speakers and addressees jointly know is categorized as *common ground* information; information that only one person knows is categorized as *privileged* information. To investigate speakers' referential behavior,

seven experiments employed a *referential communication* paradigm whereby speakers made direct reference to objects for addressees.

To be maximally effective, speakers should avoid making implicit references to privileged objects when making direct references to common ground objects. However, past reports have revealed that speakers often violate this standard (Wardlow Lane & Ferreira, 2003). The three reported studies examine influences on speakers' productions to reveal the forces that compel speakers to violate this standard.

Study 1 examines whether speakers can avoid making implicit references to privileged information with instruction. Results show that speakers make implicit references to privileged information more often when instructed not to, suggesting that speakers have little control over the degree to which they make such implicit references.

Study 2 examines the role of salience of privileged information. In a series of experiments, speakers were forced to cope with a speaker-internal (cognitive) pressure that made privileged information salient along with a speaker-external (communicative) pressure to avoid making implicit references to privileged information. Results show that speakers are more likely to make implicit references to privileged information when that information is salient even though doing so can result in communicative failure. This suggests that speaker-internal pressures exert the greatest influence on speakers' productions even at the cost of effective communication.

Study 3 examines the role of context and incentive on speakers' productions. The task was reframed as a buying-selling task. Results show that sellers made implicit references to privileged objects more often when provided with instruction and monetary incentive to avoid doing so. This suggests that grounding the task in a more practiced context and using a monetary incentive do not change speakers' productions. All three studies suggest that if speakers want to avoid mentioning something, their best tactic is to avoid thinking about it.

References

- Wardlow, L. & Ferreira, V.S. (2003). Finding Common Ground: How do speakers block privileged information? The 9th Annual Architectures and Mechanisms for Language Processing. Poster presented in Glasgow, Scotland

CHAPTER 1

INTRODUCTION

Put any two people together, involve them in a joint activity such as dialogue, and it will quickly become obvious that each person has knowledge that the other does not. Part of learning to communicate effectively is learning to take account of such knowledge differences (Nadig & Sedivy, 2002; Suzuki, 2002). Knowledge differences are particularly important with respect to referential communication, whereby speakers describe entities or pieces of information to addressees. Because addressees do not have access to speakers' intentions, their task is to try to determine what speakers are referring to based only upon words and context. One way addressees limit the set of possible referents that they think speakers might be describing is by judging what they consider to be most salient with respect to shared knowledge (i.e., common ground) (Clark, Schreuder, & Buttrick, 1983).

In order to speak cooperatively, speakers should abstain from inappropriately referring to knowledge that only they know (i.e., privileged ground) (Grice, 1975) since such references can cause communication to break down. Restricting speech to information that speakers know to be shared requires that speakers be able to efficiently take account of knowledge differences. However, determining which information is shared with an addressee may pose a large processing burden on speakers' production systems (Brown & Dell, 1987; Gary S. Dell & Brown, 1991; Horton & Keysar, 1996).

The language production system: Message encoding

Models of the language production system generally distinguish four stages: message encoding; grammatical encoding; articulation; and, a monitoring stage that operates when conditions and resources allow (Bock, 1982; Levelt, 1989). At message

encoding, speakers determine their communicative goals and engage in information selection. During information selection, speakers select the conceptual features (mental representations of pieces of information about the meanings of objects) that will allow them to communicate their intended message. Speakers map those conceptual features onto words and syntactic representations at the grammatical encoding stage, and then map those words onto sounds that are articulated at the articulation stage. If resources allow, the system will also monitor pre-verbal utterances before articulation to ensure that the to-be-articulated message maps onto the intended message.

The language production system, like many other cognitive systems, is often discussed using the notion of spreading activation (G. S. Dell, 1986; G. S. Dell, Burger, & Svec, 1997; MacKay, 1982; Stemberger, 1985). That is, at each stage of processing, representations accumulate activation, allowing them to influence processing at the following stages. These representations can become selected based upon their level of activation, with higher levels of activation making it more likely that they will be selected (G. S. Dell & O'Seaghdha, 1992). Upon being selected, a representation can undergo further processing (e.g., be bound to other representations, like a word can be inserted into a position in a sentence), and is typically assumed to receive a 'jolt' of activation that causes it to strongly influence processing at subsequent stages. The work described here, which is concerned specifically with referential communication, adopts this spreading activation approach, and incorporates the working assumption that higher levels of activation, no matter what the source, will make it more likely that a given representation will be selected and therefore be articulated.

Information Selection: Can speakers avoid references to privileged information?

With respect to activation of conceptual features, the production system cannot activate and encode just any conceptual features: it must activate and select the right conceptual features. Given the present issues, what that means is that the production system must activate only those conceptual features that both describe the intended referent, and that represent shared knowledge. The problem is that conceptual features can receive activation for a variety of reasons (or from a variety of sources), making it difficult, if not impossible, to limit activation to conceptual features that are shared. The question then becomes about intention and selection. Can speakers avoid selecting a highly salient and active conceptual feature if that feature is not an intended part of an utterance?

The work contained in this dissertation addresses the following questions: (1) What effect does the salience of privileged information have on speakers' perspective-taking behavior? Does increasing the salience of privileged information make it more likely that that information will become highly active and thus *more* likely to be mentioned? Or, does increasing the salience of privileged information highlight the fact that the information is privileged and thus make it *less* likely to be mentioned? (2) Can speakers limit their implicit references to privileged information when doing so is harmful to communication rather than when doing so is pragmatically awkward? (3) How, does the production system limit the intrusions of active privileged conceptual features in speakers' utterances? Can speakers use their real-world knowledge to help them determine which features should be mentioned? (4) Will increased effort to avoid making implicit references to privileged objects help speakers limit their utterances to

shared conceptual features? Or, will an increased effort to avoid making implicit references to privileged objects cause the representations of those objects to become highly activate, and thus more likely to be mentioned? The three studies (7 experiments) in this dissertation address these questions.

Study 1

Study 1 investigates whether active concepts that are not intended elements of speakers' utterances will nevertheless become part of speakers' utterances due to their activation. Study 1 uses a referential communication task whereby speakers are asked to describe mutually known objects for addressees. In the task, two naïve participants are shown four objects. Three of those objects are mutually visible, creating their common ground; a fourth object is visible to the speaker only, creating speakers' privileged ground. On critical trials speakers are instructed to describe a particular common ground object (hereafter, *target*) for addressees (e.g., a triangle). On *contrasting* trials the target is paired with a size-contrasting object of the same type in privileged ground (e.g., a smaller triangle). On *non-contrasting* trials the target is unique to the set (e.g., a square, so that the target triangle was the only triangle in the set). On all trials, speakers' use of size-contrasting modifiers is measured. When speakers use size-contrasting modifiers that differentiate the common ground target from the privileged ground object (e.g., when speakers call the target triangle 'large triangle'), then they have inappropriately used privileged information to ground a shared referent.

On half of trials in Experiment 1, in the *concealment* condition, speakers are given instruction and point-based incentive to avoid implicitly mentioning privileged information. On the other half of trials, speakers receive no special instructions about the

hidden object. This instruction and incentive about the privileged object in the concealment condition serves to make privileged information salient to speakers. This salience could have two possible outcomes. First, it could make the privileged status of that information more apparent to speakers. That is, the instruction and incentive could draw speakers' attention to the fact that they should avoid including mention of that privileged information in their references to targets. In contrast, the heightened salience of privileged information could result in speakers' mentioning that privileged information more often, even though they intend not to.

Results showed that making privileged objects highly salient, by instructing speakers not to mention those objects, results in speakers mentioning those objects more often. As such, the data suggest that highly active conceptual features can become part of speakers' utterances even though they are not marked as shared information and even though speakers do not intend to mention them.

Study 2

Study 2 directly examines the influence of the salience of privileged information on speakers' utterances by testing communicative and cognitive pressures. Whenever speakers produce an utterance, both communicative and cognitive pressures influence the form the utterance takes. Sometimes these *speaker-external* (communicative) and *speaker-internal* (cognitive) pressures work to opposite ends: Speaker-external pressures may dictate that speakers tailor utterances for addressees, whereas speaker-internal pressures can make tailoring utterances difficult. One way speakers can tailor utterances for particular addressees is by referring only to shared knowledge and avoiding references to information that only the speaker knows (i.e., *privileged* information;

Stalnaker, 1978). To what extent do speaker-internal cognitive pressures influence the degree to which speakers adhere to speaker-external communicative pressure to avoid references to privileged information?

In the *privileged-relevant* condition of Experiment 1, speakers are required to explicitly refer to privileged objects on some filler trials. In the *privileged-ignore* condition, there is no such requirement. Making speakers sometimes name the privileged object has two effects: First, it creates a speaker-external pressure to avoid referring to the privileged objects. If speakers name a common-ground heart *smaller heart* because they can see a privileged larger heart, then addressees cannot know whether *smaller heart* refers to the common-ground heart or to the privileged object – but only if privileged objects are sometimes named. Second, it draws attention to the privileged objects, because speakers know they sometimes have to name them. The question is, which pressure raised in the privileged-relevant condition wins: the speaker-external pressure to avoid referring to privileged objects or the speaker-internal pressure that makes privileged objects more salient? Results showed that speakers made implicit references to privileged information (when explicitly referring to common ground objects) more often in the privileged-relevant condition than in the privileged-ignore condition. Thus, speaker-internal pressure wins: Increasing the salience of privileged ground compels references to privileged objects, even though it is in this circumstance where implicit references to privileged objects can confuse addressees. Experiment 2 showed that this pattern did not arise because speakers were confused about the task.

Experiment 3 directly contrasted whether increasing the salience of privileged objects caused increased references to those objects. Trials were assessed under *low*

salience conditions, where an experimenter indicated to-be-named objects by pointing to them, or in *high salience* conditions, where speakers recognized to-be-named objects through instructions of the sort, “1 to the right of,” next to a *reference* object in the display, which on some trials was the privileged object. Results showed more references to privileged information on high-salience than low-salience trials. Lastly, Experiment 4 combined the attention manipulations of Experiments 1 and 3, and revealed them to *not* be independent. This suggests that the mechanisms responsible for references to privileged information in all experiments likely arose from the same mechanism. Overall, speakers appear particularly sensitive to speaker-internal pressures, such as are associated with increasing the salience (and activation) of privileged objects when designing their utterances.

Study 3

Study 3 investigates whether speakers’ poor performance in Studies 1 and 2, as measured by their seeming inability to avoid making implicit references to privileged information, was due to the abstract nature of the task. Past reports in cognitive psychology have shown that performance can be influenced by whether people can bring real-world knowledge to a task. One example is the Wason Card Sorting task (Wason, 1966, 1968). Wason provided participants with four two-sided cards. On each card was printed a letter on one side and a number on the other. Participants were also presented with a rule, such as, “If there is a vowel on one side, there is an odd number on the other”. Participants were instructed to turn over as few cards as possible in order to determine whether the rule was broken by the set of cards provided. Only a small percentage of participants solve this problem correctly (usually less than 10%; see Wason

& Johnson-Laird, 1972; Evans 1978; and, Johnson-Laird & Wason, 1977, for reviews of the card sorting task literature). Participants performed so poorly on this task that Wason argued that his experimental participants (undergraduate students) had not reached Piaget's stage of formal operational thought (Inhelder & Piaget, 1958).

However, Griggs and Cox (1982) used the same task, requiring the same logic, and found that participants performed quite well. Specifically, Griggs and Cox created a set of objects and a rule that allowed participants to use their real-world knowledge (or more specifically, memories of their personal history) to solve the problem. As an example, imagine seeing four cards, each with an age, either over or under 21-years, and a type of beverage, either alcoholic or non-alcoholic, on it. Accompanying these cards is the rule, "If a person is under 21-years of age, he/she cannot have an alcoholic beverage". Participants who encountered problems of this type solved the problem correctly at a much higher rate (about 73%) than in the original Wason task. Experiment 1 of Study 2 uses the same logic. Specifically, just as the original Wason task was reframed, I reframed the task reported in Study 1 so that it was a buying-selling exchange task. Therefore, the task now represented a situation that participants would have encountered in their daily life. Experiment 1 asks whether participants can use their past experiences and real-world knowledge to help control their implicit references to privileged information when given instruction and non-monetary incentive to do so.

Although the task itself was the same (i.e., speakers named objects for a listener), now speakers were told that they were sellers and addressees were told that they were buyers. Any time that a seller could name an object such that a buyer could pick it out of

the set, a transaction (i.e., a “sale”) had been made. Participants were told that the hidden object (the sellers’ privileged information) was the object that a competitor was selling.

On half of trials In Experiment 1, sellers were instructed and provided non-monetary incentive to avoid making implicit references to privileged information. On the other half of trials, no such instruction was provided. Results showed that sellers were more likely to make implicit references to privileged information when provided with instruction and incentive to avoid doing so, replicating the results from Study 1 in a more practiced context that allowed participants to bring their real-world knowledge to bear on the task.

The design of Experiment 2 replicated the first experiment except that it included an additional manipulation of incentive-type. On half of trials, sellers were provided with a lottery payment incentive. They were instructed that the top-performing seller would receive \$100. On the other half of trials, sellers were provided with a course credit incentive. They were instructed that they would receive course credit for their participation. If the results of Study 1 and Study 2 Experiment 1 were due to a lack of effort on the part of sellers, we should see sellers in the lottery payment condition avoid making implicit reference to privileged information more often than sellers in the course credit condition.

Results showed that in the lottery payment condition sellers were more likely to make implicit reference to privileged information than in the course credit condition. The results of both experiments suggest that (1) speakers’ references to privileged information are not due to the abstract nature of the task employed or a lack of effort in designing utterances; and, (2) speakers make implicit references to privileged information because

that information is made salient to them, not because of an intention to mention that information.

Together, the results of the three reported studies suggest that speaker-internal cognitive forces, and in particular, salience, influence speakers' ability to take their addressees' knowledge into account when speaking. Specifically, the more highly salient privileged information is made for speakers, and thus, the more highly activated it becomes, the more likely that information is to be selected for production. The result is that speakers are more likely to implicitly mention that information even if they intend not to.

References

- Bock, J. K. (1982). Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological Review*, 89, 1-47.
- Brown, P. M., & Dell, G. S. (1987). Adapting production to comprehension: The explicit mention of instruments. *Cognitive Psychology*, 19, 441-472.
- Clark, H. H., Schreuder, R., & Buttrick, S. (1983). Common ground and the understanding of demonstrative reference. *Journal of Verbal Learning and Verbal Behavior*, 22, 245-258.
- Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, 93, 283-321.
- Dell, G. S., & Brown, P. M. (1991). Mechanisms for listener-adaptation in language production: Limiting the role of the "model of the listener." In E. Donna Jo Napoli, E. Judy Anne Kegl & et al. (Eds.), *Bridges between psychology and linguistics: A swarthmore festschrift for lila gleitman*. (pp. 105-129): Hillsdale, NJ, USA.
- Dell, G. S., Burger, L. K., & Svec, W. R. (1997). Language production and serial order: A functional analysis and a model. *Psychological Review*, 104(1), 123-147.
- Dell, G. S., & O'Seaghdha, P. G. (1992). Stages of lexical access in language production. *Cognition*, 42, 287-314.
- Evans, J. St B. T. (1978). The psychology of deductive reasoning: Logic. In A. Burton & J. Radford (Eds.), *Thinking in Perspective*. London: Methuen.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics: Speech act* (Vol. 3, pp. 41-58). New York: Academic Press.
- Griggs, R.A. & Cox, J.R. (1982). The Elusive Thematic-Materials Effect in Wason's Selection Task. *British Journal of Psychology*, 73 (3), 407-20.

- Horton, W. S., & Keysar, B. (1996). When do speakers take into account common ground? *Cognition*, 59(1), 91-117.
- Inhelder, B. & Piaget, J. (.1958). *The growth of logical thinking from childhood to adolescence.*, NY, NY: Basic Books.
- Johnson-Laird, P.N., Legrenzi, P. & Sonino Legrenzi, M. (1972). Reasoning and sense of reality. *British Journal of Psychology*, 63, 395-400.
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- MacKay, D. G. (1982). The problems of flexibility, fluency, and speed-accuracy trade-off in skilled behavior. 89, 483-506.
- Nadig, A. S., & Sedivy, J. C. (2002). Evidence of perspective-taking constraints in children's on-line reference resolution. *Psychological Science*, 13(4), 329-336.
- Stalnaker, R. C. (1978). Assertion. In P. Cole (Ed.), *Syntax and semantics 9: Pragmatics* (pp. 315-332.). New York, NY: Academic Press.
- Stemberger, J. P. (1985). An interactive activation model of language production. In A. Ellis (Ed.), *Progress in the psychology of language* (Vol. 1, pp. 143-186). London: Erlbaum.
- Suzuki, T. (2002). Perspective-taking and comprehension of passive sentences by japanese-speaking children. *Journal of Psycholinguistic Research*, 31(2), 131-144.
- Wardlow, L., & Ferreira, V. S. (2003). *Finding common ground: How do speakers block privileged information?* Paper presented at the AMLaP 2003, Glasgow, Scotland.
- Wardlow Lane, L., Groisman, M. & Ferreira, V.S. (2006). Don't talk about pink elephants! Sellers' control over leaking private information during language production. *Psychological Science*, 17 (4), 273-277.

- Wason, P.C., & Shapiro, D. (1971). Natural and contrived experience in a reasoning problem. *Quarterly Journal of Experimental Psychology*, 23, 63-71.
- Wason, P.C. (1966). Reasoning. In B. Foss (Ed.), *New Horizons in Psychology*. London: Penguin.
- Wason, P.C. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology*, 20, 273-281.
- Wason, P.C. & Johnson-Laird, P.N. (1972). *Psychology of Reasoning: Structure and Content*. Cambridge, MA: Harvard University Press.

CHAPTER 2

DON'T TALK ABOUT PINK ELEPHANTS! SPEAKERS' CONTROL OVER
LEAKING PRIVATE INFORMATION DURING LANGUAGE PRODUCTION

Abstract

Speakers' descriptions sometimes inappropriately refer to information known only to them, thereby 'leaking' knowledge of that private information. We evaluated whether speakers can explicitly control such leakage in light of its communicative consequences. Speakers described mutually known objects (e.g., a triangle) that had size-contrasting matches that were privileged to the speaker (e.g., a larger triangle visible to the speaker only), so that use of a contrasting adjective (i.e. "small") involves referring to the privileged information. Half the time, speakers were instructed to conceal the identity of the privileged objects. If speakers can control their leaked references to privileged information, the conceal instruction should make such references less likely. Surprisingly, the conceal instruction caused speakers to refer to privileged objects more. Thus, speakers not only have difficulty not leaking privileged information, but attempts to avoid such leakage only makes it more likely.

Introduction

Though we nearly always bring our own perspective to any given situation, sometimes we behave as though we fail to appreciate that fact. This has been explored experimentally in tasks like that illustrated in Figure 1. Here, four objects are positioned between two people. One person can see three of them: a triangle, circle, and heart. The other person can additionally see a fourth object -- a larger triangle. If the second person is asked to identify the mutually visible triangle so that the first person can pick it out, she or he ought to say “triangle,” just as “circle” would describe the sole circle. Yet, sometimes speakers in this circumstance say “small triangle” instead (Horton & Keysar, 1996; Nadig & Sedivy, 2002; Wardlow & Ferreira, 2003), as if they fail to appreciate their unique perspectives.

Figure 1 About Here

Two classes of factors might affect the likelihood that speakers will disregard their knowledge of perspective differences and produce seemingly erroneous utterances like “small triangle” in this situation. One class includes low-level factors that compel speakers to pay more or less attention to the shape that only they can see (hereafter the hidden shape). (For accounts of how factors like these might operate, see Horton & Keysar, 1996; Nadig & Sedivy, 2002.) Specifically, too much attention to the hidden shape may boost its salience overwhelming the knowledge that it is hidden, leading speakers to refer to it when labeling the to-be-described (target) shape. To the extent that such low-level factors are influential, utterances like “small triangle” are like “Simon says” errors, whereby undue attention to the hidden shape

compels speakers to refer to it, even though they should not and even if they may not intend to.

The other class of factors includes knowledge of the high-level communicative consequences of producing such errors (Clark, 1996; Schober & Brennan, 2003). Note that when speakers say “small triangle” instead of “triangle,” not only have they communicated which shape their addressee ought to pick out (i.e., the mutually visible triangle), they have also potentially leaked *implicit* information. In particular, the addressee can infer that the speaker can probably see another triangle, which in the illustrated situation is likely to be the hidden shape. In most situations, this leaked information is unlikely to harm speakers’ communicative goals (Nadig & Sedivy, 2002); here, if the speaker aims to convey which triangle the addressee should select, “small triangle” works about as well as “triangle” (after all, the addressee can only see one triangle), and the leaked ‘bonus’ information is largely irrelevant. Indeed, it might be that producing utterances like “small triangle” is generally adaptive, because it allows speakers to communicate more information with fewer words (compare “small triangle” to “triangle, and I can also see a larger triangle”).

But what happens when leaked information conflicts with speakers’ goals? Assume that in a situation like in Figure 1, speakers are instructed not only to name the target shape, but also to conceal the hidden shape. Now speakers should avoid describing the target as “small triangle,” because the leaked information might cue addressees to the identity of the hidden shape. Can speakers’ high-level

communicative intentions (to name the target and conceal the hidden shape) overcome their basic tendency to sometimes violate their knowledge of perspective differences? Or are the low-level factors (e.g., salience) that compel speakers to produce utterances like “small triangle” not subject to control by speakers’ intentions?

A third possibility, which comes out of research on Ironic Processes (Wegner, 1994), suggests that speakers may actually provide *more* information about the hidden object precisely because of their intention to conceal it. Ironic Processes suggests a dual-process for performing actions: an operator process and a monitor process. While the operator attempts to create the desired state, the monitor checks the system for signs of failure. The operator thus brings failure conditions into consciousness, which can ironically cause behaviors that fulfill those conditions, thereby causing a precisely counter-intentional behavior. For example, subjects instructed to hold a pendulum steady when under cognitive load will swing it along a particular axis more when given emphatic instructions not to do just that. So here the instruction to conceal could invoke a monitor which could then lead to *more* references to the hidden object.

The experiment used a referential communication task (Hanna, Tanenhaus, & Trueswell, 2003; Keysar, Barr, Balin, & Brauner, 2000) like that illustrated in Figure 1, where speakers described mutually visible shapes on *target* cards while trying to ignore hidden shapes on *foil* cards. On *critical* trials, the object on the target card was medium-sized (see Figure 2). On half of the critical trials (*contrasting* trials), foils

and targets were the same shape, but contrasted in size. Thus, contrasting trials were designed to elicit utterances that included modifiers that contrasted the target with the hidden shape. On the other half of critical trials (*non-contrasting* trials), the foil was a different shape from the target. Non-contrasting trials thus assessed how often utterances included modifiers irrespective of the contrast to the hidden shape.

Figure 2 About Here

Speakers were tested in two order-counterbalanced blocks. In *baseline* blocks, speakers were instructed to identify the target so that the addressee could select it from the mutually visible set. Speakers and addressees kept scores, each receiving one point whenever addressees successfully selected the target. In *conceal* blocks, participants were given additional instructions that encouraged speakers to hide the identity of the foil when identifying the target. Specifically, after addressees selected the target, they were allowed to guess the identity of the foil. A point was added or subtracted from addressees' scores if they guessed correctly or incorrectly (respectively). Speakers were instructed not to allow addressees to gain additional points. Therefore, speakers should avoid any behavior that might cue the identity of the hidden shape such as producing modified utterances like "small triangle," because the modifying adjective could cue the identity of the foil.

Performance in the conceal condition should determine whether speakers can control the production of leaked information in light of its communicative consequences. If so, then the instruction to conceal the hidden shape should reduce the mention of foil-contrasting modifiers relative to performance in the baseline

block. However, if information is leaked as an uncontrollable consequence of low-level factors such as attention increasing the salience of the hidden shape, then the instruction to conceal that hidden shape should not decrease modifier production. In fact, the conceal instruction might actually increase its salience; if so, speakers might use modifiers even more in the conceal block than in the baseline block.

Methods

Participants. Participants were 88 undergraduates at the University of California, San Diego, 44 as speakers and 44 as addressees. All participants were native speakers of English.

Materials and Design. Participants were tested on 288 cards. Each card displayed one simple line drawing of a familiar object. The objects varied in size across and within trials such that the actual size of a large (or small) object (relative to the size of the other objects in that set) on one trial might be small (or large) on another trial. Target cards were paired with foils that were either larger or smaller than the target. Each object type was used only on one trial and never occurred with more than one other object of the same type.

Two manipulations were used: contrast-type (contrasting versus non-contrasting) and concealment (conceal versus baseline). On contrasting trials, the privileged object on the hidden card contrasted in size with the to-be-named mutually visible object. On non-contrasting trials, the to-be-named mutually visible object was unique. On conceal trials, speakers were told not to provide addressees with any

information about the hidden shape. On baseline trials, speakers were not given any special instructions regarding the hidden shape.

Four experimental conditions were assigned to each critical object by crossing the levels of contrast type and concealment. Both factors were manipulated within speakers and items in counterbalanced fashion. Subjects were presented with 36 critical trials, half contrasting and half non-contrasting. Additionally, half were in the conceal condition and half were in the baseline condition. Conceal and baseline trials were blocked, with the order of blocks counterbalanced across subjects. Thirty-six additional filler trials, half composed of two pairs of cards, and half composed of one pair and two unique cards, were administered. Speakers were asked to identify a mutually visible shape on all trials.

Procedures. A coin toss randomly assigned participants to the roles of speaker and addressee. Participants sat at opposite sides of a table where the speaker could see a computer monitor, but the addressee could not. At the beginning of each trial, the addressee was instructed to close his or her eyes while the experimenter placed four cards on the table. The speaker then looked at a computer monitor that displayed a schematic of the four blank cards, one of which had an arrow above it with the instruction, *Block this card*. The speaker blocked the corresponding actual card by positioning an occluder between the card and the addressee so that the addressee could not see the card. Next, the speaker looked back at the computer screen, which showed the four-card schematic with an arrow pointing at a different card with the instruction, *Identify this card*. The speaker was instructed to describe the

corresponding card with just enough information so that the addressee could identify it. Upon hearing the speaker's description, the addressee opened his or her eyes and attempted to identify the target card. If the addressee identified the correct card, each participant received one point; otherwise neither participant received a point.

On conceal trials, addressees were told that after trying to identify the described card, they could guess the identity of the hidden shape. If addressees guessed correctly, they received an extra point, whereas if they guessed incorrectly, they lost a point. Participants were instructed to individually earn as many points as possible, and on conceal blocks, speakers were instructed not to allow addressees to receive extra points.

Results. Target descriptions on all trials were transcribed and coded for whether they included a modifier that contrasted the size of the target with that of the foil used on contrasting trials (e.g., the description of the medium-sized triangle was coded for whether it was described as “small triangle” both on contrasting and non-contrasting trials). The proportions of targets described in each experimental condition with such modifiers were computed for each subject. Statistics were conducted on these proportions with repeated-measures analyses of variance (ANOVAs) using subjects ($F1$) and items ($F2$) as random factors (analyses carried out using arcsine-transformed proportions yielded the same pattern of significance). The ANOVA design was 2 x 2, with the factors contrast type and concealment. We used planned comparisons to assess performance on contrasting versus non-contrasting trials separately under each concealment instruction. All significant

effects achieved the .05 level unless otherwise specified. We report variability with repeated-measures 95% confidence-interval halfwidths (CIs) based on single degree-of-freedom comparisons (Loftus & Masson, 1994).

Figure 3 shows the mean percentages of target descriptions that included specified modifiers (i.e., “small triangle”) as a function of contrast type and concealment instruction. As expected, speakers produced more modifying adjectives overall on contrasting trials (10%), than on non-contrasting trials (0.9%), $F(1,43) = 16.3$, $CI = \pm 4.5\%$; $F(1,35) = 35.9$, $CI = \pm 3.0\%$, $\eta^2 > .275$, showing that modifiers were used specifically in response to the size contrast. Surprisingly, speakers also produced more modifying adjectives overall in conceal blocks (7.9%) than baseline blocks (3.0%), $F(1, 43) = 6.5$, $CI = \pm 3.9\%$; $F(1,35) = 10.1$, $CI = \pm 3.3\%$, $\eta^2 > .131$. In fact, speakers produced 13% more adjectives on contrasting trials (14.4%) than on non-contrasting trials (1.4%) in the conceal block, but only 4.9% more adjectives on contrasting trials (5.4%) than on non-contrasting trials (0.5%) in the baseline block, leading to a significant interaction, $F(1, 43) = 5.3$, $CI = \pm 5.1\%$; $F(1,35) = 6.4$, $CI = \pm 4.4\%$, $\eta^2 > .275$. Planned comparisons revealed both of these differences to be significant (with the baseline comparison marginal by subjects), $F(1,43) = 26.9$, $F(1,35) = 34.1$ and, $F(1,43) = 3.8$, $p < .06$, $F(1,35) = 5.0$, respectively. In short, speakers tend to modify target descriptions with respect to hidden information, but they do so even more when instructed to conceal the hidden information.

Figure 3 About Here

Thus, at least under these task conditions, speakers were unable to control whether they leaked hidden information as a function of the communicative consequences of that reference. In fact, the opposite was observed: When speakers were provided with instructions and incentives not to leak information about a hidden shape, they were even more likely to do so. These results support the idea that when speakers fail to account for their unique perspective, it is because relatively autonomous low-level processes sometimes cause privileged information to be unintentionally incorporated into their descriptions.

Discussion. Speakers' lack of communicative control over leaking privileged information carries theoretical and practical implications. Concerning theory, models of language production distinguish conceptual processing and grammatical encoding (Bock, 1982; Levelt, 1989). Within such accounts, it is at conceptual processing that speakers encode their communicative intentions and define a to-be-conveyed message. That message consists of those conceptual features that are accessible and whose production will allow speakers to achieve their communicative objective. Importantly, additional conceptual features that are not needed to convey the intended message may also be accessible (e.g., contextually activated but unimportant information, or relevant but private information). The question here amounts to whether grammatical encoding processes encode only those conceptual features that comprise the intended message, or whether they can also encode features that, though accessible, are not intended to be expressed. Given the results, we suggest that being part of a communicative intention is not a necessary condition for grammatical

encoding to be sensitive to an accessible conceptual feature.

These results also carry important implications for our understanding of privacy and communication. Consider the relationship between the present results and the well-known observation that when directed, “don’t think of a pink elephant,” people inevitably do just that. The latter observation illustrates that people do not have total explicit control over what thoughts come to mind. The present results incorporate this observation, as the direction to conceal the hidden shape evidently only made that hidden shape more salient. However, the present results go further, by showing that when directed to conceal the hidden shape, speakers were more likely not only to *think* of it, but also to *refer* to it. The observation that private information is sometimes leaked despite explicit attempts to avoid doing so suggests not only that leaked information may sometimes be information speakers might want to keep private, but that attempts to conceal that private information might make its leakage even more likely. If so, these results are likely to be relevant to interactions involving everything from interpersonal interactions to adversarial negotiation.

Acknowledgements

This research was supported by National Institutes of Health grant R01 MH64733.

The authors express appreciation to Kristy Lawson for assistance with data collection and Bob Slevc and Tamar Gollan for helpful discussions.

Chapter 2, in full, is a reprint of the material as it appears in Wardlow Lane, L., Groisman, M. & Ferreira, V.S. (2006). Don't talk about pink elephants! Speakers' control over leaking private information during language production. *Psychological Science*, 17(1), 273-277. Reprinted with permission from the Association for Psychological Science.

Figure Captions

Figure 1. Example experimental set up.

Figure 2. Example stimuli.

Figure 3. Percentage of target descriptions including foil-contrasting modifiers as a function of whether the foil and target were the same (contrasting trials) or different shapes (non-contrasting trials), and whether speakers were given instructions to conceal the hidden shape (conceal trials) or not (baseline trials). Error bars illustrate 95% confidence interval of the interaction by speakers.

Chapter 2, Figure 1
Example experimental set up.

Contrasting Trials:

Identify this card

Identify this card

Non-contrasting Trials:

Identify this card

Identify this card

Chapter 2, Figure 2
Example stimuli.

Chapter 2, Figure 3

Percentage of target descriptions including foil-contrasting modifiers as a function of whether the foil and target were the same (contrasting trials) or different shapes (non-contrasting trials), and whether speakers were given instructions to conceal the hidden shape (conceal trials) or not (baseline trials). Error bars illustrate 95% confidence interval of the interaction by speakers.

References

- Bock, J. K. (1982). Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological Review*, 89, 1-47.
- Clark, H. H. (1996). *Using language*. Cambridge, England: Cambridge University Press.
- Hanna, J. E., Tanenhaus, M. K., & Trueswell, J. C. (2003). The effects of common ground and perspective on domains of referential interpretation. *Journal of Memory & Language*, 49(1), 43-61.
- Horton, W. S., & Keysar, B. (1996). When do speakers take into account common ground? *Cognition*, 59(1), 91-117.
- Keysar, B., Barr, D. J., Balin, J. A., & Brauner, J. S. (2000). Taking perspective in conversation: The role of mutual knowledge in comprehension. *Psychological Science*, 11(1), 32-38.
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review*, 1, 476-490.
- Nadig, A. S., & Sedivy, J. C. (2002). Evidence of perspective-taking constraints in children's on-line reference resolution. *Psychological Science*, 13(4), 329-336.
- Schober, M. F., & Brennan, S. E. (2003). Processes of interactive spoken discourse: The role of the partner. In A. C. Graesser & M. A. Gernsbacher (Eds.), *Handbook of discourse processes* (pp. 123-164). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.
- Wardlow, L., & Ferreira, V. S. (2003). Finding common ground: How do speakers block privileged information? The AMLaP Conference, Glasgow, Scotland.
- Wegner, D. M. (1994). Ironic processes of mental control. *Psychological Review*, 101(1), 34-52.
- Wegner, D. M., Ansfield, M., & Pilloff, D. (1998). The putt and the pendulum: Ironic effects of the mental control of action. *Psychological Science*, 9(3), 196-199.

CHAPTER 3

SPEAKER-EXTERNAL VERSUS SPEAKER-INTERNAL FORCES ON UTTERANCE FORM: DO COGNITIVE DEMANDS OVERRIDE THREATS TO REFERENTIAL SUCCESS?

Abstract

To what extent do speaker-external communicative pressures versus speaker-internal cognitive pressures affect utterance form? In four experiments, we measure speakers' references to privately known (i.e., *privileged*) objects when naming mutually known (i.e., *common ground*) objects. Although speaker-external communicative pressures demand that speakers avoid inappropriate references to privileged objects, the first two experiments show that speakers often ignore this demand when it co-exists with a speaker-internal pressure to attend to those privileged objects. We hypothesize that our implementation of the speaker-external demand acted to increase the salience of privileged objects (a speaker-internal pressure). Experiment 3 shows directly that boosting the salience of privileged objects does increase the likelihood that speakers will inappropriately refer to those objects. Finally, Experiment 4 suggests that the mechanism responsible for increasing speakers' references to privileged objects in Experiments 1 and 2 is the same as the mechanism causing such references in Experiment 3. Together, the results suggest that the language production system is especially sensitive to cognitive pressures even when such pressures cause utterances to violate communicative requirements about the forms utterances should take.

Introduction

The form any particular utterance takes is the result of pressures from sources both external to and internal to the speaker. Speaker-external pressures are communicative in nature – for example, how speakers’ word choices convey different messages to addressees. Speaker-internal pressures are cognitive in nature – for example, how speakers’ word choices are influenced by the allocation of attention. Often, speaker-external and speaker-internal pressures work to similar ends; for example, allocation of attention might compel speakers to make word choices that are optimally tailored to particular addressees (i.e., if a speaker’s attention is drawn to something, their addressee’s attention likely is as well, so that commenting on attention-drawing elements is likely communicative). Sometimes, however, speaker-external and speaker-internal pressures work to opposite ends, when speaker-external communicative pressures dictate that speakers design utterances tailored for addressee needs, whereas speaker-internal cognitive pressures can make production of tailored utterances difficult.

The process of tailoring utterances to addressee needs is generally referred to as *audience design* (Clark, 1996; Clark & Carlson, 1982; Clark & Murphy, 1982). Exactly how interlocutors engage in audience design is controversial (Brown & Dell, 1987; Horton & Gerrig, 2002; Horton & Keysar, 1996; Keysar & Horton, 1998; Lockridge & Brennan, 2002; Polichak & Gerrig, 1998b; Schober & Brennan, 2003). However, what is clear is that speakers do take the specific needs of their addressees into account in some ways. For example, adults use exaggerated prosody when

speaking to children and non-native speakers (DePaulo & Coleman, 1986; Fernald & Mazzie, 1991); when giving directions to a location speakers provide more information if the direction seeker appears to be from out-of-town (Kingsbury, 1968 as cited in Krauss & Fussell, 1991); speakers revise their content when provided with evidence that addressees are having difficulty understanding (Garvey, 1975); speakers continue to use the same term to refer to the same entity even if the term is overly specific in a new context (Brennan & Clark, 1996); and speakers modify their utterances when they estimate that their addressees have more limited processing capacity such as when speaking to cognitively-impaired addressees (DePaulo & Coleman, 1986).

Another way speakers tailor utterances for addressees is by taking into account whether the to-be-discussed information is information that addressees know about. The information that both speakers and addressees both know makes up *common ground* (Clark & Marshall, 1981), and the information that only one person knows makes up that person's *privileged ground* (Stalnaker, 1978). To communicate effectively, speakers should adhere to a speaker-external communicative demand to avoid inappropriate references to privileged information (Grice, 1975). For example, imagine a situation where speakers can see two hearts, one of which is the intended referent or *target*. Addressees, however, can only see one heart. If speakers say something of the form, "It's the heart on the right," the result may be confused addressees who are wondering whether the heart they do know about is the target. For

speakers to tailor their speech to an addressee's knowledge, they must keep track of common versus privileged information.

But it is not clear *how* and *when* the language production system is able to incorporate such information. To inform these questions, in the present experiments, we contrast two types of influences on speakers' ability to account for common versus privileged ground: speaker-internal (i.e., cognitive) pressures, and speaker-external (i.e., communicative) pressures. We look at the extent to which each of these pressures might enable speakers to avoid inappropriate references to privileged information. This distinction between speaker-internal and speaker-external influences on utterance formation is critical to all theories of common ground processing, the specifics of which we address in the General Discussion. Our approach is to delineate the relative roles of these two types of information, using the results to inform the viability of current theoretical approaches.

Here, we use a referential communication task to set up a situation in which the external pressure should lead speakers to tailor their utterances to take into account their addressees' knowledge. Referential communication refers to that aspect of communication whereby speakers describe entities (referents) so that they can be selected from some set of possibilities (e.g., one person from all possible people in the world). The goal of referential communication is to successfully refer: to produce an utterance whose form allows addressees to determine exactly which referent, out of all referents in some implicit or explicit contrast set, is intended. Following the above example, assume that a speaker and an addressee share knowledge of the

existence of one heart (i.e., the referent). However, the speaker also has privileged knowledge of another heart. In order to make reference to the common ground heart, the speaker should say something of the form, “It’s the heart” – after all, the addressee only knows of one heart. If instead the speaker says something of the form, “It’s the heart on the right,” contrasting the location of the common ground and privileged hearts, addressees may be confused. Under certain circumstances (to be detailed below), an inappropriate reference to privileged information such as this may actually cause referential failure.

Everyday communication generally succeeds, suggesting that speakers are quite good at taking common and privileged ground into account in designing referential expressions. This contrasts with laboratory demonstrations which show that speakers do not reliably exclude privileged information from their utterances (Horton & Keysar, 1996; Nadig & Sedivy, 2002; Wardlow Lane, Groisman & Ferreira, 2006). How can the inconsistency between what appears to happen in everyday conversation and these laboratory demonstrations be reconciled? One possibility is that everyday conversation (but not laboratory settings) is situated in contexts that enable ready computation of common ground (Clark & Marshall, 1981) lessening the speaker-internal cognitive demands of common ground processing. However, there is at least some evidence that even when laboratory situations are made quite naturalistic, and specifically when computation of common ground is made simple, speakers occasionally but consistently include privileged information in their utterances at a fairly high rate (Wardlow Lane et al, 2006; baseline condition).

Another explanation for why speakers refer to privileged information in naturalistic conversation might be that they only do so when speaker-external communicative pressures are minimal. For instance, perhaps speakers refer to privileged information when doing so will not harm communicative success (and therefore it largely goes unnoticed). The point of this paper is to investigate this last possibility: Is speakers' communicative efficacy sensitive to speaker-external communicative demands that require them to avoid referring to privileged information when pitted against speaker-internal demands that cause them to attend to privileged information?

To test this question, we set up a situation where inclusion of privileged information (in references to common ground information) threatens communicative success. One such situation (see Figure 1), which will be detailed here, was previously tested in Wardlow Lane, Groisman and Ferreira (2006). Two naïve participants, a speaker and an addressee, were shown four line drawings of objects. Both participants could see three of those objects (i.e., they were in common ground). At the beginning of each trial, the speaker occluded the fourth object so that it was not visible to the addressee (i.e., the speaker placed a blank opaque picture frame in front of the object). As such, the fourth object's privileged ground status was naturalistically marked because it was visible only to the speaker. After occluding the privileged object, the speaker named a particular target object, which could have been any of the three common ground objects. On *contrasting* trials, the target was the same type of object as the privileged object, but contrasted in size (i.e., a common

ground heart with a size-contrasting privileged heart foil). On *non-contrasting* trials, the target was unique. Our measure was the percentage of time that speakers used size-contrasting modifiers (i.e., *big* or *small*) to refer to the target object on contrasting trials (e.g., when speakers called the target heart “small heart” because there was a larger heart in privileged ground). This was compared to the percentage of time that speakers used size-contrasting modifiers on non-contrasting trials (e.g., calling the target heart “a small heart” even when there was no other heart in the set of objects). To the extent that speakers used such adjectives more on contrasting than non-contrasting trials, they used privileged information when forming their referring expressions.

In the Wardlow Lane et al (2006) study, we altered speakers’ goals (what they were trying to accomplish [or avoid] by speaking). Specifically, speakers were told that on half of trials, addressees would be allowed to guess the identity of the privileged object and that speakers should do what they could to stop addressees from guessing correctly. Because references to targets that were grounded in privileged information (i.e., “small heart”) provided addressees with implicit information about the hidden object, such references should have been avoided. That is, if speakers said “small heart” to name a common ground heart when there was a larger heart in privileged ground, then they essentially hinted to addressees that the hidden object was likely a large heart. The experiment revealed that the instruction and incentive to avoid making reference to privileged information did not act to reduce the inclusion of such information. In fact, the opposite was observed: Speakers were *more* likely

to refer to privileged information (implicitly, by modifying target objects to contrast with the hidden object) when specifically instructed not to.

Why might speakers implicitly refer to privileged ground more when instructed not to? One explanation for this finding is that the production system may not be sensitive to the type of communicative pressure that was tested. In that experiment, speakers had two goals: A primary goal of successfully referring to targets; and, a secondary goal of not providing information about privileged objects. When speakers ground references to targets in privileged information (i.e., when they say “small heart” to name a common ground heart when there was a larger heart in privileged ground) their primary goal of conveying which was the target object was not threatened (because only mutually visible shapes could be targets); only their secondary goal of concealing the hidden shape was threatened. Importantly, the production system may not be particularly sensitive to whether inclusion of some feature in an utterance (such as a modifier) threatens a more remote (secondary) goal, especially when that goal is unrelated to the primary goal of successful reference.

In the current experiment, we manipulate communicative success in a more fundamental way. Here, speakers have only one goal: To successfully refer to a target object so that addressees can identify it. In one condition, only if speakers consistently avoid referring to privileged ground (e.g., avoid saying “small heart”) will they be able to use a descriptive strategy that can always successfully refer to target objects (details below). If speakers do sometimes refer to privileged ground, they will undermine this descriptive strategy and therefore refer to targets less

successfully. In another condition, we remove this threat, such that even when speakers ground references to targets in privileged information, they can still successfully refer. If speakers are sensitive to this external pressure (a threat-to-successful-reference), then they should make privileged ground references less often when doing so can harm their successful referential strategy. In this way, we have made the more general communicative pressure to avoid inappropriate references to privileged information primary. Now, instead of only (perhaps) causing momentary confusion because of an inappropriate reference to privileged information, such references can actually result in utterances whose form does not allow addressees to identify the intended target.

Another possible explanation of the Wardlow Lane et al (2006) result is that the production system may be particularly sensitive to speaker-internal (i.e., cognitive) pressures. Specifically, the instruction and incentive to avoid mentioning privileged information may have inadvertently heightened speakers' attention to privileged objects, and such heightened attention might have led to increased references to those privileged objects. This raises an interesting possibility: By setting up a task so that grounding references to targets in privileged information will cause referential failure, we can concurrently increase attention to privileged information. The question of interest is, what does such heightened attention to privileged information do? Does it highlight the fact that the information is in fact privileged, thereby causing speakers to even better respect the external pressure to avoid referring to the privileged information? If so, this would suggest that when speakers

in general refer to privileged information, it is not because a speaker-internal attentional pressure is compelling them to, but rather because external pressures are not sufficient to compel speakers to avoid referring to privileged information. The alternative possibility is that heightened attention to privileged information might increase the likelihood that that privileged information will be referred to, even when doing so risks referential failure. If so, this would suggest that the speaker-internal pressure can override the influence of the speaker-external pressure.

In Experiment 1, we include both a speaker-external communicative pressure and a speaker-internal cognitive pressure on utterance form that work to opposite ends. Specifically, we include a threat-to-referential success that constitutes a speaker-external pressure to avoid including privileged information in utterances and a speaker-internal pressure demanding attention to privileged information. We explore in more detail the role that such attention-paying plays in references to privileged ground in later experiments, and tie this notion explicitly to our implementation of threat-to-referential success.

Experiment 1

Method

Participants. Participants were 24 native English-speaking undergraduates at the University of California, San Diego. Each participated as both a speaker and an addressee.

Materials and Design. Each pair of participants viewed 288 simple line drawings of common objects, four of which were displayed on any given 8 1/2 x 11

inch page in a binder. Objects varied in size across and within trials such that a large (or small) object on one trial (relative to the size of the other objects in that set) might be small (or large) on another trial. Each object type was used on only one trial and never occurred with more than one other object of the same type.

Two manipulations were used: contrast-type (contrasting versus non-contrasting) and target-type (privileged-ignorable target versus privileged-relevant-target). Four experimental conditions were assigned to each critical object by crossing the levels of the contrast-type and the target-type. Contrast-type was manipulated within speakers and items in counterbalanced fashion and target-type was manipulated between speakers and within items in counterbalanced fashion. On critical trials speakers were instructed to describe a particular common ground object (hereafter, *target*) for addressees (e.g., a heart). On *contrasting* trials the target was paired with a size-contrasting object of the same type in privileged ground (e.g., a larger heart). On *non-contrasting* trials the target was unique to the set (e.g., a privileged square, so that the target heart was the only heart in the set). Speakers in the privileged-ignorable target condition only named targets that were in common ground. Speakers in the privileged-relevant-target condition named targets that were in privileged ground (on 18 out of 36 filler trials). This resulted in both a threat to referential success and heightened attention to privileged ground objects. All other targets in this condition were in common ground (54 trials, for a total of 72 trials in all).

On all trials, speakers' use of size-contrasting modifiers was measured. When speakers use size-contrasting modifiers (e.g., when speakers call the target heart "small heart" when there is a larger heart in privileged ground), then they have used privileged information to ground a common-ground referent. Their use of modifiers on contrasting versus non-contrasting trials provides a measure of their references to privileged information.

Procedure

A coin toss randomly assigned participants to the roles of speaker and addressee. After completing a full set of stimuli in these roles, participants changed roles and completed another full set of stimuli (from a different condition). Participants sat at opposite sides of a table on which a binder was placed. At the beginning of each trial, addressees closed their eyes while speakers opened the binder to reveal a set of four objects. Speakers looked at those objects for two seconds, after which the experimenter pointed to one of the objects. Speakers then blocked that object by positioning an occluder (a 3-inch x 5-inch picture frame) between the object and addressees so that addressees could not see the object when viewing the display. After a two-second delay, the experimenter pointed to one of the four objects. Speakers were instructed to identify that object for addressees without using eye movement, head movement or direction (i.e., "the object on the left").

Upon hearing the description, addressees opened their eyes and pointed to the object that they thought speakers identified. Speakers then removed the occluder so addressees could see all objects, and indicated to addressees whether the choice was

correct. Addressees kept score of the number of correct choices made. Speakers and addressees were asked to maximize this score across the experiment. Questions and comments were allowed and encouraged during the practice session, but discouraged during the real trials.

Analyses

For each experimental condition, the percentage of targets described with direction-appropriate size-contrasting modifiers was computed for each subject. These percentages were submitted to repeated measures analyses of variance (ANOVAs) using subjects (E_1) and items (E_2) as random factors. (Analyses carried out using arcsine-transformed proportions yielded the same pattern of significance as reported here.) The ANOVA design was 2 x 2, with the factors of contrast-type and target-type. All significant effects achieved the .05 level unless otherwise specified. We report variability with repeated measures 95% confidence-interval half-widths (CIs) based on single-degree-of-freedom comparisons (Loftus & Masson, 1994).

Results

Figure 3 shows the mean percentages of target descriptions that included specified modifiers (i.e., “small triangle”) as a function of contrast-type and target-type. Speakers modified most often when the target had a size-contrasting match in privileged ground and speakers sometimes had to identify privileged objects, as compared to when speakers never had to identify privileged objects. In contrast, speakers never used modifiers when the target object was unique to the set regardless of whether they had to identify privileged objects.

Statistical analyses support these observations. The main effect of contrast-type was significant: Speakers produced significantly more modifiers in the contrasting condition (13.2%) than in the non-contrasting condition (0%), $F(1,22) = 7.9$, $CI = \pm 6.9\%$; $F(1,70) = 55.3$, $CI = \pm 3.1\%$. The main effect of target-type was also significant. Speakers produced significantly more modifiers in the privileged-relevant-target condition (12.3%) than in the privileged-ignorable target condition (0.9%), $F(1, 22) = 5.9$, $CI = \pm 6.9\%$; $F(1,70) = 39.1$, $CI = \pm 3.1\%$. The interaction between these two factors was also significant, $F(1, 22) = 5.9$, $CI = \pm 9.7\%$; $F(1,70) = 39.1$, $CI = \pm 4.4\%$. In the privileged-relevant-target condition, speakers modified target descriptions 24.5% more often on contrasting trials (24.5%) than on non-contrasting trials (0%), whereas in the privileged-ignorable target condition, speakers modified only 1.9% more on contrasting trials (1.9%) than on non-contrasting trials (0%).

Because participants acted as both speakers and addressees in this experiment, half of participants acted as speakers only after acting as addressees, whereas the other half acted as speakers naively, at the beginning of the experimental session. It is possible that when speakers previously experienced the task as addressees, they would better understand the communicative contingencies and so respect the privileged-common distinction more effectively. To investigate this, we reconducted the above analyses after introducing a *speaker order* factor, breaking speakers into groups depending on whether they acted as speakers first or second. These two populations acted similarly, producing more modifiers on contrasting trials than on

non-contrasting trials, but especially so in the privileged-relevant target condition. On contrasting trials (recall that no speaker ever produced a modifier on non-contrasting trials), participants who acted as speakers after acting as addressees produced modifiers 30.3% more in the privileged-relevant target condition (31.1%) than in the privileged-ignorable target condition (0.8%), whereas speakers who acted as speakers at the beginning of the experimental session produced modifiers 16.5% more in the privileged-ground condition (19.8%) than in the common-ground condition (3.3%). Thus, if anything, participants who acted as speakers only after acting as addressees were *more* likely to include privileged information in their utterances with greater communicative threat (and when more attention is drawn to privileged ground, in the privileged-relevant target condition). However, speaker order did not significantly affect modifier use, neither as main effect, nor as interaction with any other factor (all F s < 1).

Discussion

The results of Experiment 1 are that speakers made *more* references to privileged information when they had to cope with both a speaker-external pressure to avoid such references and a speaker-internal pressure that made privileged information salient. As such, a speaker-external pressure to avoid referring to privileged information did not act to decrease such references, at least not when paired with a speaker-internal pressure that worked to the opposite end. The observed pattern suggests that speakers are highly sensitive to speaker-internal pressures.

Thus, Experiment 1 suggests that when a manipulation heightens attention to privileged ground, speakers are only more likely to refer to that privileged ground. However, an alternative explanation for the pattern observed in Experiment 1 is that speakers in the privileged-relevant target condition may have used more modifiers because they were unable to develop an effective referential strategy due to the complexity of the task. Specifically, speakers might have been confused about how to name privileged objects when those objects had pair mates that were visible to addressees (i.e., it was difficult to name a privileged heart when there was another heart in common ground). This confusion may have carried over into their descriptions of common ground objects.

To illustrate, imagine that speakers are asked to identify a privileged ground heart when there is a smaller heart that is in common ground. Speakers can either say “heart,” or “large heart”. If speakers say, “heart,” addressees should assume that the target is the visible (common ground) heart, as they have no a priori reason to think that there is another heart in the set. An utterance of this form would result in a failure to make a correct reference because the target was actually the privileged heart. If speakers say, “large heart,” addressees will have to guess whether the target is the visible (common ground) heart (which in our experiments was always medium-sized) or the hidden object (this is the penalty or threat to referential success that was part of the design of this condition). (Of course, from a Gricean perspective [i.e., Grice, 1975], the correct strategy is to call the privileged larger heart ‘large heart,’ as including the extra information with respect to common ground should compel

addressees to seek additional information, which in this context would naturally be the privileged-ground heart.) Nonetheless, interlocutors are not always perfectly Gricean (see Engelhardt, Bailey, & Ferreira, 2006). To be maximally effective, speakers must consistently name common ground hearts, “heart,” and larger privileged ground hearts, “large heart,” an approach that may have been difficult for speakers to realize. Or relatedly, speakers may have determined that because they sometimes violated this general strategy (say, because cognitive pressures sometimes caused them to erroneously refer to the hidden heart), the entire strategy was undermined and so they chose not to pursue the strategy at all. In either case, the fact that speakers sometimes had to describe privileged-ground objects that had size-contrasting matches in common ground may have made it more difficult for speakers to determine or implement the optimal description strategy. Experiment 2 tests this possibility.

Experiment 2

In order to contrasting whether the results of Experiment 1 were due to participant confusion, Experiment 2 required all speakers to name privileged targets on some filler trials and manipulated (between subjects) whether those privileged target objects had pair mates in common ground or whether they were unique to the set. Privileged target referents that have pair mates in common ground (i.e., when speakers are asked to name a privileged heart when there is another heart in common ground) must be consistently grounded with size-contrasting information (just as in Experiment 1). Privileged target referents that are unique do not – if speakers are

asked to name a privileged circle when no other circle exists in the set, they need only say “circle.”

The effect of this manipulation is that it provides half of speakers with a simple strategy for referring to privileged objects and the other half with a more difficult strategy. If the results of Experiment 1 are due to the difficulty in determining a referential strategy, speakers who are required to name privileged ground targets that have size-contrasting pair mates in common ground should produce more modifiers than those that do not have the same requirement.

Experiment 2 used a slightly different procedure for presenting subjects with stimuli. Unlike Experiment 1, in which outlines of cards were printed on binder pages, Experiment 2 used actual individual cards. This difference in procedure should not affect the referential demands of the task; if it in fact does not, then we should see comparable performance between Experiment 1 and Experiment 2, at least in the analogous conditions (i.e., when speakers in Experiment 2 must identify privileged-ground cards that have shape matches in common ground).

Method

Participants. Participants were 56 undergraduates at the University of California, San Diego, 28 as speakers and 28 as addressees. Data were excluded from one pair due to experimenter error. All participants were native speakers of English.

Materials and Design. Each pair of participants viewed 288 different cards. Each card displayed one simple line drawing of a familiar object (the same objects

used in Experiment 1). Cards were placed on a table between speakers and addressees in groups of four.

Two manipulations were used: contrast-type (contrasting versus non-contrasting) and privileged–target-type (privileged-pair versus privileged-unique). Contrast-type was manipulated within subjects and items in counterbalanced fashion. Contrasting and non-contrasting trials were as described in Experiment 1. Privileged-target-type was manipulated between subjects and within items in counterbalanced fashion. Speakers in the privileged-pair condition were required to name targets in privileged ground that had size-contrasting pair mates in common ground on some filler trials. Speakers in the privileged-unique condition only named privileged ground targets that were unique to the set. For all speakers, 18 out of 36 filler trials included privileged ground targets. In all, 72 trials were administered.

Procedure

A coin toss randomly assigned participants to the roles of speaker and addressee. Participants sat at opposite sides of a table such that the speaker could see a computer monitor, but the addressee could not. At the beginning of each trial, the addressee closed his or her eyes while the experimenter placed four cards on the table. The speaker then looked at a computer monitor that displayed a schematic of four blank cards, one of which had an arrow above it with the instruction, *Block this card*. The speaker then blocked the corresponding actual card by positioning an occluder (a 5-inch x 5-inch picture frame) between the card and the addressee so that the addressee would not be able to see the card when viewing the display. Next, the

speaker looked back up at the computer monitor that displayed a new schematic of four blank cards. This time one of the cards had an arrow above it with the instruction, *Identify this card*. Speakers were instructed to identify the object on that card as in Experiment 1. Points were awarded as in Experiment 1.

Analyses

The analyses were carried out as in Experiment 1.

Results

Figure 4 shows the mean percentages of target descriptions that included specified modifiers as a function of contrast-type and privileged-target-type. Speakers were just as likely to use size-contrasting modifiers when a target had a size-contrasting match in privileged ground regardless of whether on filler trials speakers had to identify privileged targets that were unique to the set or privileged targets that had size-contrasting matches in common ground.

Statistical analyses support this observation. The main effect of contrast-type was significant. Speakers produced significantly more modifiers in the contrasting condition (25.9%) than in the non-contrasting condition (3.6%), $F(1,26) = 10.5$, $CI = \pm 10\%$; $F(1,70) = 60.5$, $CI = \pm 5.1\%$. The main effect of privileged-target-type was not significant. Speakers produced modifiers equally often in the privileged-pair condition (14.5%) as in the privileged-unique condition (15.0%), $F(1, 26) = <1$, $CI = \pm 10\%$; $F(1,70) = <1$, $CI = \pm 5.1\%$. The interaction between these two factors was also not significant, $F(1, 26) = <1$, $CI = \pm 14.1\%$; $F(1,70) = <1$, $CI = \pm 7.7\%$. In the privileged pair condition, speakers produced modifiers 21.8% more on contrasting

trials (25.4%) than on non-contrasting trials (3.6%), whereas in the privileged unique condition, they produced modifiers 22.7% more on contrasting trials (26.3%) than on non-contrasting trials (3.6%).

Discussion

The results of Experiment 2 show that modification rate was not influenced by whether a privileged ground target had a pair mate in common ground or was unique to the set. This suggests that speakers' use of modifiers in Experiment 1 was not influenced by any difficulty determining an effective referential strategy that is caused by having privileged-ground objects contrast in size with common-ground objects. Together, the results of Experiments 1 and 2 suggest that a manipulation specifically designed to raise a threat to communicative success (a speaker-external pressure) did not lead to fewer references to privileged ground when that threat co-existed with an internal pressure that caused speakers to allocate more attention to privileged ground.

Specifically, the requirement to name privileged objects, which was necessary to establish our speaker-external pressure of a threat to referential success (at least in the way it is currently implemented), presumably caused the presence of privileged objects to become highly salient to speakers. Highly salient privileged objects represent a speaker-internal cognitive pressure as they demand attentional resources. With respect to this task, the more highly salient privileged information became, the more difficult it may have been for speakers to avoid referring to it. Likewise, and as mentioned above, increased salience may be the reason that speakers' references to

privileged information increase with explicit instruction to avoid such references (Wardlow Lane et al, 2006). Note that this could have worked out the other way (both in Experiments 1 and 2 and in Wardlow Lane et al, 2006): Heightened attention to privileged information could have highlighted the privileged status of that information. The fact that heightened attention to privileged information only increases its likelihood of mention illustrates that speaker-internal pressures directly influence the processes that formulate utterances, rather than influencing speakers' utterances through representations that already encode common- versus privileged-ground distinctions.

Of course, having speakers sometimes name privileged-ground objects is a somewhat indirect manipulation of a speaker-internal pressure to increase the salience of privileged-ground objects. For instance, note that even when speakers sometimes had to name privileged-ground objects on filler trials, speakers did not have any special pressure to process privileged ground objects on target trials. Therefore, a demonstration that speakers' references to privileged ground are sensitive to a more direct manipulation of attention is warranted. Experiment 3 included such a more direct manipulation. The manipulation increased the degree to which speakers paid attention to privileged objects by drawing their attention to those privileged objects immediately before utterance production on critical trials. Specifically, on high salience contrasting trials speakers were required to look at the privileged object immediately before producing their reference to a common ground object. If increased salience of privileged information increases the likelihood that that

information will be mentioned, we should see an increased use of context-relevant modifiers in the high salience condition. This manipulation was arbitrary with respect to conversational pressure.

Note also that Experiments 1 and 2 exhibited very similar performance, despite the slight differences in procedure (Experiment 1 using a binder presentation with “cards” on a page, Experiment 2 using actual cards). The binder presentation is more efficient to implement, and so is used in subsequent experiments.

Experiment 3

Experiment 3 was designed to directly test whether a speaker-internal pressure, increased attention to privileged ground objects, increases the likelihood that speakers will identify those common ground objects with privileged ground information. To do this, Experiment 3 required speakers to identify common ground targets under low salience or high salience conditions. In the low salience condition, speakers were shown which object to name in a straightforward manner, as in Experiment 1: The experimenter indicated which object was the target by pointing to one of the three mutually visible objects. In the high salience condition, the identity of the target was revealed in a different way: Speakers opened a page in a binder that displayed instructions regarding which object to name. The instructions on the binder page were of the sort, “1 to the right of,” next to a reference object (i.e., a picture of one of the objects in the set). Using Figure 5 as an example, the instructions would say “1 to the right of” next to the picture of the large heart.

On all contrasting trials in the high salience condition, the reference object was the privileged object, drawing speakers' attention to privileged ground immediately before utterance production. Therefore, in order to determine which object was the target, speakers first had to look at the privileged object. Thus, speakers' attention to privileged ground objects was highest in the high salience condition. If attention to privileged ground objects, a speaker-internal cognitive pressure, causes increased mention of those objects when speakers name common ground targets, we should see an increased use of context-relevant modifiers in the high salience condition.

Method

Participants. Participants were 74 native English-speaking undergraduates from the University of California, San Diego who participated for experimental credit. Four participants were excluded due to experimenter error.

Materials and Design. The materials and presentation format were the same as those described in Experiment 1. The stimuli binder was laid flat on a table so that there was a top sheet and a bottom sheet of paper. On the top sheet of paper a set of four objects was displayed. In the high salience condition, the paper on the bottom half of the binder, when unfolded, indicated which object was the target object.

Two manipulations were used: contrast-type (contrasting versus non-contrasting) and attention-type (low salience versus high salience). Contrast-type and attention-type were manipulated within subjects and items in counterbalanced

fashion. Contrasting and non-contrasting trials were the same as in the previous experiments. On all trials the experimenter pointed to an object to indicate that it should be hidden (i.e., made privileged). On low salience trials, the experimenter waited two seconds and then pointed to another object to indicate that it was the target. On high salience trials the speaker unfolded the bottom sheet of the binder to reveal instructions. Instructions were typed across the bottom page and indicated a reference point, which was one of the objects on the top page. To the left of the indicated object were instructions such as “1 to the left of” or “2 to the right of,” revealing the identity of the target only in relation to the indicated object. On all experimental trials the indicated object was the privileged object. Thus, in order to determine which object to identify, speakers first needed to look at the privileged object, increasing the salience of that object.

Procedure

The procedure was the same as that described in Experiment 1, except for the implementation of the high salience condition as described.

Analyses

Analyses were conducted as in the previous experiments.

Results

Figure 6 shows the mean percentages of target descriptions that included specified modifiers as a function of contrast type and attention-type. Speakers used modifiers most often when the target had a size-contrasting match in privileged

ground and when the location of the target was revealed by its relationship to the privileged object as compared to when the location of the target was indicated by the experimenter. Speakers never used modifiers when the target was unique to the set regardless of how the location of the target was revealed.

Statistical analyses support these observations. The main effect of contrast type was significant. Speakers produced significantly more modifiers in the contrasting condition (6.8%) than in the non-contrasting condition (0%), $F1(1,69) = 25.0$, $CI = \pm 2.7\%$; $F2(1,35) = 27.6$, $CI = \pm 2.6\%$. The main effect of salience was also significant. Speakers produced significantly more modifiers in the high salience condition (4.9%) than in the low salience condition (2.1%), $F1(1, 69) = 5.4$, $CI = \pm 2.3\%$; $F2(1,35) = 10.8$, $CI = \pm 1.9\%$. The interaction between these two factors was also significant, $F1(1, 69) = 3.3$, $CI = \pm 3.3\%$; $F2(1,35) = 6.6$, $CI = \pm 2.6\%$. In the high-salience condition, speakers used modifiers 9.4% more on contrasting trials (9.4%) than on non-contrasting trials (0%), whereas in the low salience condition, speakers used modifiers only 4.3% more on contrasting trials (4.3%) than on non-contrasting trials (0%).

Discussion

Experiment 3 replicates the first two experiments in that speakers used more context-relevant modifiers when they named targets that had size-contrasting pair mates in privileged ground than when they named targets that were unique to the set. Importantly, this effect was larger in the high salience condition, when speakers' attention was drawn to the privileged object immediately before utterance production.

This shows that increasing the salience of privileged ground objects, which produces a speaker-internal cognitive pressure, results in more references to privileged information.

With respect to the results of Experiment 1, we cannot yet claim that the speaker-internal manipulation of attention to privileged information used in Experiment 3 involves the same cognitive mechanism as the speaker-internal manipulation of attention used in Experiment 1. If these mechanisms are different, then Experiment 3, though successfully demonstrating that attention to privileged information can increase references to that information, does not necessarily demonstrate that the manipulation used in Experiment 1 increased references to privileged information because it increased attention to privileged information. To test whether the manipulation of attention to privileged information used in Experiment 3 and the manipulation of attention to privileged information used in Experiment 1 are the same, Experiment 4 will introduce the manipulation used in Experiment 3 under circumstances where attention to privileged information was (hypothetically) already increased through the manipulation used in Experiment 1. If the cognitive bases of the two manipulations are independent, then the manipulation used in Experiment 3 should still be effective, even when attention to privileged information is already increased as was done in Experiment 1 (in the privileged-target condition); in fact, we would expect the same size effect of the manipulation in Experiment 4 as was observed in Experiment 3. However, if the cognitive bases of

the two manipulations are interrelated, then the salience manipulation should influence performance differently in Experiment 4 versus Experiment 3.

Experiment 4

Method

Participants. Participants were 72 undergraduate native English speakers from the University of California, San Diego. Each participated as both speaker and addressee. Five participants were excluded due to experimenter error.

Materials and Design. Materials were the same as those used in Experiment 3. The only difference in design between Experiments 3 and 4 was that speakers identified privileged ground objects (that were unique to the set) on six filler trials (8% of all trials) in Experiment 4 whereas they never identified privileged objects in Experiment 3.

Two manipulations were used: contrast-type (contrasting versus non-contrasting) and salience-type (low salience versus high salience). Contrast-type and attention-type were manipulated within subjects and items in counterbalanced fashion. Contrasting and non-contrasting trials were the same as in the previous experiments. Low and high salience trials were the same as in Experiment 3.

Procedure

The procedure was the same as in Experiment 3.

Analyses

Analyses were conducted as in the previous experiments.

Results

Overall Performance. Figure 7 shows the mean percentages of target descriptions that included specified modifiers as a function of contrast-type and salience-type. Speakers' use of modifiers when the target had a size-contrasting match in privileged ground was not influenced by whether the location of the target was revealed by its relationship to the privileged object or whether the location of the target was indicated by the experimenter. Speakers never used modifiers when the target was unique to the set regardless of how the location of the target was revealed.

Statistical analyses support these observations. The main effect of contrast-type was significant. Speakers produced significantly more modifiers in the contrasting condition (31.0%) than in the non-contrasting condition (0%), $F(1,66) = 44.5$; $CI = \pm 9.0\%$; $F(1, 35) = 305.4$, $CI = \pm 3.6\%$. The main effect of salience-type was not significant. Speakers produced about equal numbers of modifiers in the high salience condition (15.0%) than in the low salience condition (16.0%), $F(1,66) = < 1$; $CI = \pm 3.2\%$; $F(1, 35) = < 1$, $CI = \pm 3.4\%$. The interaction between these two factors was also not significant, $F(1,66) = < 1$; $CI = \pm 4.6\%$; $F(1, 35) = < 1$, $CI = \pm 4.9\%$. In the high-salience condition, speakers used modifiers 29.5% more often on contrasting trials (29.5%) than on non-contrasting trials (0%), whereas in the low salience condition, speakers used modifiers 31.5% more often on contrasting trials (31.5%) than on non-contrasting trials (0%).

To determine whether performance in Experiment 3 versus Experiment 4 was statistically different, we examined only the contrasting trials in a combined analysis

of the two experiments. The ANOVA design was 2×2 , with the factors salience (high vs. low) and experiment (3 vs. 4). If the 5.1% difference observed between salience conditions in Experiment 3 was larger than the 2.0% reversed difference observed in Experiment 4, this ANOVA should reveal a significant interaction. In fact, the interaction was significant, $F(1,135) = 6.5$, $CI = \pm 3.9\%$, $F(1, 70) = 8.1$, $CI = \pm 3.5\%$.

Discussion

Experiment 4 found that speakers were equally likely to make references to privileged ground objects when their attention was drawn to those privileged objects immediately before utterance production (high salience condition) as when it was not (low salience condition). This finding is in direct contrast to the results of Experiment 3 where rate of privileged ground references increased in the high salience condition. Furthermore, a combined analysis of Experiment 3 and Experiment 4 data revealed that the tendency to modify more in the high-salience condition was greater in Experiment 3 than in Experiment 4. In short, introducing a privileged ground target eliminated the salience manipulation. This suggests that the hypothetical manipulation of attention in Experiment 1 (requiring subjects to sometimes name privileged cards on filler trials) and the more face-valid manipulation of attention in Experiment 3 indeed relied on interrelated cognitive mechanisms. If the two manipulations were completely distinct, then each manipulation should have affected the use of size-contrasting modifiers independently. This suggests that indeed, subjects in the privileged-relevant target condition of Experiment 1 used size-

contrasting modifiers more often because the requirement to sometimes name privileged-ground targets specifically heightened attention to privileged-ground information overall.

General Discussion

Summarizing, Experiment 1 showed that the presence of a speaker-external pressure (a threat to referential success) to avoid inappropriately mentioning privileged information did not reduce the rate at which speakers made such references when it co-existed with a speaker-internal pressure that caused privileged information to be more salient. In fact, speakers used modifiers at a fairly high rate when both pressures co-existed. The results of Experiment 2 rule out the possibility that privileged ground references were made because speakers were confused about how to identify privileged objects that had size-contrasting matches in common ground. Together, Experiments 1 and 2 suggest that the discrepancy between the seemingly low rate at which speakers refer to privileged information in their everyday conversations and the seemingly higher rate at which they make such references in laboratory tasks is not because laboratory tasks fail to include communicative imperatives to exclude privileged information from utterances that are present in everyday conversations (at least in the situation where such imperatives co-exist with speaker-internal pressures to attend to privileged information). Because threats to referential success may commonly occur with heightened attention to privileged information in everyday conversation, the current manipulations do not necessarily represent an uncommon situation. Furthermore, heightened attention to obviously

privileged information could have had the effect of reinforcing that privileged status, resulting in fewer references to that information. The fact that the opposite was observed suggests that formulation processes are directly sensitive to speaker-internal pressures, rather than having speaker-internal pressures influence formulation through representations already tagged for common- versus privileged status.

The results of Experiment 3 show that a direct manipulation of a speaker-internal cognitive pressure (increased attention to privileged information) results in more inappropriate references to privileged information. This implies that the high rate of privileged ground references in Experiments 1 and 2 could indeed have occurred because the speaker-external communicative demand (a threat to referential success) served to increase a speaker-internal pressure by increasing the attention speakers allocated to privileged objects. Lastly, the results of Experiment 4 suggest that the mechanisms responsible for the high rate of privileged ground references in Experiments 1, 2 and 3 were likely the result of the operation of the same mechanism. Taken together, these experiments suggest that the more salient privileged information is, the more difficult it is for speakers to exclude it from their utterances, even when inclusion of privileged information threatens the fundamental goal of successful communication.

In order to give speakers the strongest chance of performing in a laboratory context as they do in everyday conversations, we incorporated a number of features into our experimental setting to increase naturalness while allowing for a tightly controlled experimental setting. Following the recent emphasis in psycholinguistics

on the importance of situating language users in naturalistic contexts where speakers produce their utterances for another naïve participant (Lockridge & Brennan, 2002), we included both a naïve speaker and naïve addressee in our task (i.e., we did not use a confederate participant). Additionally, we made our stimuli physically co-present because of the suggestion that speakers will be more likely to properly encode the distinction between information in common and privileged grounds if there is physical co-presence of stimuli (Clark & Marshall, 1981; Polichak & Gerrig, 1998a). Furthermore, we had speakers themselves make objects privileged, and the way privileged status was indicated was with an obvious physical barrier. As such, we attempted to give speakers every advantage to take the distinction between common and privileged knowledge into account so that they could exclude privileged information from their utterances. Nonetheless, we continued to see high rates of references to privileged information whenever speakers' attention was drawn to that privileged information.

When speakers make references to privileged information, they have failed in some sense to take into account the differences between what they know and what their partner knows (or, at the very least, they have failed to use that information properly when designing their utterances). Perhaps it is not surprising that speakers make this mistake so often, given the complexity of the task and the various speaker-internal and speaker-external pressures that may be present. The complexity of taking knowledge differences into account is revealed by the various theories that attempt to explain this process. To date there are two general processing theories that explain

how speakers are able to determine the common ground status of information and use that determination in real-time language production.

One theory, often called *monitoring and adjustment*, argues for a modularized and specialized two-stage processing strategy (Horton & Keysar, 1996; Keysar & Barr, 2002). The first stage is quick and automatic, wherein interlocutors take their own perspective on a given situation (i.e., speakers consider all information including both common and privileged ground). Thus, speakers initially design their utterances without taking into account that some information is likely privileged. In a task like the one used in the present experiments, speakers' initial utterance form would be something such as, "large triangle," even when the smaller triangle is privileged knowledge. Forming an "egocentric" utterance is presumably easy because speakers obviously have ready access to the information that they know themselves. The second stage, however, is an effortful and controlled monitoring stage in which interlocutors attempt to correct for their unique perspective, if given adequate time and resources. After correction for the fact that the smaller triangle is privileged knowledge, speakers' utterances should be of the form, "It's the triangle". Evidence for the dual-process model is that speakers produce more egocentric utterances – utterances that do not take perspective into account – when they are speeded (Horton & Keysar, 1996) and when they are under an increased cognitive load (Roßnagel, 2000).

A second theory takes a non-modular *probabilistic* approach to the problem of determining common ground. This theory claims that the distinction between

common ground and privileged ground is only one of many linguistic and non-linguistic constraints that speakers use when formulating utterances. In contrast to monitoring and adjustment, a probabilistic approach claims that whether information is in common or privileged ground has a partial but immediate effect on utterance form. The production system is said to integrate information from multiple sources using cue-weighting mechanisms that serve to provide probabilistic evidence for alternative interpretations (such as whether a given piece of information should be included in an utterance) that are activated in parallel. These constraints are weighted according to their salience and reliability and their effects are related to the probability that each cue will be important for communication (Hanna & Tanenhaus, 2004; Hanna, Tanenhaus, & Trueswell, 2003). These sources serve to update the common ground versus privileged ground distinction continuously as the interlocutor is engaged in conversation (Nadig & Sedivy, 2002).

Furthermore, recent work supporting this theory suggests that common ground processing falls out of more general cognitive functions (Horton & Gerrig, 2005a; Horton & Gerrig, 2005b). Horton and Gerrig (2005a) showed that speakers tailored their utterances to addressee needs only when necessary memory representations were readily accessible. This suggests that common ground processing (or audience design) can be explained by appealing to ordinary cognitive processing, specifically to the accessibility of memory representations, without the necessity of specialized and modular processing mechanisms.

The present results have implications for both of these general processing theories. First, the present results are consistent with the general spirit of monitoring-and-adjustment theories, in that monitoring-and-adjustment theories predict that utterance formation involves a measurable egocentric component. Because heightening speakers' attention to privileged information caused them to refer more to that privileged information (rather than making them better able to respect the privileged status of that information), the present results reveal a form of egocentricism. That said, it is important to note that the most distinctive aspect of monitoring-and-adjustment is that it predicts a time-course to speaker-based versus addressee-based influences on production; given enough time, speakers should be able to accommodate their addressees' distinct perspectives. In our experiments, we did not impose any time pressure on speakers, and thus, even according to monitoring-and-adjustment, speakers should have been able to consistently accommodate their addressees' distinct perspectives. This was in fact not observed, as speakers included references to privileged information up to 30% of the time (in conditions of heightened attention).

Also, the present results extend in informative ways the account of common ground processing proposed by Horton and Gerrig (2006), suggesting that common versus privileged-ground processing are critically underpinned by general cognitive mechanisms. Horton and Gerrig focused on the effects of memory retrieval, showing that when a cue to memory retrieval (namely, a particular conversational participant) was a highly effective cue for retrieving partner-specific information,

speakers efficiently took their addressees knowledge into account. The present results instead focus on a different general cognitive mechanism, namely attention allocation. Experiments 1 and 2 showed that when an environmental cue (namely, a visual occluder) is a highly effective cue to reducing attention (because speakers can always ignore information so cued – in the privileged-ignorable target conditions of Experiments 1), speakers successfully excluded privileged information. But when that same cue was not a highly effective cue to reducing attention (because speakers could not always ignore the cued information – in the privileged-relevant target conditions of Experiments 1, 2, and 4), speakers less successfully excluded privileged information. In short, just as the effectiveness of retrieval cues can affect speakers' ability to engage in audience design, so too does the effectiveness of attentional cues.

The results of these experiments might be taken to show that speakers are quite poor at taking the distinction between shared and privileged knowledge into account during utterance production. After all, speakers did inappropriately refer to privileged information quite often. This position might further suggest that the production system is suboptimal with regard to communicative efficacy. However, it should be noted that a system that forced speakers to tailor each utterance according to addressee needs might actually be suboptimal. Because addressees can signal their understanding to speakers and speakers can take an active role in establishing reference (Clark & Brennan, 1991), there is no need for speakers to shoulder the entire cognitive burden. Furthermore, given that it is likely that

comprehending language is at least somewhat less demanding than producing language, it should be more efficient to allow comprehenders to carry a larger burden for effective communication. As such, it is perhaps reassuring that external pressures to tailor utterances exert less force on utterance form than internal processing pressures.

Interestingly, it seems that the harder we try to get speakers to exclude privileged information, the more they *include* it (Wardlow Lane et al, 2006). However, what each of the manipulations we have used have in common is that they all increase the salience of privileged information; or more precisely, they increase the salience of the contrast between the privileged object and its size-contrasting pair mate in common ground. This finding is important for researchers investigating speakers' propensities to include privileged information in their utterances. Subtle aspects of experimental design that act to draw speakers' attention to privileged information may interfere with speakers' abilities to exclude that information causing an elevated rate of such references.

In conclusion we return to our initial question: To what extent do speaker-external pressures (i.e., communicative pressures) versus speaker-internal pressures (i.e., cognitive demands) affect utterance form? Based upon our results, we suggest that speakers are directly sensitive to speaker-internal pressures when designing their utterances.

Acknowledgements

This research was supported by National Institutes of Health grant R01 MH64733.

The authors express appreciation to Kristy Lawson for assistance with data collection and Bob Slevc, Tanya Kraljic and Tamar Gollan for helpful discussions.

Chapter 3, in full, has been submitted for publication of the material as it appears in the manuscript: Wardlow Lane, L. & Ferreira, V.S. Speaker-external versus speaker-internal forces on utterance form: Do cognitive demands override threats to referential success?

Figure Captions

Figure 1. Example experimental set up for Experiments 1 & 2

Figure 2. Example experimental set up for the Privileged-relevant target condition (where some filler targets are in privileged ground) and the Privileged-ignorable target condition (where all targets are in common ground).

Figure 3. Experiment 1. Percentages of target descriptions including foil-contrasting modifiers as a function of whether the foil and target were the same (contrasting trials) or different shapes (non-contrasting trials), and whether speakers were required to name some privileged ground targets on filler trials. Error bars illustrate 95% confidence interval of the interaction by speakers.

Figure 4. Experiment 2. Percentages of target descriptions including foil-contrasting modifiers as a function of whether the foil and target were the same (contrasting trials) or different shapes (non-contrasting trials), and whether speakers were asked to identify privileged objects that had pair mates in common ground (privileged-pair trials) or that were unique to the set (privileged-unique trials). Error bars illustrate 95% confidence interval of the interaction by speakers.

Figure 5. Example experimental set up for Experiments 3 and 4. Note that the binder was angled such that the instructions were only visible to the speaker.

Figure 6. Experiment 3. Percentages of target descriptions including foil-contrasting modifiers as a function of whether the foil and target were the same (contrasting trials) or different shapes (non-contrasting trials), and whether speakers were shown the target by a point (low salience trials) or through the use of a picture of the privileged object (high salience trials). Error bars illustrate 95% confidence interval of the interaction by speakers.

Figure 7. Experiment 4. Percentage of target descriptions including foil-contrasting modifiers as a function of whether the foil and target were the same (contrasting trials) or different shapes (non-contrasting trials), and whether speakers were shown the target by a point (low salience trials) or through the use of a picture of the privileged object (high salience trials). Error bars illustrate 95% confidence interval of the interaction by speakers.

Chapter 3, Figure 1
Example experimental set up for Experiments 1 & 2

Chapter 3, Figure 2

Example experimental set up for the Privileged-relevant target condition (where some filler targets are in privileged ground) and the Privileged-ignorable target condition (where all targets are in common ground).

Chapter 3, Figure 3

Experiment 1. Percentages of target descriptions including foil-contrasting modifiers as a function of whether the foil and target were the same (contrasting trials) or different shapes (non-contrasting trials), and whether speakers were required to name some privileged ground targets on filler trials. Error bars illustrate 95% confidence interval of the interaction by speakers.

Chapter 3, Figure 4

Experiment 2. Percentages of target descriptions including foil-contrasting modifiers as a function of whether the foil and target were the same (contrasting trials) or different shapes (non-contrasting trials), and whether speakers were asked to identify privileged objects that had pair mates in common ground (privileged-pair trials) or that were unique to the set (privileged-unique trials). Error bars illustrate 95% confidence interval of the interaction by speakers.

Chapter 3, Figure 5

Example experimental set up for Experiments 3 and 4. Note that the binder was angled such that the instructions were only visible to the speaker.

Chapter 3, Figure 6

Experiment 3. Percentages of target descriptions including foil-contrasting modifiers as a function of whether the foil and target were the same (contrasting trials) or different shapes (non-contrasting trials), and whether speakers were shown the target by a point (low salience trials) or through the use of a picture of the privileged object (high salience trials). Error bars illustrate 95% confidence interval of the interaction by speakers.

Chapter 3, Figure 7

Experiment 4. Percentage of target descriptions including foil-contrasting modifiers as a function of whether the foil and target were the same (contrasting trials) or different shapes (non-contrasting trials), and whether speakers were shown the target by a point (low salience trials) or through the use of a picture of the privileged object (high salience trials). Error bars illustrate 95% confidence interval of the interaction by speakers.

References

- Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(6), 1482-1493.
- Brown, P. M., & Dell, G. S. (1987). Adapting production to comprehension: The explicit mention of instruments. *Cognitive Psychology*, 19, 441-472.
- Clark, H. H. (1996). *Using language*. Cambridge, England: Cambridge University Press.
- Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick & J. M. Levine (Eds.), *Perspectives on socially shared cognition* (Vol. 13, pp. 127-149). Washington, D.C.: American Psychological Association.
- Clark, H. H., & Carlson, T. B. (1982). Context for comprehension. In J. Long & A. Baddeley (Eds.), *Attention and performance* (pp. 1-37.). New York, NY: Academic Press.
- Clark, H. H., & Marshall, C. (1981). Definite reference and mutual knowledge. In A. Joshi, B. Webber & I. Sag (Eds.), *Elements of discourse understanding* (pp. 10-63). Cambridge, England: Cambridge University Press.
- Clark, H. H., & Murphy, G. L. (1982). Audience design in meaning and reference. In J. F. Le Ny & W. Kintsch (Eds.), *Language and comprehension* (pp. 287-299). Amsterdam, Holland.
- DePaulo, B. M., & Coleman, L. M. (1986). Talking to children, foreigners, and retarded adults. *Journal of Personality and Social Psychology*, 51, 945-959.
- Engelhardt, P. E., Bailey, K. G. D., & Ferreira, F. (2006). Do speakers and listeners observe the Gricean Maxim of Quantity? *Journal of Memory & Language*, 54, 554-573.
- Fernald, A., & Mazzie, C. (1991). Prosody and focus in speech to infants and adults. 27, 209-221.
- Garvey, C. (1975). Requests and responses in children's speech. *Journal of Child Language*, 2, 41-63.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics: Speech act* (Vol. 3, pp. 41-58). New York: Academic Press.

- Hanna, J. E., & Tanenhaus, M. K. (2004). Pragmatic effects on reference resolution in a collaborative task: Evidence from eye movements. *Cognitive Science*, 28(1), 105-115.
- Hanna, J. E., Tanenhaus, M. K., & Trueswell, J. C. (2003). The effects of common ground and perspective on domains of referential interpretation. *Journal of Memory & Language*, 49(1), 43-61.
- Horton, W. S., & Gerrig, R. (2002). Speakers' experiences and audience design: Knowing *when* and knowing *how* to adjust utterances to addressees. *Journal of Memory & Language*, 47, 589-606.
- Horton, W. S., & Gerrig, R. (2005a). The impact of memory demands on audience design during language production. *Cognition*, 96, 127-142.
- Horton, W. S., & Gerrig, R. J. (2005b). Conversational common ground and memory processes in language production. *Discourse Processes*.
- Horton, W. S., & Keysar, B. (1996). When do speakers take into account common ground? *Cognition*, 59(1), 91-117.
- Keysar, B., & Barr, D. (2002). Self anchoring in conversation: Why language users don't do what they "should". In T. Gilovich, D. W. Griffin & D. Kahneman (Eds.), *Heuristics and biases: The psychology of judgment* (pp. 150-166). New York, NY: Cambridge University Press.
- Keysar, B., & Horton, W. S. (1998). Speaking with common ground: From principles to processes in pragmatics: A reply to polichak and gerrig. *Cognition*, 66, 191-198.
- Kingsbury, D. (1968). Manipulating the amount of information obtained from a person giving directions: Harvard University, Cambridge, MA.
- Lockridge, C. B., & Brennan, S. E. (2002). Addressees' needs influence speakers' early syntactic choices. *Psychonomic Bulletin & Review*, 9(3), 550-557.
- Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review*, 1, 476-490.
- Nadig, A. S., & Sedivy, J. C. (2002). Evidence of perspective-taking constraints in children's on-line reference resolution. *Psychological Science*, 13(4), 329-336.
- Polichak, J., & Gerrig, R. (1998a). Common ground and everyday language use: Comments on horton and keysar (1996). *Cognition*, 66, 183-189.

- Polichak, J., & Gerrig, R. (1998b). Common ground and everyday language use: Comments on horton & keysar. *Cognition*, 66(2), 182-189.
- Roßnagel, C. (2000). Cognitive load and perspective-taking: Applying the automatic-controlled distinction to verbal communication. *European Journal of Social Psychology*, 30, 429-445.
- Schober, M. F., & Brennan, S. E. (2003). Processes of interactive spoken discourse: The role of the partner. In A. C. Graesser & M. A. Gernsbacher (Eds.), *Handbook of discourse processes* (pp. 123-164). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.
- Stalnaker, R. C. (1978). Assertion. In P. Cole (Ed.), *Syntax and semantics 9: Pragmatics* (pp. 315-332.). New York, NY: Academic Press.
- Wardlow Lane, L., Groisman, M. & Ferreira, V.S. (2006). Don't talk about pink elephants! Speakers' control over leaking private information during language production. *Psychological Science*.

CHAPTER 4

BE WARY OF PAYING PEOPLE TO KEEP SECRETS: EXAMINING THE USE OF INCENTIVES TO CONCEAL CONFIDENTIAL INFORMATION

Abstract

A firm's value is tied to its ability to protect confidential information. We ask whether strategies used to minimize intentional leaks of confidential information, such as direct instruction and incentives, may actually increase unintentional leaks. We examined this question in two experiments using a competitive buyer-seller product exchange task. Sellers' goal was to sell products to buyers without revealing, or leaking, confidential information about the availability of a pseudo-competitor's product (which was better than the product being sold in all cases). Experiment 1 showed that sellers were more likely to leak confidential information when provided with instruction and non-monetary incentive to avoid doing so. Experiment 2 showed that monetary incentive further exacerbated rather than diminished the effects found in Experiment 1. Results suggest that firms should avoid stressing what information should be kept confidential – a very common practice – because highlighting confidential information may increase its likelihood of being unintentionally revealed.

Introduction

The importance of managing confidential information could not be made clearer than in a recent corporate scandal surrounding the Fortune 500 computer and office equipment maker Hewlett-Packard (HP). Preventing confidential information leaks was of such importance to the board's chairperson that she reportedly hired third parties to illegally monitor board members (who were opposed to her business strategies) and journalists suspected of exchanging HP's confidential information (Kanellos, 2007). The chairperson's actions led to her indictment by California's attorney general and ultimate resignation from HP on September 22, 2006. Days before her resignation, the chairperson defended her actions, emphasizing the value of preventing the release of confidential information: "The recent events that have taken place follow an important investigation that was required after the board sought to resolve the persistent disclosure of confidential information from within its ranks. These leaks had the potential to affect not only the stock price of HP but also that of other publicly traded companies" (HP Press Release, September 12, 2006).

Motivated by recent scandals such as the one described above, the current paper examines methods of preventing or controlling leaks of confidential information. Very broadly, leaks of confidential information can be categorized as *intentional* or *unintentional* (e.g., Demeski, Lewis, Yao & Yildirim, 1999; Flynn, 2006). Intentional information leaks are by definition "wanted". For example, as in the HP scandal, an intentional leak may come from an anonymous board member of a corporation interested in disseminating confidential information to the perceived

benefit of the firm. To this end, it cannot be clearly argued when and how such leaks should be prevented or controlled. Still, firms frequently implement safeguards to prevent intentional leaks such as those experienced at HP (e.g., at the board level). For example, firms may label information as confidential or classified to highlight what should not be revealed. Furthermore, for highly sensitive information, some firms pay premiums to minimize the risk that confidentiality will be breached (e.g., to an accounting or consulting firm).

Conversely, unintentional leaks are by definition “unwanted”. For example, if a salesperson unintentionally leaks information about the relative efficacy of a competitor’s product (if superior) to a potential buyer, the leaked information may prevent the sale, reducing earnings for both the salesperson and the firm. In such a context, it can be safely argued that it is in the firm’s best interest to prevent, or at the very least control, unintentional information leakage. Here we focus on unwanted, unintentional leaks of confidential information.

Leaks of confidential information can be further categorized as *explicit* (directly expressed) or *implicit* (indirectly expressed; e.g., McKenzie, Liersch, Finkelstein, 2006; Sher & McKenzie, 2006; McKenzie, 2004; Wardlow Lane, Groisman, & Ferreira, 2006). If, for example, a salesperson says something to a customer such as, “this product is not as good as other products on the market,” that salesperson would have *explicitly* leaked information that the product they were selling was relatively bad. If a salesperson instead says something like, “this product is OK,” that salesperson would have *implicitly* leaked information about the product

in question. While explicit leaks deserve attention, it would be somewhat unsurprising if explicit leaks had an impact on those receiving the information (e.g., the customer in the example above). Conversely, it is unclear whether those receiving implicit information leaks can absorb and utilize that information. It is also unclear how to train employees to avoid making implicit leaks of confidential information. For this reason, we examine implicit information leaks in the current paper. In sum, unlike HP's information leaks, which were intentionally and explicitly communicated to journalists by HP board members, we focus here on leaks of confidential information that are *unintentional* and *implicit* in nature.

While methods such as labeling and highlighting what information *should not* be revealed are often used to *decrease intentional* leaks of confidential information, recent psychological research suggests that these same methods may inadvertently *increase unintentional* leaks of confidential information. Specifically, Wardlow Lane, Groisman and Ferreira (2006) showed that people had an overall tendency to make unintentional and implicit references to confidential information. Most importantly people were more likely to leak confidential information when given instruction and non-monetary incentive to avoid doing so. Presumably these information leaks were unintentional because incentives to reduce leaks should, from a rational perspective, never cause people to intentionally increase levels of information leakage (e.g., Camerer & Hogarth, 1999). In other words, highlighting information that should have been kept confidential, even in the form of instructions to not mention that information, led to more (rather than less) unintentional leaks. We call this the Pink

Elephant effect because it is akin to the well-known psychological finding that instructing people *not to think* about something, such as a pink elephant, causes them *to think* about it more (Wegner, 1994b).

Wardlow Lane et al. (2006) hypothesized that the mere act of bringing attention to confidential information made it more readily accessible, and therefore more likely to be (unintentionally) revealed (see Wegner, 1994a, 1998, for a discussion of ironic processes). Furthermore, recent work has confirmed that making confidential information more salient does increase the chances that that information will be mentioned even when doing so results in an unwanted outcome (Wardlow Lane & Ferreira, 2007). This result suggests that the more a firm attempts to prevent intentional leakage of confidential information (by, for instance, highlighting what information should not be revealed), the more confidential information will be revealed unintentionally – an extremely counterintuitive proposition.

However, making such inferences from Wardlow Lane et al. (2006) to the business world may be premature because both tasks could be construed as abstract in nature. Indeed, past reports in cognitive psychology have shown that task performance can be influenced by whether people can bring real-world knowledge to bear in solving the task. Importantly, abstract tasks, which tend to make it more difficult to apply such real-world knowledge, often result in poorer performance than tasks that allow participants to apply such knowledge (see Griggs & Cox, 1982 and McKenzie, 2005). Thus, we reframed the original Wardlow Lane et al. task so that it

represented a real-world buyer-seller transaction to examine whether the original findings would generalize to real-world business, rather than just laboratory, contexts.

In Wardlow Lane et al.'s original task, two people, labeled a speaker and a listener, sat across from one another at a table. Four cards were placed on the table each with simple line drawings of common objects (such as stars, hearts and circles). Three of the objects were visible to both the speaker and the listener, making up their shared knowledge. A fourth object was visible only to the speaker, making its identity confidential to the speaker. For example, imagine that the speaker and listener could both see a small triangle, a large triangle, and a circle (see Figure 1). Another, smaller, circle was only visible to the speaker. The objects varied in size such that a large instance of one object in a pair could be equivalent to the small instance of one object of a different pair.

Figure 1 About Here

The goal of the task was for the speaker to name a target object, which was always medium-sized and one of the three mutually visible objects, for the listener. The speaker was instructed to use as much information as necessary, but not more information than was necessary for the listener to be able to pick it out of the set of three mutually visible objects. If the target object was the sole mutually visible circle, we might expect the speaker to say, "it's the circle". However, speakers sometimes said, "it's the large circle," implicitly revealing information about the identity of a confidential small circle.

On half of trials, in what was labeled the “conceal” condition, speakers were instructed to conceal the identity of the confidential object (i.e., they were instructed not to reveal any confidential information to the listener). Furthermore, in the conceal condition, a point system was implemented that penalized speakers when listeners were able to correctly determine the identity of the confidential object. Recall that the results showed that speakers were *more* likely to leak confidential information in the conceal condition as compared to a baseline condition that included no instructions about the confidential object.

In the current studies, we attempted to allow participants to bring their real-world knowledge to bear by changing Wardlow Lane et al’s (2006) abstract speaker-listener object identification task to a less abstract buyer-seller product exchange task (see Table 1).

Table 1 About Here

Specifically, two participants were assigned to act as a seller or a buyer. Participants were shown pictures of products, some of which are typically sold by salespeople working on a commission basis, such as electronics, automobiles and jewelry. Three of these products were visible to both the seller and buyer. A fourth, confidential, product was visible to the seller only. Participants were told that the confidential product was a product that a pseudo-competitor was selling. Furthermore, they were told that on all trials, the pseudo-competitor’s product could be assumed to be a “better” product. If buyers could correctly identify the confidential, pseudo-competitor’s, product, they could purchase that product rather

than the sellers' product (i.e., the seller would lose the sale). As such, the pseudo-competitor's product was confidential information that the seller did not want to leak to the buyer. The task was for sellers to identify the objects that they were selling so that buyers could pick them out of the set of mutually visible objects. For example, if sellers were asked to sell a ring, we might expect them to say that they were selling, "a ring." However, when sellers identified their products by comparing them to the pseudo-competitors' products (i.e., when they said "a small ring" because of the presence of a pseudo-competitor's confidential, larger ring), they implicitly leaked confidential information to buyers.

Experiment 1 asks whether instruction and non-monetary incentive to not leak confidential information will influence the degree to which sellers can keep that information confidential in a more real-world task than Wardlow Lane et al's (2006). Drawn from the Wardlow Lane et al result, the *Pink Elephant* hypothesis says that sellers will leak confidential information more often when provided with instruction and incentive not to leak that information. In contrast, the *Thematic-Materials* hypothesis says that sellers will leak confidential information less often when provided with instruction and incentive not to leak that information. This hypothesis is named after the Thematic-Materials Effect which resulted from research which showed that participants perform better on tasks requiring inferential logic, such as Wason's Card Sorting task, when the task is framed in a manner that allows participants to bring their real-world experiences to bear (Griggs & Cox, 1982; Johnson-Laird, Legrenzi & Sonino Legrenzi, 1972; Wason & Shapiro, 1971).

Experiment 1 tests these hypotheses. The Pink Elephants hypothesis will be supported if we find that sellers leak confidential information more often when provided with an instruction and incentive to avoid doing so. The Thematic-Materials hypothesis will be supported if we find that sellers leak confidential information less often when provided with an instruction and incentive to do so.

Furthermore, we examine buyers' performance to determine whether buyers are sensitive to implicitly leaked confidential information (see McKenzie, 2005), an aspect of behavior that Wardlow Lane et al did not report. After all, if buyers are unaware of, or unable to use, implicitly leaked information, then such information is essentially benign. That is, the conditions under which people are likely to unintentionally and implicitly leak confidential information is only instructive if other people are able to use that leaked information.

Experiment 1

Method

Participants. Sixty-four undergraduates from the University of California, San Diego, each participated as a seller and as a buyer.

Materials and design. Each pair of participants saw 144 simple line drawings of familiar objects, hereafter *products*, four of which were displayed in a row on an 8 1/2 x 11" page in a binder. The products varied in size across and within trials such that the actual size of a large (or small) product (relative to the size of the other objects in that set) on one trial might be small (or large) on another trial. The products were items that are typically bought and sold at retail establishments, often resulting

in a commission for the salesperson. Examples of such items are jewelry, real estate, automobiles and clothing (see Figure 2 for an example of a typical trial).

Figure 2 About Here

Two manipulations were used: contrast-type (contrasting versus non-contrasting) and pay-type (commission versus salary). On contrasting trials, mutually visible targets were paired with confidential *foils* that were the same type of object as the target, though they contrasted in size. On non-contrasting trials, mutually visible targets were objects that were unique to the set. On commission trials, sellers were told not to provide buyers with any information about the hidden (confidential) object. On salary trials, sellers were not given any special instructions regarding the hidden object. Sellers were asked to identify a mutually visible object on all trials.

Four experimental conditions were assigned to each critical product by crossing the levels of contrast-type and pay-type. Both factors were manipulated within participants and items in counterbalanced fashion.

Procedure

Each participant played both the role of the buyer and seller during the course of the experimental session. A coin toss randomly assigned participants to the roles of seller and buyer for the initial round. Participants sat at opposite sides of a table on which a binder was placed. At the beginning of each trial, buyers were instructed to close their eyes while sellers opened the binder to reveal a set of four products. Sellers were instructed to look at the products for two seconds, after which the experimenter pointed to one of them. Sellers then blocked that product by positioning

an occluder (a 3-inch x 5-inch picture frame) between the product and the buyer so that buyers could not see the product when viewing the display (i.e., they made the product confidential). The experimenter waited two seconds before pointing to another product. Sellers were instructed to identify that product for buyers without using eye movement, head movement or direction (i.e., “the object on the left”). Buyers opened their eyes upon hearing the description, and pointed to the product that they thought sellers identified. Sellers then removed the occluder so buyers could see all of the products.

In the *salary condition*, when the buyer purchased the seller’s product (i.e., when buyers could correctly identify the product that the seller named), each participant received .5 point (see Table 2 for a full description of the score structure). When a purchase was not made (i.e., when buyers could not correctly identify the product that the seller named), neither participant received a point.

Table 2 About Here

In the *commission condition*, buyers and sellers participated in a two-step process. In the first step, sellers named the target product and buyers attempted to choose that product. In the second step, buyers were instructed that they could guess the identity of the hidden product (i.e., the competitor’s product) if they had correctly identified the sellers’ product in the first step. If buyers chose not to guess what the competitor was selling, but a transaction had been made in the first step, then the seller received one point and the buyer received half a point. Sellers received more points, because they received an extra .5 point as commission. However, if buyers

correctly guessed the identity of the hidden product, sellers lost that sale and commission, and because of that they received only their salary (.5 point) for that trial. In this circumstance, buyers were rewarded with one point for buying the competitor's hidden product. Because buyers' only means for determining the identity of the hidden product was when sellers made implicit reference to it (i.e., when they said, "small diamond ring"), sellers should avoid leaking any information about the hidden product to buyers in this condition.

If buyers guessed incorrectly, neither participant received any points. This point structure was implemented to discourage sellers from intentionally misleading buyers. The sellers' task was simply to name the product without providing any information that could help buyers determine the identity of the hidden product. Additionally, buyers had no incentive to guess randomly as the result of an incorrect guess was zero points for both participants on that trial.

Finally, it is important to note that, although sellers and buyers each earned separate scores (see Table 2 for a full description of the score structure), each participants' goal was to maximize their *combined* buyer and seller score since they were playing the role of both buyer *and* seller during the course of the experiment.

Analysis

Utterances were coded according to whether or not a context-relevant modifier (i.e., "small" or "large") was used. Modifier use was assessed on contrasting and non-contrasting trials, and the degree to which modifiers were used more on contrasting than on non-contrasting trials measures references to confidential

information. The raw percentage of utterances that included a context-relevant modifier was calculated in each condition for each participant.

Statistics were conducted with repeated-measures analyses of variance (ANOVAs) using subjects (F) as random factors. The ANOVA design was 2 x 2, with the factors contrast-type and pay-type. Planned comparisons assessed performance on contrasting versus non-contrasting trials separately under each pay-type instruction. All effects reported as significant reached the .05 level or better. Variability is reported with repeated-measures 95% confidence-interval halfwidths (CIs) based on single degree-of-freedom pairwise comparisons (Loftus & Masson, 1994).

Results

Leakage of confidential information. Figure 3 shows the mean percentages of target product descriptions that included context-relevant modifiers (i.e., “small diamond ring”) as a function of contrast-type and pay-type. Sellers modified most often when (1) the target had a size-contrasting confidential match and (2) they were instructed and provided point-based incentive not to leak confidential information -- as compared to when sellers had a size-contrasting confidential match and no instruction or point-based incentive not to leak confidential information. In contrast, sellers rarely used modifiers when the target product was unique to the set regardless of whether they were given instruction and incentive to avoid leaking confidential information.

Statistical analyses support these observations. The main effect of contrast-type was significant: Sellers produced significantly more modifiers in the contrasting condition (12.2%) than in the non-contrasting condition (0.8%) $F(1,63) = 13.0$, $CI = \pm 8.4\%$. The main effect of pay-type was also significant. Sellers produced significantly more modifiers in the commission condition (4.8%) than in the salary condition (0.8%) $F(1,63) = 5.6$, $CI = \pm 3.8\%$. The interaction between these two factors is marginally significant, $F(1, 63) = 1.9$, $p = 0.17$, $CI = \pm 4.7\%$. In the commission condition, sellers modified target descriptions 14.5% more often on contrasting trials (14.7%) than on non-contrasting trials (0.2%), whereas in the salary condition, sellers modified only 9.6% more on contrasting trials (9.6%) than on non-contrasting trials (0%). The planned comparison between contrasting trials in the commission and salary condition was significant, $F(1, 63) = 8.4$.

Figure 3 about here

Sellers' performance. Table 3 shows the number of trials on which sellers and buyers earned 0.0, 0.5 or 1.0 points as a function of pay-type and modifier use. Overall, sellers were 2.7 times more likely to *lose* points when they leaked confidential information (as measured by their use of modifiers on contrasting trials) in the commission condition: When sellers leaked confidential information in the commission condition, they received 0.5 point on 69.1% of trials, whereas they received 1.0 full point on 25.5% of trials. In contrast, sellers were 1.9 times more likely to *earn* 1.0 full point when they did not leak confidential information in the commission condition (46.6% of trials as compared to 0.5 points on 25% of trials).

Furthermore, revealing confidential information did not carry a penalty in the salary condition. Sellers received the maximum possible points per trial (0.5) on 99% of trials when no confidential information was leaked and on 100% of trials when confidential information was leaked.

Table 3 about here

Table 4 shows the average number of points earned by sellers and buyers on contrasting trials according to modifier use. Calculating the average number of points received by sellers and buyers on contrasting trials in the commission condition shows that sellers outscored buyers by 0.11 points when sellers did not leak confidential information. Importantly, buyers outscored sellers by 0.24 points on trials where sellers did leak confidential information.

Table 4 about here

Buyers' performance. The pattern for buyers' scores was the opposite of the pattern for sellers' scores. That is, buyers were more 2.8 times more likely to *gain* 0.5 point when sellers leaked confidential information in the commission condition. When sellers leaked confidential information in the commission condition buyers received 1.0 full point on 71% of trials, whereas they received only 0.5 points on 25.5% of trials. In contrast, buyers' scores suffered when sellers did not leak confidential information in the commission condition. Specifically, buyers received 1.0 full point on 24.1% of trials, 0.5 point on 47.5% of trials and they earned 0 points on 28.7% of trials.

Additionally, buyers' guessing behavior was influenced by sellers' use of modifiers. Buyers guessed the identity of the confidential product on 75% of trials when sellers leaked confidential information in the commission condition. Those guesses were correct on 95% of trials. In contrast, when sellers did not leak confidential information in the commission condition, buyers guessed the identity of the confidential product 52% of the time. Importantly, those guesses were correct less often as compared to when confidential information was leaked. Buyers made correct guesses on 54% of trials and incorrect guesses on 46% of trials when confidential information was not leaked.

Discussion

The results of Experiment 1 show that sellers are *more* likely to make an implicit reference to confidential information when given instruction and non-monetary incentive *not* to do so. Importantly, sellers revealed confidential information even though doing so was harmful to their performance and beneficial to buyers' performance. Furthermore, the results show that buyers are able to make use of implicitly conveyed confidential information. Specifically, buyers' scores improved when sellers revealed confidential information because buyers were able to make correct guesses about the identity of the confidential object more often when sellers implicitly leaked information about those products.

Surprisingly, buyers guessed the identity of the confidential product often, even when sellers did not reveal confidential information. Because this research was designed to investigate the factors that cause speakers to reveal or conceal

confidential information, we did not measure all potential influences on buyers' guessing behaviors. For example, sellers may have revealed information through differences in the timing and/or stress patterns of their speech. Additionally, there may be individual differences in buyers' risk-taking behaviors that are independent of this task. We leave these potentially informative investigations for future research. What is important for the current analysis is that buyers were more likely to guess, and more likely to guess correctly, when sellers revealed confidential information.

These results replicate the Wardlow Lane et al (2006) Pink Elephant effect using a task that allows participants to use real-world knowledge to perform the task. In the current study (and in Wardlow Lane et al, 2006), we extend the finding that people are more likely to *think* about an object when told not to by showing that instructions and non-monetary *not to talk* about something causes them *to talk* about it more. Ultimately, the Pink Elephant effect – when applied to the current study's business context – suggests that employees who are given specific instructions and non-monetary incentives not to reveal confidential information may actually be more likely to leak that confidential information than employees who are given no special instructions or incentives.

However, this result does not address whether employees' monetary incentives counteract this Pink Elephant effect. As noted previously, it is common for employees to receive financial incentives for good performance. In Experiment 2 we ask whether people are better able to keep confidential information private when provided with a monetary incentive to do so. If the monetary incentive only serves to

heighten the Pink Elephant effect (i.e., if more effort to conceal confidential information results in the operation of an ironic process more often) then we might expect a monetary incentive to *increase* the amount of confidential information that people leak. It is possible that this could occur despite the fact that the monetary incentive is specifically provided in order to *decrease* the amount of confidential information that people leak. In contrast, if financial incentives to keep confidential information private do prove effective, we would expect people to reveal less confidential information when given a monetary incentive to do so. Experiment 2 tests these hypotheses.

Experiment 2

Method

Participants. Participants were 48 pairs of native English-speaking undergraduates from the University of California, San Diego who participated in pairs for experimental credit. Participants acted either as sellers or as buyers. Two pairs of participants were excluded for not following directions.

Materials and Design. The materials were the same as those used in Experiment 1. But in contrast to Experiment 1, the design was modified in the current experiment such that participants only played one role (i.e., they did not switch roles half-way through the experiment as they did in Experiment 1). In this way, the individual, rather than combined, effects of information leakage could be examined on sellers and buyers. Also, sellers performed the task two times, using different stimuli groups.

Three manipulations were used: contrast-type (contrasting versus non-contrasting), pay-type (commission versus salary) and incentive-type (course credit or lottery payment). The contrast-type and pay-type conditions were implemented as in Experiment 1. Incentive-type was manipulated within subject. On course credit trials, participants performed the entire experiment one time for course credit. On lottery payment trials, participants performed the entire experiment one time for the opportunity to win \$100. Order of incentive type was counterbalanced across subjects. Participants were informed that the seller who earned the highest score and the buyer who earned the highest score across the entire experiment would receive a \$100 payment.

Eight experimental conditions were assigned to each critical product by crossing the levels of contrast-type, pay-type and incentive-type. All factors were manipulated within participants and items in counterbalanced fashion.

Procedure. The procedure was the same as that described in Experiment 1, except for the implementation of the incentive-type condition. To implement the incentive-type condition, participants were told that they would perform the task two times. They performed one round (a full 36 trials) for course credit. Scores on this round did not count toward a pay-off. Participants performed the other round (a full 36 trials) for a potential lottery payment (the \$100 payment described above). Order of incentive-type factor was counterbalanced and participants were informed of incentive-type condition before the start of each round.

Analyses. Analyses were conducted as in the previous experiments.

Results

Leakage of confidential information. Figure 4 shows the mean percentages of target descriptions that included specified modifiers (i.e., “small diamond ring”) as a function of contrast-type and target-type. Sellers modified most often when the target had a size-contrasting match in privileged ground and sellers were instructed and provided a monetary incentive not to leak confidential information, as compared to when there was a size-contrast and sellers were given a non-monetary incentive. In contrast, sellers rarely used modifiers when the target product was unique to the set regardless of whether they were given instruction and either incentive to avoid leaking confidential information.

Statistical analyses support these observations. The main effect of contrast-type was significant. Sellers produced significantly more modifiers in the contrasting condition (14.2%) than in the non-contrasting condition (0.3%) $F(1, 45) = 22.2$, $CI = \pm 5.9\%$. The main effect of pay-type was significant. Sellers produced significantly more modifiers in the commission condition (10.4%) than in the salary condition (4.1%) $F(1, 45) = 16.9$, $CI = \pm 3.1\%$. The main effect of incentive-type was not significant. Sellers produced modifiers on 8.1% of trials in lottery payment condition and they produced modifiers on 6.4% of trials in the course credit condition, $F(1, 45) = 1.0$, $CI = \pm 3.3\%$. The interaction between these three factors is marginally significant, $F(1, 45) = 2.3$, $CI = \pm 4.6\%$. In the contrasting condition, speakers modified target descriptions 6.4% more often on lottery payment trials (23.7%) than on course credit trials (17.3%) in the commission condition, whereas in the salary

condition, speakers modified at an equivalent rate on lottery payment trials (8.0%) and course credit trials (8.0%). The planned comparison between contrasting trials in the commission and salary condition was significant $F(1,45) = 4.4$.

Figure 4 about here

Sellers' performance. Table 5 shows the number of trials on which sellers and buyers earned 0.0, 0.5 or 1.0 points as a function of pay-type, incentive-type and modifier use. Overall, sellers were 4 times more likely to *lose* points when they revealed confidential information (as measured by their use of modifiers on contrasting trials) in the commission condition. When sellers revealed confidential information in the commission condition they received only 0.5 point on 77.7% of trials, whereas they received 1.0 full point on 19.6% of trials. In contrast, sellers were 2.6 times more likely to *earn* 1.0 full point when they did not reveal confidential information in the commission condition (55.7% of trials as compared to 0.5 points on 18.9% of trials).

Revealing confidential information did not carry a penalty in the salary condition. Sellers received the maximum possible points per trial (0.5) on 99% of trials when no confidential information was revealed and on 100% of trials when confidential information was revealed.

Table 5 about here

Table 6 shows the average number of points earned by sellers and buyers on contrasting trials according to incentive-type and modifier use. Calculating the average number of points received by sellers and buyers on contrasting trials in the

commission condition collapsed across incentive-types, shows that sellers outscored buyers by 0.18 points when sellers did not leak confidential information. Importantly, buyers outscored sellers by 0.30 points on trials when sellers did leak confidential information.

Table 6 about here

Buyers' performance. The pattern for buyers' scores was the opposite of the pattern for sellers' scores. Because the patterns between scores in both incentive-type conditions (i.e., lottery payment and course credit) were the same, the reported scores are collapsed across incentive-type. When sellers revealed confidential information in the commission condition, buyers were 4.0 times more likely to *gain* 0.5 point (i.e., earn a full point). When sellers leaked confidential information in the commission condition buyers received 1.0 full point on 77.7% of trials, whereas they received only 0.5 points on 19.6% of trials and they earned 0 points on 2.7% of trials. In contrast, buyers' scores suffered when sellers did not leak confidential information in the commission condition. Buyers were 2.9 times more likely to receive only 0.5 point when sellers did not reveal confidential information. Buyers received 1.0 full point on 19.3% of trials, 0.5 point on 56.7% of trials and they earned 0 points on 24.0% of trials.

Additionally, buyers' guessing behavior was influenced by sellers' use of modifiers. Buyers guessed the identity of the confidential product on 80% of trials when sellers leaked confidential information in the commission condition. Those guesses were correct on 97% of trials. In contrast, when sellers did not leak

confidential information in the commission condition, buyers guessed the identity of the confidential product 43% of the time. Importantly, those guesses were correct less often as compared to when confidential information was leaked. Buyers made correct guesses on 44% of trials and incorrect guesses on 56% of trials when confidential information was not leaked.

Discussion

The results of Experiment 2 show that sellers are not only more likely to make an implicit reference to confidential information when given instruction not to do so, but that they will reveal confidential information even more often when the instruction is accompanied by a monetary incentive not to reveal the confidential information. Furthermore, sellers continue to reveal confidential information even though doing so is harmful to their performance and beneficial to buyers' performance.

These results again extend the Wardlow Lane et al (2006) Pink Elephant effect using a task that is situated in a real-world context and that provided participants with a salient monetary incentive to improve performance. This finding suggests that the harder firms try to influence their employees to keep confidential information private, the more likely they are to influence their employees to reveal that confidential information.

General Discussion

Experiment 1 showed that people are more likely to implicitly and unintentionally leak confidential information when provided with an instruction and

non-monetary incentive to avoid doing so. This finding replicated earlier work (Wardlow Lane et al, 2006) using a task that is situated in a more common (and less abstract) buyer-seller context. Experiment 2 showed that people are also more likely to implicitly and unintentionally leak confidential information when provided with instruction and a salient monetary incentive to avoid doing so. Furthermore, the instruction not to leak confidential information and the monetary incentive not to leak confidential information had additive influences on sellers' behaviors. As such, we can conclude that these strategies often employed by firms to minimize leaks of confidential information (i.e., instructing employees not to reveal pieces of information and rewarding their behavior with increases in pay) may actually result in more implicit and unintentional leaks of confidential information.

It should be noted that the instruction we gave to sellers to avoid referring to confidential information was of the form, "do not provide any information about the hidden product." This instruction made no direct reference to sellers' use of modifiers. It is possible that sellers were unaware that their use of modifiers resulted in a leak of the information that they wanted to keep private. However, the analyses of guessing behavior and points reveals that sellers nonetheless lost points when they used modifiers that leaked confidential information. Furthermore, this loss of points did not lessen the extent to which they used modifiers. This is in itself interesting. Outside of the laboratory, in a business context, firms and their employees will potentially pay a price when employees implicitly leak confidential information even if employees are unaware of how the leak happened. Moreover, with respect to the

specific instruction that we used, current research suggests that providing speakers with a direct instruction not to use an adjective to compare a mutually visible object to a confidential object does not lessen the degree to which speakers include an adjective that leaks confidential information (Wardlow Lane & Ferreira, unpublished data).

Furthermore, that we found no influence of the reframing of the original task might suggest that the manipulation was not strong enough to reveal the Thematic-Materials effect (i.e., allow participants to bring their real-world knowledge to bear on the task). Although we agree that future research that implements a stronger practiced context would be valuable, the extent to which we reframed Wardlow Lane et al.'s task from an abstract speaker-listener object identification task to a buyer-seller product exchange game was equivalent to prior demonstrations of the Thematic Materials effect.

For example, the original demonstration of the Thematic Materials effect reframed the famous Wason Card Sorting task (Wason, 1966, 1968). Wason provided participants with four two-sided cards. On each card was printed a letter on one side and a number on the other. Participants were also presented with a rule, such as, "If there is a vowel on one side, there is an odd number on the other". Participants were instructed to turn over as few cards as possible in order to determine whether the rule was broken by the set of cards provided. In order to solve this problem, participants need to use inferential logic to determine which two of the four cards needed to be turned over to test the rule.

Only a small percentage of participants solve this problem correctly (usually less than 10%; see Wason & Johnson-Laird, 1972; Evans 1978; and, Johnson-Laird & Wason, 1977, for reviews of the card sorting task literature). Participants performed so poorly on this task that Wason argued that his experimental participants (undergraduate students) had not reached Piaget's stage of formal operational thought (Inhelder & Piaget, 1958). However, Griggs and Cox (1982) used the same task, requiring the same logic, and found that participants performed quite well. Specifically, Griggs and Cox created a set of objects and a rule that allowed participants to use their real-world knowledge (or more specifically, memories of their personal history) to solve the problem. As an example, imagine seeing four cards, each with an age, either over or under 21-years, and a type of beverage, either alcoholic or non-alcoholic, on it. Accompanying these cards is the rule, "If a person is under 21-years of age, he/she cannot have an alcoholic beverage". Participants who encountered problems of this type solved the problem correctly at a much higher rate (about 73%) than in the original Wason task. This effect became known as the Thematic-Materials effect (see Griggs & Cox, 1982; Johnson-Laird, Legrenzi & Sonino Legrenzi, 1972; Wason & Shapiro, 1971) and serves as an exemplar of the finding that people are better able to solve abstract and complex cognitive tasks when they can bring their real-world experience to those tasks to derive a solution – even though the nature of the Wason task itself did not change – just the materials from an abstract game involving letters and numbers to a more real-world context involving alcohol consumption. Much in the same way, we did not change the nature of

Wardlow Lane et al's (2006) task, just the materials from an abstract speaker-listener context to a more real-world buyer-seller scenario.

What then, if anything, can firms do to discourage employees from unintentionally leaking confidential information? One strategy is to focus employees' attention on the information that they do want to reveal, rather than on the information that they do not want to reveal, to increase the salience of the information that should be mentioned and lower (or at least not raise) the salience of confidential information. For example, employees' attention should be directed at the positive aspects of the products that they are selling rather than on the negative aspects or on the performance or value of their product as compared to other products.

Author Note

This research was supported by National Institutes of Health grant R01 MH64733 and R01 HD051030. We thank Brenna Gall, Amber Latronica, Erin Shea and Ray Suen for assistance with data collection and Victor Ferreira and Craig McKenzie for helpful discussions. Please address correspondence to Liane Wardlow Lane at the Department of Psychology 0109, University of California, San Diego, La Jolla, CA, 92093-0109. Email: lwardlow@ucsd.edu.

Chapter 4, in full, is being prepared for publication of the material as it appears in the manuscript: Wardlow Lane, L. & Liersch, M.J. *Be wary of paying people to keep secrets: Examining incentives to conceal privileged information.*

Table Captions

Table 1. Experiment Framing. The original Wardlow Lane et al (2006) task was converted to the more naturalistic and practiced context of buying and selling items.

Table 2. Point Structure for Experiments 1 and 2.

Table 3. Experiment 1. Number of contrasting trials on which 0, 1/2 or 1 point was earned according to participant role and pay-type.

Table 4. Experiment 1. Average points earned by sellers and buyers on each commission trial according to whether a modifier was or was not used. Sellers earned more points than buyers when modifiers were not used. Sellers earned fewer points than buyers when modifiers were used.

Table 5. Experiment 2. Number of contrasting trials on which 0, 1/2 or 1 point was earned according to participant role, pay-type and incentive-type.

Table 6. Experiment 2. Average points earned by sellers and buyers on each commission trial according to whether a modifier was or was not used. Sellers earned more points than buyers on trials that did not include a modifier. Sellers earned fewer points than buyers on trials that did include a modifier across both incentive-type conditions.

Chapter 4, Table 1

Experiment Framing. The original Wardlow Lane et al (2006) task was converted to the more naturalistic and practiced context of buying and selling items.

	Wardlow Lane et al (2006)	Current Experiment 1	Current Experiment 2
Environment	Speaker – listener	Seller – buyer	Seller – buyer
Reward	Object identification	Product exchange	Product exchange
System			
Items	Shapes	Products	Products
Incentive	Points	Points	Points & Money

Chapter 4, Table 2

Point Structure for Experiments 1 and 2.

	Seller	Buyer
Salary		
Initial transaction failed	0	0
Initial transaction correct	.5	.5
Commission		
Initial transaction failed	0	0
Initial transaction correct...		
... and Buyer does not guess	1	.5
... and Buyer guesses incorrectly	0	0
... and Buyer guesses correctly	.5	1

Chapter 4, Table 3

Experiment 1. Number of contrasting trials on which 0, 1/2 or 1 point was earned according to participant role and pay-type.

	Commission				Salary			
	No Modifier		Modifier		No Modifier		Modifier	
	Seller	Buyer	Seller	Buyer	Seller	Buyer	Seller	Buyer
0.0 pt	92	92	3	2	4	4	0	0
0.5 pt	81	154	38	14	340	340	39	39
1.0 pt	151	78	14	39	0	0	0	0

Chapter 4, Table 4

Experiment 1. Average points earned by sellers and buyers on each commission trial according to whether a modifier was or was not used. Sellers earned more points than buyers when modifiers were not used. Sellers earned fewer points than buyers when modifiers were used.

	No Modifier	Modifier
Seller	.59	.60
Buyer	.48	.84

Chapter 4, Table 5

Experiment 2. Number of contrasting trials on which 0, 1/2 or 1 point was earned according to participant role, pay-type and incentive-type.

	Lottery Payment							
	Commission				Salary			
	No Modifier		Modifier		No Modifier		Modifier	
	Seller	Buyer	Seller	Buyer	Seller	Buyer	Seller	Buyer
0.0 pt	47	47	2	2	0	0	0	0
0.5 pt	41	120	46	17	250	250	22	22
1.0 pt	120	41	17	46	0	0	0	0

	Course Credit							
	Commission				Salary			
	No Modifier		Modifier		No Modifier		Modifier	
	Seller	Buyer	Seller	Buyer	Seller	Buyer	Seller	Buyer
0.0 pt	54	54	1	1	2	2	0	0
0.5 pt	40	118	41	5	253	253	23	23
1.0 pt	118	40	5	41	0	0	0	0

Chapter 4, Table 6

Experiment 2. Average points earned by sellers and buyers on each commission trial according to whether a modifier was or was not used. Sellers earned more points than buyers on trials that did not include a modifier. Sellers earned fewer points than buyers on trials that did include a modifier across both incentive-type conditions.

	Lottery Payment	
	No Modifier	Modifier
Seller	0.68	0.62
Buyer	0.49	0.84

	Course Credit	
	No Modifier	Modifier
Seller	0.64	0.54
Buyer	0.29	0.93

	Collapsed across both incentive-types	
	No Modifier	Modifier
Seller	0.66	0.58
Buyer	0.48	0.88

Figure Captions

Figure 1. Example experimental set-up for Wardlow Lane et al (2006).

Figure 2. Example experimental stimuli for Experiments 1 and 2.

Figure 3. Experiment 1. Percentage of target descriptions including foil-contrasting modifiers as a function of whether the foil and target were the same (contrasting trials) or different products (non-contrasting trials), and whether sellers were given instructions to conceal the confidential product (commission trials) or not (salary trials). Error bars illustrate 95% confidence interval of the interaction.

Figure 4. Experiment 2. Percentage of target descriptions as a function of whether trials included foil-contrasting modifiers (contrasting trials) or not (non-contrasting) and whether sellers were given instructions to conceal the confidential product (commission trials) or not (salary trials) and whether they were given a lottery payment incentive or a course credit incentive. Error bars illustrate 95% confidence interval of the interaction.

Chapter 4, Figure 1
Example experimental set-up for Wardlow Lane et al (2006).

Chapter 4, Figure 2
Example experimental stimuli for Experiments 1 and 2.

Chapter 4, Figure 3

Experiment 1. Percentage of target descriptions including foil-contrasting modifiers as a function of whether the foil and target were the same (contrasting trials) or different products (non-contrasting trials), and whether sellers were given instructions to conceal the confidential product (commission trials) or not (salary trials). Error bars illustrate 95% confidence interval of the interaction.

Chapter 4, Figure 4

Experiment 2. Percentage of target descriptions as a function of whether trials included foil-contrasting modifiers (contrasting trials) or not (non-contrasting) and whether sellers were given instructions to conceal the confidential product (commission trials) or not (salary trials) and whether they were given a lottery payment incentive or a course credit incentive. Error bars illustrate 95% confidence interval of the interaction.

References

- Camerer, C. F. & Hogarth, R. M. (1999). The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of Risk and Uncertainty*, 19, 7-42.
- Demeski, J.S., Lewis, T.R., Yao, D., & Yildirim, H. (1999). Practices for managing information flows within organizations. *Journal of Law, Economics and Organization*, 15(1), 107-131.
- Evans, St. B.T. (1978). The psychology of deductive reasoning: Logic. In A. Burton and J. Radford (Eds.) *Thinking in Perspective*. Londong: Metheun.
- Flynn, F.J. (2006). Covert disclosures: Unauthorized leaking, public officials and the public sphere. *Journalism Studies*, 7(2), 256-273.
- Griggs, R.A. & Cox, J.R. (1982). The Elusive Thematic-Materials Effect in Wason's Selection Task. *British Journal of Psychology*, 73 (3), 407-20.
- HP Press Release (September 12, 2006). Patricia Dunn to remain HP chairman through January 2007 board meeting; board appoints Mark Hurd as successor. HP.com.
- Inhelder, B. & Piaget, J. (.1958). *The growth of logical thinking from childhood to adolescence.*, NY, NY: Basic Books.
- Johnson-Laird, P.N., Legrenzi, P. & Sonino Legrenzi, M. (1972). Reasoning and sense of reality. *British Journal of Psychology*, 63, 395-400.
- Johnson-Laird, P.N. & Wason, P.C. (1977). A theoretical analysis of insight into a reasoning task. In P.N. Johnson-Laird and P.C. Wason (Eds) *Thinking: Readings in Cognitive Science*. Cambridge: Cambridge University Press.
- Kanellos, M. (February 27, 2007). HP board 'was split by battle for control'. CNET News.com.
- Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review*, 1, 476-490.
- McKenzie, C. R. M. (2005). Judgment and decision making. In K. Lamberts & R. L. Goldstone (Eds.), *Handbook of cognition* (pp. 321-338). London: Sage.
- McKenzie, C.R.M., Liersch, M.J. & Finkelstein, S.R. (2006). Recommendations Implicit in Policy Defaults. *Psychological Science*, 17 (5), 414-420.

- McKenzie, C. R. M. (2004). Framing effects in inference tasks -- and why they are normatively defensible. *Memory and Cognition*, 32, 874-885.
- Sher, S., & McKenzie, C. R. M. (2006). Information leakage from logically equivalent frames. *Cognition*, 101, 467-494
- Wardlow Lane, L. & Ferreira, V.S. (2007). Speaker-external versus speaker-internal forces on utterance form: Do cognitive demands override threats to referential success? Unpublished manuscript.
- Wardlow Lane, L., Groisman, M. & Ferreira, V.S. (2006). Don't talk about pink Elephant! Sellers' control over leaking private information during language production. *Psychological Science*, 17 (4), 273-277.
- Wason, P.C. & Johnson-Laird, P.N. (1972). *Psychology of Reasoning: Structure and Content*. Cambridge: MA: Harvard University Press.
- Wason, P.C., & Shapiro, D. (1971). Natural and contrived experience in a reasoning problem. *Quarterly Journal of Experimental Psychology*, 23, 63-71.
- Wegner, D.M. (1994a). Ironic processes of mental control. *Psychological Review*, 101, 34-52.
- Wegner, D.M. (1994b). Pink elephant tramples white bear: The evasion of suppression. *Psychology*, 5(40), Paradoxical Cognition, 2.
- Wegner, D.M., Ansfield, M., & Pilloff, D. (1998). The putt and the pendulum: Ironic effects of the mental control of action. *Psychological Science*, 9, 196-199.

CHAPTER 5

GENERAL DISCUSSION

Together, the three studies in this dissertation characterize influences on the forms speakers' utterances take with respect to perspective-taking behavior. The results of Study 1 demonstrate that instructing speakers to avoid making implicit reference to privileged objects does not reduce their tendency to do so. In fact, the opposite pattern was observed; when speakers were instructed to avoid making implicit references to privileged objects when naming common ground objects, they were even more likely to do so. This finding suggests that speakers do not have complete communicative control over the extent to which they leak information that they want to keep private. This lack of control suggests that active conceptual features may become part of speakers' utterances even if speakers intend *not* to include those particular features. This result implies that attempts to avoid revealing some particular piece of information makes it more likely that that information will be revealed. This result may inform theories pertaining to the types of social interactions whereby speakers attempt to conceal information such as adversarial negotiations.

The results of Study 1 suggested that the factor driving speakers' implicit references to privileged information was the salience of privileged information. Study 2 directly assessed this possibility. The results of Study 2 demonstrate that when speakers must cope with speaker-internal cognitive pressures that serve to increase the salience of privileged information along with speaker-external communicative pressures that make implicit references to privileged information communicatively harmful, they are more sensitive to the speaker-internal cognitive pressure. Increasing the salience of privileged information makes it more likely that speakers will make

implicit reference to that information even though doing so can be communicatively harmful. This finding suggests that speakers cannot fully limit their utterances to shared information when privileged information is made salient.

The results of Studies 1 and 2 suggest that speakers are quite poor at keeping privileged information privileged. However, because the task employed was arguably abstract in nature, the question of whether the results would generalize to other contexts remained. Study 3 assessed whether speakers' performance could be improved with the use of a less abstract task and with a highly salient incentive. The results of Study 3 demonstrate that speakers cannot limit their utterances to privileged information even when they are able to bring their real-world knowledge to bear on a task. The results further suggest that using a salient monetary incentive to increase speakers' efforts to avoid making implicit reference to privileged information does not aid them in their attempts to do so. In fact, a monetary incentive increases the likelihood that speakers will make implicit references to privileged information. This finding suggests that in the business world, firms should avoid stressing what information should be kept privileged because highlighting privileged information may increase its likelihood of being unintentionally revealed.

In the Introduction, I proposed seven questions that this dissertation research is meant to address. Here, I return to those original questions and attempt to answer them in light of the results of the three reported studies. (1) What effect does the salience of privileged information have on speakers' perspective-taking behavior? Does increasing the salience of privileged information make it more likely that that

information will become highly active and thus *more* likely to be mentioned? Or, does increasing the salience of privileged information highlight the fact that the information is privileged and thus make it *less* likely to be mentioned? The results of all three studies show that making privileged information salient results in that information making its way into speakers' utterances at a higher rate than when privileged information is not made salient. These results suggest that increasing the salience of privileged information causes that information to become more highly active, and thus more likely for selection. Furthermore, the results suggest that highlighting privileged information does not serve to highlight the fact that that information is privileged to speakers – or at the very least not in a meaningful way that reduces the extent to which that information is selected for production.

(2) Can speakers limit their implicit references to privileged information when doing so is potentially harmful to communication rather than when doing so is pragmatically awkward? The results of all three studies suggest that speakers cannot control the degree to which they make implicit references to privileged information even when doing so is potentially harmful to communication (Study 2) or harmful to their task-based goals (Studies 1 and 3). In fact, in all of the studies, increasing the potential harm of making implicit references to privileged ground increased the amount that speakers did so.

(3) How does the production system limit the intrusions of active privileged conceptual features in speakers' utterances? Can speakers use their real-world knowledge to help them determine which features should be mentioned? The results

of Study 3 suggest that speakers are unable to use their real-world knowledge and experiences to help them avoid making implicit references to privileged information. However, this result should be further studied using a task or task environment that is even closer to participants' real-world experiences to ensure that it generalizes.

(4) Will increased effort to avoid making implicit references to privileged objects help speakers limit their utterances to shared conceptual features? Or, will an increased effort to avoid making implicit references to privileged objects cause the representations of those objects to become highly active, and thus more likely to be mentioned? The results of Study 3 suggest that increasing incentive to avoid mentioning privileged information, which should lead to an increase in effort on the part of speakers, does not result in fewer implicit mentions of that information. In fact, the opposite pattern was observed: Monetary incentive increased speakers' implicit mentions of privileged information as compared to a non-monetary incentive. This result, along with the results of Study 2, suggests that the monetary incentive served to increase the salience of the privileged object. Even though the increased salience was the result of an intention to avoid mentioning that information, it likely resulted in a boost of activation of the conceptual features associated with that information which, in turn, resulted in that information being mentioned more often.

The thrust of this research focuses on an error that speakers make in that they fail to account for their addressees' different perspective when they make implicit references to privileged information. However, I caution readers from using this research to support a claim that speakers are unskilled (or just plain bad) at taking

their addressees' perspectives into account when speaking. Notice that in the majority of experiments, there was a baseline condition in which speakers were quite good at taking their addressees' perspectives into account. It was only the conditions whereby speakers were so overwhelmed by the cognitive pressure to pay attention to privileged objects that they so often failed to take perspective differences into account.

Though performance suffered in this task when privileged information was made especially salient (and though that finding is informative about the causes of speakers' inappropriate references to privileged information), it could easily be true that when speakers are moving about in their daily lives, the things that their attention is drawn to happen also to be the things that they are, or should be, talking about. Furthermore, in everyday conversation, speakers and addressees are able to work together to establish understanding. Errors such as those reported in this work may be less harmful in everyday conversation because addressees can typically signal their understanding and speakers can make adjustments based upon that feedback.

Although I caution the reader not to draw a conclusion that speakers are bad at taking perspective differences into account, these three studies are still informative in that they suggest that speakers have little control over whether they include privileged information in their utterances especially under conditions that make privileged information salient. This lack of control occurs even in situations where making references to privileged information is potentially harmful to the communicative intention of producing the utterance in the first place and when doing so is harmful to

speakers' task-based goals. These results suggest that speakers who can effectively ignore privileged information will have the best chance of avoiding references to that privileged information.