Managing Privileged Knowledge about Identity in Language Comprehension

by

Mindaugas Mozuraitis

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Department of Psychology
University of Toronto

© Copyright by Mindaugas Mozuraitis (2015)
Managing Privileged Knowledge about Identity in Language Comprehension

Mindaugas Mozuraitis
Doctor of Philosophy
Department of Psychology
University of Toronto
2015

Abstract

A central claim in research on language comprehension is that assumptions about other’s knowledge are essential to guide language users’ understanding of utterances in both interactive conversations and fictional narratives. Past studies have suggested that listeners or readers are able to suppress privileged knowledge about the existence of entities from the earliest moments of processing. In the present dissertation, it is investigated whether this ability extends to cases where comprehenders possess privileged knowledge about the identity of an entity. This question is motivated by the central role played by identity in various semantic phenomena, as well as children’s well-documented difficulties in managing “identity” discrepancies in theory of mind tasks.

Using a visual world eye-tracking paradigm, Experiment 1 shows that knowledge that is in conflict with an object's outward features (e.g., knowing that the object that looks like a lightbulb is actually a candle) is readily accessible to comprehension systems as listeners hear object descriptions unfold in time. Experiment 2 then verifies that, without the foreknowledge,
the visually-misleading objects are identified in terms of their outward appearance. The validated stimulus objects are then used in Experiments 3 and 4, which evaluate the extent to which listeners suppress their privileged knowledge about object identity during real-time conversation. Similarly, Experiments 5 and 6 are conducted to explore whether listeners suppress privileged knowledge about the identity of entities during the comprehension of fictional narratives. In this case, interpretation patterns are examined as listeners hear the discourse “spoken” by fictional characters whose knowledge states are again varied.

The results from both kinds of tasks reconfirm previous findings showing that language users rapidly take into account a speaker’s awareness of the existence of a referential entity. In contrast, however, the experiments show that comprehenders do not suppress private knowledge about the identity of an entity during referential processing—an effect that occurs even when the perspective of another individual is emphasized and regardless of whether the context involves conversational interaction or fictional narratives. Together, the results reveal important differences in the way in which distinct types of perspective-based cues are integrated in real-time language comprehension.
Acknowledgments

This work would not have been possible without the guidance, support, and gentle humour of my supervisors, Craig Chambers and Meredyth Daneman. I am forever in debt to them for constantly challenging me to strive for originality of thought and creativity in my research. I could not have asked for better mentors. My gratitude also goes to my committee member Eyal Reingold for his wisdom and advice throughout the research and writing process and Patricia Ganea, Sarah Brown-Schmidt, and, especially, Daphna Heller for their insightful feedback on this dissertation.

In addition, I would like to thank my two academic big sisters Cara Tsang and Christine Burton. Their friendship and guidance helped me navigate the sometimes-bewildering rules and conventions of academic as well as social life in graduate school. My thanks also goes to Ronda Lo, Piumi Rajasekera, Daniel De Santis, Robert Redford, and Joshua Romano for their assistance with data collection and data coding.

Finally, I would like to express my gratitude to my parents, my sister, and, especially, my wife Inga Dievulyte, with whom I had the privilege of sharing the happy moments of my graduate career with and whose support and constant encouragement helped me through the hard times of the program. I have been incredibly lucky to have them rooting for me for the past 6 years.
Table of Contents

Acknowledgments .................................................................................................................. iv
Table of Contents .................................................................................................................. v
List of Tables .......................................................................................................................... viii
List of Figures .......................................................................................................................... ix
List of Appendices ................................................................................................................... xi

Chapter 1: Background and Overview ................................................................................. 1
1.1 Introduction ....................................................................................................................... 1
1.2 Common ground in communication .............................................................................. 4
1.3 Common ground in real-time language processing ....................................................... 8
  1.3.1 Incrementality in real-time language processing ................................................. 8
  1.3.2 Time course of common ground integration .................................................... 11
1.4 Dissertation overview .................................................................................................... 17

Chapter 2: Appearance vs. Reality in Real-Time Referential Processing ....................... 20
2.1 Experiment 1 ...................................................................................................................... 23
  2.1.1 Method ...................................................................................................................... 25
    2.1.1.1 Participants ....................................................................................................... 25
    2.1.1.2 Materials and design ...................................................................................... 25
    2.1.1.3 Procedure ....................................................................................................... 29
  2.1.2 Results and Discussion ............................................................................................. 30
    2.1.2.1 Coding ........................................................................................................... 30
    2.1.2.2 Statistical modeling ...................................................................................... 30
    2.1.2.3 Eye movement data ...................................................................................... 31
2.2 Summary ......................................................................................................................... 33

Chapter 3: Privileged Knowledge about Object Existence vs. Object Identity in Conversation ......................................................................................................................... 36
3.1 Experiment 2 ...................................................................................................................... 42
3.1.1 Method .................................................................................................................. 44
  3.1.1.1 Participants........................................................................................................ 44
  3.1.1.2 Materials and design ......................................................................................... 44
  3.1.1.3 Procedure ........................................................................................................ 45
  3.1.1.4 Coding............................................................................................................... 45
3.1.2 Results and Discussion ......................................................................................... 45

3.2 Experiment 3 ........................................................................................................... 46
  3.2.1 Method ............................................................................................................... 46
    3.2.1.1 Participants........................................................................................................ 46
    3.2.1.2 Materials and design ......................................................................................... 47
    3.2.1.3 Procedure ........................................................................................................ 50
    3.2.1.4 Predictions ....................................................................................................... 52
  3.2.2 Results and Discussion ....................................................................................... 53
    3.2.2.1 Coding............................................................................................................... 53
    3.2.2.2 Statistical modeling.......................................................................................... 53
    3.2.2.3 Eye movement data........................................................................................ 54

3.3 Experiment 4 ........................................................................................................... 57
  3.3.1 Method ............................................................................................................... 59
    3.3.1.1 Participants........................................................................................................ 59
    3.3.1.2 Materials and design ......................................................................................... 60
    3.3.1.3 Procedure ........................................................................................................ 62
  3.3.2 Results and Discussion ....................................................................................... 62
    3.3.2.1 Coding............................................................................................................... 62
    3.3.2.2 Statistical modeling.......................................................................................... 62
    3.3.2.3 Eye movement data........................................................................................ 62

3.4 Summary .................................................................................................................. 64

Chapter 4: Knowledge Discrepancies Involving Identity in Narrative Comprehension...... 68

4.1 Experiment 5 ........................................................................................................... 74
  4.1.1 Method ............................................................................................................... 74
    4.1.1.1 Participants........................................................................................................ 74
    4.1.1.2 Materials and design ......................................................................................... 74
    4.1.1.3 Procedure ........................................................................................................ 79
  4.1.2 Results and Discussion....................................................................................... 79
List of Tables

Table 2.1. List of target objects and the three types of competitors used in critical trials in Experiments 1, 3, and 4. .................................................................27

Table 3.1. Knowledge State manipulation in Experiment 3. ...........................50

Table 4.1. Example of the three versions of experimental narratives used in Experiments 5 and 6 and corresponding critical comprehension questions used in Experiment 5. ......................78
List of Figures

Figure 1.1. Illustration of a critical display from the addressee's and the speaker's (director's) perspectives, used in Keysar et al. (2000). Source: Keysar et al. (2000). ........................................ 12

Figure 1.2. Illustration of a critical display from the addressee's perspective in Hanna et al. (2003). Source: Hanna et al. (2003). ................................................................. 14

Figure 1.3. Illustration of a critical display from the addressee's perspective in Heller et al. (2008). Source: Heller et al. (2008). ................................................................. 15

Figure 2.1. A regular candle (Panel A), a candle that looks like a lightbulb (Panel B), and a regular lightbulb (Panel C), illustrating the type of competitor manipulation common to Experiments 1, 2, 3, and 4. ................................................................. 26

Figure 2.2. Illustration of a critical display with target (candy), competitor (*candle that looks like a lightbulb), and unrelated object (tennis ball) in Experiment 1. ........................................ 27

Figure 2.3. Likelihood of a saccade to the competitor as a function of competitor type, Experiment 1. Note: Although the analyses were conducted in logit space, raw likelihood values are reported for easier interpretation. ................................................................. 32

Figure 3.1. Illustration of a critical display (from the addressee’s perspective) as a function of knowledge state (the display was identical for the shared-knowledge and privileged-identity conditions) and competitor type as well as an example of the critical part of the verbal instructions given by the confederate speaker. Notes: (1) Asterisk indicates visually misleading object; (2) the black background within a display compartment represents an occluder panel whereby an object was visible to the listener but not to the speaker. ............................. 48

Figure 3.2. Bird’s-eye view of the physical arrangement of the participant (addressee) and confederate speaker in Experiments 3 and 4. ................................................................. 52

Figure 3.3. Likelihood of a saccade to the competitor as a function of competitor type and addressee’s understanding of the speaker’s knowledge state, Experiment 3. ............................ 55

Figure 3.4. Example sequencing of the context and critical trials with the knowledge-based competitor object and corresponding instructions in Experiment 4. ............................ 61

Figure 3.5. Likelihood of a saccade to the competitor as a function of competitor type and linguistic context, Experiment 4. ................................................................. 63

Figure 4.1. Example critical display used in Experiment 6. Top-left: side-character; Top-right: target object; Bottom-left: unrelated object; Bottom-right: protagonist. ............................ 83
Figure 4.2. Likelihood of a saccade to the side-character as a function of narrative type, Experiment 6. .................................................................
List of Appendices

Appendix A. The proportion of fixations to the target and competitor objects over time, across conditions (Experiment 1). The box with dashed lines represents the analysis window (200-700 ms after target noun [e.g., candy] onset). ........................................................................................................109

Appendix B. The proportion of fixations to the target and competitor objects over time, across conditions (Experiment 3). The box with dashed lines represents the analysis window (200-700 ms after target noun [e.g., candy] onset). ........................................................................................................110

Appendix C. The proportion of fixations to the target and competitor objects over time, across conditions (Experiment 4). The box with dashed lines represents the analysis window (200-700 ms after target noun [e.g., candy] onset). ........................................................................................................112

Appendix D. The proportion of fixations to the target and side-character over time, across conditions (Experiment 6). The box with dashed lines represents the analysis window (150-550 ms after target noun [e.g., raincoat] onset). ........................................................................................................113

Appendix E. All critical narratives used in Experiments 5 and 6 and critical comprehension questions used in Experiment 5. ........................................................................................................114
Chapter 1: Background and Overview

1.1 Introduction

Language comprehension requires the access and integration of many different streams of information within a short period of time. For example, in the case of spoken language, speech sounds in an unfolding utterance must be identified and mapped onto items in the mental lexicon, and the hierarchical relationships among lexical items must be identified according to the syntactic rules of the language. Further, the broader meaning of the sentence must be computed by enriching and integrating the linguistic content with discourse-level information. For example, in a conversation, one needs to infer other interlocutors’ intentions and adjust to their knowledge and beliefs about the world. When someone asks “Could you open the window?”, the listener will often understand that the intention of the utterance is a request to open the window rather than a question about the ability to open it. On the other hand, if the same utterance is directed to a toddler, the purpose of the utterance could, in fact, be a question about the ability to open the window rather than a polite request to actually do it. Understanding and tracking mental states is also an important aspect of understanding fictional stories. Indeed, comprehension of more complex prose would be impossible, if we could not differentiate between the beliefs we hold as readers or listeners and the beliefs that different characters of a story have about their reality.

Contemporary perspectives on language comprehension adopt the view that the mechanisms underlying language processing are characterized by considerable immediacy and incrementality: a range of processes are initiated from the earliest moments of encountering the input, and the evolving mental representations are continuously adjusted as additional information becomes available. This view is informed by a large number of studies exploring a variety of phenomena such as syntactic, semantic, and referential processing and rests on
evidence drawn from a wide range of experimental techniques (for reviews, see Dahan & Magnuson, 2006; Rayner & Clifton, 2009). Consequently, a central task of psycholinguistics becomes to delineate how these different processing tasks work together as language is encountered in time. For example, one classic question is whether, as suggested by a number of researchers (e.g., Frazier & Clifton, 1996), syntactic analysis receives some kind of priority before other types of higher-level processes can take place or, in contrast, whether syntactic, semantic, and pragmatic processes are interdependent and operate on unfolding operations simultaneously as language input is encountered (e.g., Tanenhaus & Trueswell, 1995).

A more contemporary question in studies of real-time comprehension, which is more central to the present dissertation, concerns individuals’ ability to coordinate their knowledge states with those of others. This is an important question because a number of linguistic phenomena such as figurative language as well as the process of choosing among possible referring expressions cannot be interpreted without explicit mental representation of a speaker’s knowledge or intentions. For example, the correct interpretation of a pronoun such as “he” is possible only if the addressee can identify the intended referent. Although there is broad consensus that perspective taking of this sort is a necessary component of the semantic interpretation of language (e.g., Clark & Marshall, 1981), the questions for mechanistic models of interpretation involve the details of how comprehenders bypass their own privileged knowledge as the various components of linguistic processing occur. According to one view, inferences about the speaker’s knowledge state receive low priority and are considered only if other analyses fail to produce a coherent interpretation (e.g., Keysar 1994; Keysar, Barr, Balin, & Brauner, 2000; Keysar, Lin, & Barr, 2000; Barr, 2008). Other researchers argue that knowledge differences are taken into account at the earliest moments of processing in live conversation between two individuals (e.g., Brown-Schmidt, Gunlogson, & Tanenhaus, 2008; Hanna,
Tanenhaus, & Trueswell, 2003; Heller, Grodner, & Tanenhaus, 2008), as well as in the context of narrative comprehension by a single individual, where the comprehender must keep track of fictional characters’ knowledge state (e.g., Gerrig, Brennan, & Ohaeri, 2001; Gerrig, Ohaeri, & Brennan, 2000). Live-conversation studies addressing this question typically involve two individuals whose knowledge about the referential entities is either the same or different. The status of knowledge is typically manipulated by varying the speaker’s awareness of physical co-presence of the entities that are available for reference. For example, an object may be visually occluded from the perspective of the speaker, thereby preventing that partner from knowing whether there is an object behind the occluding surface (and, consequently, what it is). In the narrative comprehension studies, the question is typically addressed using offline questionnaires that probe readers’ assumptions about the story characters’ knowledge, or by investigating whether readers experience processing difficulties when linguistic input is inconsistent with the expectations about the story characters’ knowledge.

A distinguishing feature of the work reported in this dissertation is a focus on inferences about the speaker’s knowledge state in terms of objects’ “identity” in situations where objects’ physical co-presence is already established. In experiments that use conversational contexts, this is achieved through the use of objects whose outward appearance is potentially misleading when it comes to the task of determining the category to which the object belongs (e.g., an object which looks like a lightbulb but is a candle). In experiments that explore narrative comprehension, knowledge about the features pertaining to an entity’s identity is manipulated by varying how much a narrative protagonist is stated to know about another character in the story (e.g., the protagonist knows a man but does not know that he is a ranger). By focusing on knowledge discrepancies involving the identity of an entity and by directly comparing them to more well-studied cases involving knowledge about the co-presence of entities, it is possible to
provide new and important insights into the underlying processes involved in managing
inferences about knowledge discrepancies in conversational and narrative contexts. In the
remainder of this chapter, I discuss the relevant literatures on perspective taking in
conversational and narrative contexts in greater detail, and outline the series of studies that will
be presented in the following chapters.

1.2 Common ground in communication
“Common ground” is defined as the information and beliefs assumed to be shared among
conversational participants. As suggested by Clark and Marshall (1981), common ground can be
formed on the basis of community membership, the contents of linguistic interactions
(sometimes called “linguistic co-presence”), shared physical environments (physical co-
presence), or a combination of all of these sources. For example, we assume that people who live
in the same community (e.g., are citizens of Lithuania) will share certain knowledge (e.g., will
know the current president of Lithuania). In addition, it is usually reasonable to assume that if
something was mentioned in a conversation, all participants of the conversation know that piece
of information. Finally, if all conversational partners can see an object in front of them, they can
safely assume that they all know about the object’s presence in the situation.

Furthermore, Clark, Schreuder, and Buttrick (1983, see also Clark, 1996; Clark &
Wilkes-Gibbs, 1986) have proposed that common ground serves as the primary context for
language understanding. According to this view, speakers design their utterances in such a way
that they have a good reason to believe that addressees can readily and uniquely process what
was meant on the basis of how the utterance relates to what is known to all interlocutors versus
what is the speaker’s privileged knowledge. For example, if a speaker utters “There’s that girl”,
she/he believes that the intended addressee can identify the unique and intended girl based on the
information they both share. Similarly, on the comprehension side of the communicative process,
addressees are understood to evaluate a speaker’s utterances (e.g., “that girl”) against the established common ground (e.g., by performing a kind of search for the girl believed to be mutually known with the speaker). Also, consider another example in which “that girl” occurs in a question, as when the speaker asks, “Who is that girl?”. In this case, a rational addressee would expect the description to refer to someone that is not mutually known (and more specifically someone who the speaker does not know about, see Grice, 1975). All these examples illustrate that successful communication requires keeping track of what is known to all conversational partners (what is in common ground) and what is known only to oneself (privileged information).

Psycholinguistic studies on this topic have provided various kinds of empirical evidence showing that speakers take common ground into account when constructing their utterances (e.g., Clark & Wilkes-Gibbs, 1986; Fussell & Krauss, 1989; 1992), and that addressees use common ground to contour their understanding of utterances (e.g., Gerrig & Littman, 1990; Gibbs, Mueller, & Cox, 1988; Lea, Mason, Albrecht, & Birch, 1998). For example, in Clark and Wilkes-Gibbs (1986) participants were seated at tables separated by an opaque screen. Each participant saw a set of cards depicting ambiguous complex figures. For the speaker, the cards were already arranged in a target order while, for the addressee, an identical set of cards was arranged in a different order. The speaker’s task was to provide instructions to the addressee so that he/she could rearrange his/her cards to match the target order as fast and as accurately as possible. The participants were allowed to talk back and forth without any restrictions. Clark and Wilkes-Gibbs found that, as the experiment progressed, participants were able to rearrange the cards faster, with fewer turns, fewer confusions, and shorter referring expressions. These results showed that, throughout the course of the experiment, people were able to establish common ground about the cards and were effective in using the information that was in common ground
to understand as well as to construct increasingly efficient descriptions.

In another study, Lea et al. (1998) explored readers’ sensitivity to other’s knowledge in narrative comprehension. The authors asked their participants to read the following stories:

(1) Jane was dreading her dinner with her cousin, Marilyn. She complained loudly to her roommate Gloria. “Every time I go to dinner at my cousin’s I get sick.” Gloria asked, “Why did you agree to go?” Jane said, “Because I’m too wimpy to say no.” Jane went off to have dinner. Gloria decided to cook something nice for herself for dinner. “As long as I’m home alone,” she thought, “I’ll eat well.” Gloria searched her refrigerator for ingredients. She found enough eggs to make a quiche. After dinner, she put the dishes in the dishwasher. Gloria was still up when Jane arrived home about midnight. Gloria asked Jane, “Did she play you old disco records?” Jane chuckled and said, “I can’t get Disco Inferno out of my mind.”

(2) Jane and her roommate Gloria were leaving work. “Are you headed home?” asked Gloria. “No, but I’ll see you later tonight,” replied Jane. Jane drove off to have dinner with her cousin, Marilyn. As she drove, she started to have regrets. She usually got sick when she ate at her cousin’s. Jane wondered why she had agreed to go. She decided she was just too wimpy to say no. Meanwhile, Gloria went home and decided to cook something nice for herself. “As long as I’m home alone,” she thought, “I’ll eat well.” Gloria searched her refrigerator for ingredients. She found enough eggs to make a quiche. After dinner, she put the dishes in the dishwasher. Gloria was still up when Jane arrived home about midnight. Gloria asked Jane, “Did she play you old disco records?” Jane chuckled and said, “I can’t get Disco Inferno out of my mind.”

The above narratives differ in terms of the degree of mutual knowledge shared by the main
characters of the story (e.g., in (1) Gloria knows that Jane is going to have dinner with her cousin whereas in (2) she does not know anything about Jane’s dinner plans). The target sentence of the story (e.g., *Gloria asked Jane: “Did she play you old disco records?”*) was congruent with the target character’s (e.g., Gloria’s) knowledge state in the first context because it is perfectly reasonable for Gloria to ask about Jane’s cousin and the dinner when she knows about the event. However, it was incongruent with Gloria’s knowledge state in the second context because Gloria should not be inquiring about the dinner when there is no previous evidence that she knows about it. The reading times for the target sentence were shorter when the test sentence was congruent (as in [1]) with the target character’s knowledge state than when it was not (as in [2]). Lea et al. took this as evidence that the readers were sensitive to the fact that in (2), the test sentence was inappropriate because the target character had no basis for knowing who is the entity referred to with the pronoun (e.g., “she”). This finding suggests that comprehenders were sensitive to the story characters’ knowledge states when interpreting the characters’ utterances. However, the paradigms and measures used by Lea et al. and Clark and Wilkes-Gibbs cannot establish how language users process utterances as they are read or heard in real time, particularly at the precise point of encountering incongruent information. For example, it is not clear whether the processing difficulty observed in Lea et al. happened at the pronoun or later on in the target sentence. As mentioned above, answering these questions is critical in order to establish how common ground information is used within the core processes involved in language comprehension. Work on this topic is described in the next section, where I first briefly illustrate how experimental methodologies are used to explore the time-course of spoken language processing at the level of individual words.
1.3 Common ground in real-time language processing

1.3.1 Incrementality in real-time language processing

Spoken word recognition is characterized by the incremental evaluation of each unfolding sound segment in the phonetic signal against the sound pattern of different entries in the mental lexicon. For example, as the word “candy” unfolds, initially the sounds are consistent with a number of lexical entries (e.g., /kæn…/ is consistent with words such as “candle”, “Canada”, and “canteen” among others). As additional sounds become available, the number of viable candidates will gradually decrease until a single candidate is identified. This has led spoken word recognition to be understood in terms of lexical “competition” where the activation level of competitors rises and falls dynamically as the phonetic signal unfolds.

One of the studies which provides evidence for the immediate and incremental processing of speech information during spoken word recognition, using the same paradigm used in the current dissertation, is a study by Allopenna, Magnuson, and Tanenhaus (1998). In this study participants followed spoken instructions, which asked them to interact with computer displays containing multiple clip-art objects. The critical displays contained a target object (e.g., a beaker), a phonological competitor whose name started with the same speech sounds as the target’s name (e.g., a beetle), a rhyme competitor whose name rhymed with the target’s name (e.g., a speaker), and a phonologically unrelated distractor (e.g., a stroller). As participants carried out a spoken instruction to click on the target object and move it to a specific location on the display, their eye movements were monitored. Allopenna et al. observed that, approximately 150-200 ms after the onset of the target name (e.g., “beaker”), the proportion of fixations to the target object (e.g., a beaker) started to increase. Furthermore, during the first part of the target name, the proportion of fixations to the phonological competitor (e.g., a beetle) increased to the same extent as to the target object, and significantly more than to the unrelated distractor (e.g., a
stroller). Once the uniqueness point was encountered (e.g., /k/ in “beaker”), the proportion of fixations to the phonological competitor started to decrease, while the proportion of fixations to the target object continued to increase.

Given that it takes around 150-200 ms to plan and execute an eye movement (Hallet, 1986), these results show that eye movements are closely time-locked to speech processing and that visual world eye-tracking paradigm provides a way to capture the immediate and incremental processing of spoken language. From a methodological point of view, another important feature of visual world eye-tracking paradigm used in Allopenna et al. is interactivity with a concrete and behaviourally relevant referential context. In other words, it allows the use of a situation where participants are not only required to evaluate a speech signal against the sound pattern of different entries in the mental lexicon, but where they also need to evaluate this incoming information against potential referents (e.g., objects existing in an external visual world) in order to successfully perform the task (e.g., to click on or pick up the intended referent). This provides a way to study not only the mapping between the speech signal and stored lexical forms but also the effects of situation-specific factors, such as previous linguistic context, on the mapping between the speech signal and referents available in the visual domain.

Previous work has suggested that despite the match with unfolding speech input, lexical candidates that are semantically incompatible with earlier linguistic input show reduced activation. For example, Dahan and Tanenhaus monitored eye movements as participants listened to Dutch sentences (e.g., “Nog nooit klom een bok zo hoog” [Never before climbed a goat so high]) and moved objects mentioned in those sentences (e.g., a goat) to specific locations on a visual display. Along with the target object, the displays contained a phonological competitor (e.g., bot [bone]), a semantic competitor (e.g., spin [spider]), and an unrelated distractor (e.g., eiland [island]). Critically, the verb preceding the target name was either
constraining (e.g., “Nog nooit klom een bok zo hoog” [Never before climbed a goat so high]) or not constraining (e.g., “Nog nooit is een bok zo hoog geklommen” [Never before has a goat climbed so high]). In the non-constraining-verb condition, the proportion of fixations to the phonological competitor (e.g., bot [bone]) gradually increased until the point at which it was differentiated from the target word (as in Allopenna et al.). In contrast, in the constraining-verb condition, the proportion of fixations to the phonological competitor did not increase. In fact, in this case, the pattern of eye movements for the phonological competitor was similar to the one for the unrelated distractor (e.g., eiland [island]). The findings suggested that verb semantics is used to restrict the referential domain simultaneously with the available phonetic information. In other words, although “bok” (goat) and “bot” (bone) share initial sounds, “bot” cannot perform the action of climbing and, thus, receives only weak consideration as the target name “bok” unfolds.

The above-mentioned findings suggest that visual world eye-tracking paradigm is sensitive to the immediacy and incrementality of combining semantic information from previous lexical items during the unfolding speech signal to narrow the consideration of viable referents from the set available within a visual context. Similarly, using the same paradigm, other work has shown that listener’s understanding of possible actions in the physical context also has the potential to narrow the referential domain (e.g., Chambers, Tanenhaus, Eberhard, Filip, & Carlson, 2002). In other words, the visual world eye-tracking paradigm was demonstrated to be sensitive enough to assess whether the language processing system is able to engage various kinds of contextual processes fast enough to affect interpretation as it occurs in real time. It is therefore perhaps not surprising that researchers started to use this paradigm to study the time course of common ground integration in conversational interaction.
1.3.2 Time course of common ground integration

In a typical live-conversation task, the information that is shared versus privileged is manipulated either by visual co-presence (e.g., some objects can be seen by both participants and some can be seen by only one of the interlocutors), or linguistic co-presence (e.g., some information is mentioned to both participants but some is mentioned only to one of them). By creating situations where, unlike addressees, speakers are unaware of the presence of some of the objects (see Barr, 2008; Brown-Schmidt, 2012; Brown-Schmidt et al., 2008; Ferguson & Breheny, 2011, 2012; Hanna et al., 2003; Heller et al., 2008; Keysar et al., 2000; Keysar, Lin, & Barr, 2003; Lin, Keysar, & Epley, 2010; Nadig & Sedivy, 2002), it is possible to examine whether addressees are sensitive to the knowledge discrepancy between their knowledge and speakers’ knowledge when understanding speakers’ instructions. For example, in Nadig and Sedivy, critical displays contained multiple objects from the same category such as two glasses, one of which was not visible to the speaker. By monitoring addressees’ eye movements as they interpreted the instructions by the speaker to move “the glass”, the authors were able to measure to what extent addressees looked at the glass that was not visible to the speaker. The amount of visual consideration of the object that is privileged to the addressee in this way (e.g., a glass visible to the addressee but not to the speaker) indicates to what extent listeners are implicitly and temporarily considering them as potential referents (addressees rarely make mistakes in their final judgment of the object intended by the speaker).

Initial investigations of the time course of common ground integration in language processing yielded contradictory results. Some studies suggested that a listener's egocentric perspective predominates during initial stages of language processing, and that the integration of common ground occurs only after this egocentric phase (e.g., Barr, 2008; Keysar, Barr, & Horton, 1998; Keysar et al., 2000; Keysar et al., 2003; Kronmüller & Barr, 2007; Pickering &
Garrod, 2004). This view was motivated by the assumption that the process of tracking what is common and what is privileged information is resource intensive, and that in most cases the information shared between two interlocutors will overlap anyway, rendering the demanding task of perspective taking unnecessary. For example, in Keysar et al. (2000), critical displays contained multiple objects from the same category (see Figure 1.1.), such as candles that differed in size (e.g., small, medium, and large). On critical trials, the addressee (participant) was instructed by the speaker (confederate) to follow instructions like “pick up the small candle” in a situation where the smallest candle visible to the addressee (the smallest candle of three) was not visible to the speaker. Keysar et al. (2000) reported that addressees’ eye movement behavior reflected a tendency to consider a candidate referent that was not visible to the speaker. Specifically, as the target referring expression (e.g., “small candle”) was unfolding in time, addressees initial eye movements to the candle not visible to the speaker were faster than to the smallest candle visible to both the addressee and the speaker (e.g., the medium candle from the addressee’s perspective). Only well after the offset of the target referring expression did participants direct their attention to and perform the requested action with the intended candle (e.g., the medium candle). The authors took these findings as evidence for the “late” integration of common ground.

*Figure 1.1.* Illustration of a critical display from the addressee’s and the speaker’s (director’s) perspectives, used in Keysar et al. (2000). Source: Keysar et al. (2000).
More recent evidence for the “late” integration view comes from studies providing a more direct demonstration of the idea that taking another person’s perspective is resource intensive. For example, following Keysar et al. (2003), Lin, Keysar, and Epley (2010) asked their participants to follow instructions provided by the speaker to manipulate objects, such as “pick up the tape” in contexts where the visual display included a cassette tape that was visible to both interlocutors as well as a roll of adhesive tape which was hidden from the speaker. During the interpretation of “tape”, participants were more likely to visually consider the adhesive tape that was in privileged ground than in a corresponding condition where the adhesive tape was replaced with a control object (e.g., a battery). Furthermore, the authors found that participants who scored lower on a working memory span task, or who were under increased cognitive load, were even more likely to look to the object in privileged ground. These findings suggest that it should be difficult to routinely monitor and integrate common ground into initial language processing.

However, evidence from a number of other studies suggests that perspective differences are taken into account in the “early” moments of processing (Brown-Schmidt et al., 2008; Hanna et al., 2003; Hanna & Tanenhaus, 2004; Heller et al., 2008; Nadig & Sedivy, 2002). For example, in Hanna et al. (2003) Experiment 1, the participants’ task was to rearrange cardboard geometric shapes according to instructions produced by a live speaker who was hidden behind a divider. The participants were told that the (confederate) speaker could see the same shapes as them (common ground) with the exception of one shape that was hidden in an envelope and whose identity had been made apparent only to the participants (privileged information). On critical trials, participants followed instructions such as “put the blue triangle on the red one” when the display contained, for example, a blue triangle, two red triangles, a green circle, and a yellow square (see Figure 1.2. Panel A). In a corresponding condition, one of the red triangles was in the addressee’s privileged ground (e.g., the shape in the envelope, see Figure 1.2. Panel
B). Hanna et al. observed that, when both red triangles were in common ground, addressees showed equivalent visual consideration of the two red triangles upon hearing the word “red”. In contrast, when one of the red triangles was in common ground and the other was in the envelope (addressees’ privileged ground), addressees primarily looked at the red triangle that was in common ground within the first few hundred milliseconds after the onset of the adjective. However, the “privileged” red triangle was considered more than in control conditions where the envelope contained shapes unrelated to the target description, such as a yellow triangle (see Figure 1.2. Panel C). Hanna et al. took these findings as evidence that common ground information can immediately work to restrict the domain of interpretation for the unfolding description. However, the effect is not absolute and objects that match the referring expression can somewhat interfere with reference resolution even if they are not in common ground.

![Figure 1.2](image)

*Figure 1.2.* Illustration of a critical display from the addressee’s perspective in Hanna et al. (2003). Source: Hanna et al. (2003).

In another study, Heller et al. (2008) also used temporarily ambiguous referential expressions to examine whether addressees are sensitive to common ground during real-time reference resolution. Critical displays contained two sets of objects contrasting in size (e.g., big box, small box, big duck, small duck). One of the objects, however, (e.g., the small box) was covered from the speaker’s view (see Figure 1.3. Panel D). Both members of the other set (e.g., big duck and small duck) were, however, visible to the addressee and to the speaker. In this
situation, a “size contrast” was therefore apparent only for the objects that were mutually known. Heller et al. found that upon hearing the temporarily ambiguous contrastive adjective (e.g., “big”) in instructions such as “pick up the big duck”, addressees rapidly limited interpretation to the intended target object. In other words, the addressees recognized that in the absence of knowledge about the small box, the speaker should not use the contrastive adjective “big” to refer to the big box and would instead use a bare noun such as “the box”. This suggests that integration of speaker’s perspective happens comparatively “early” during language processing.

![Figure 1.3](image-url)

*Figure 1.3. Illustration of a critical display from the addressee’s perspective in Heller et al. (2008). Source: Heller et al. (2008).*

This “early” integration view was further supported by a number of studies exploring the interpretation of questions in unscripted dialogue (e.g., Brown-Schmidt et al., 2008; Brown-Schmidt 2009) and in narrative comprehension (e.g., Gibbs, Mueller, & Cox, 1988). A felicitous question should ask about information that is not known to the speaker but should be available to the addressee. In Brown-Schmidt et al., speakers asked addressees questions such as “*what’s above the cow that’s wearing shoes?*” in a situation where the addressee’s as well as speaker’s displays contained two cows (one with shoes and the other one with glasses). Critically, the animal above the cow with glasses (the competitor) was either visible on the addressee’s and
speaker’s displays (in common ground) or only on the addressee’s display (addressee’s privileged information). As the speaker’s uttered noun “cow” unfolded in time, addressees visually considered the competitor cow significantly more when the animal above it was addressee’s privileged information than when it was in common ground. Thus, addressees showed sensitivity to the fact that the speaker should ask about the information she/he does not know by directing visual attention towards entities that were in their privileged ground. This sensitivity appeared to be reliable during the processing of the temporarily ambiguous target noun (e.g., “cow”), even though in principle the disambiguating information was not encountered until the following prepositional phrase (e.g., “with shoes”). Thus, the growing body of evidence suggests that common ground has a potential to influence language processing as unfolding sounds are mapped onto potential referents.

What is the basis for the contradictory results found regarding the time course of common ground integration? One proposal involves the notion of referential fit. According to Hanna et al. (2003), the specific information chosen by a speaker to include in an utterance is usually one of the most reliable cues for identifying an intended referent. Consequently, even when an object is in privileged ground, if it “fits” a referring expression better than the intended referent in common ground, it is likely that this privileged object will attract the listener’s attention. Recall that, in the studies showing late common ground integration, Keysar et al. (2000) and Lin et al. (2010) used a methodology where a privileged object was always a “better fit” to the referring expression than the intended object in common ground in a strictly objective sense (e.g., the smallest candle in the display was the privileged object and a medium candle was the [shared] target referent on the trial where the target description was “the small candle”). Hence, it is not surprising that these authors observed comparatively late integration of common ground.
Another proposal relies on the idea of the interactivity of an experimental task. As suggested by a number of researchers, the use of pre-recorded materials (Brown-Schmidt, 2009), non-naive confederates (Kuhlen & Brennan, 2013), and highly structured tasks (Brennan & Hanna, 2009) may remove any potential for rich interaction, reduce the utility of taking the speaker’s perspective and, in turn, weaken the potential to detect common ground integration.

In summary, then, the above-mentioned studies on referential processing provide invaluable insight regarding the time course of common ground integration. However, they also leave several open questions such as whether all types of knowledge discrepancies affect language processing in the same way in different contexts (e.g., in conversation and narrative comprehension). These unanswered questions are addressed in the present dissertation.

1.4 Dissertation overview
As the previous studies illustrate, in a typical task addressing sensitivity to common ground in language comprehension, the status of knowledge as privileged versus shared is manipulated by varying the physical co-presence of the entities that are available for reference. For example, an object is usually visually occluded from the perspective of the speaker, thereby preventing him/her from knowing whether there is an object behind the occluder (and what it is). The object is, nevertheless, visible to the addressee. By incorporating the speaker’s perspective, the addressee should, in turn, reduce any expectation for that object to be part of the mutually-accessible referential domain. In the current dissertation, a different approach is adopted, focusing on knowledge discrepancies in terms of understanding of the identity of entities in a situation where their physical co-presence is already established.

To illustrate, consider the cases where listeners have to deal with references to objects that exhibit a discrepancy between what they look like and what they really are (e.g., a candle
that looks like a lightbulb). This situation allows at least two ways for a speaker to conceptualize and subsequently refer to the object. On the one hand, the object could be called “the lightbulb”, if the conceptualization of the object is based on its appearance. On the other hand, it could be referred to as “the candle”, if the conceptualization is based on what the object “truly” is, as determined by the function that the object can actually perform. Importantly, the potential to use the two terms also depends on the speaker’s knowledge state. If speakers do not know that the object is “truly” a candle because all that is available to them are visual features of the object, they should not use the noun “candle” to refer to it.

Another case where knowledge differences in terms of the identity of entities can have direct consequences for the design of an appropriate referring expression is the case of so-called referentially opaque contexts. For example, in some contexts, alternative terms such as “the man” and “the plumber” could refer to the same entity but, in other contexts, they cannot be substituted without creating a false statement (e.g., Davidson, 1984; Roberts, 1993; Ludwig & Ray, 1998; Ray, 1980). For example, the statement Susan believed that the man was polite cannot necessarily be substituted with Susan believed that the plumber was polite if Susan did not know that the man in question was a plumber. Thus, an important aspect to consider in referentially opaque contexts is the set of identity features one ascribes to the potential referents (e.g., in the example, whether Susan knows that the man is a plumber). As pointed out by Kamawar and Olson (1999), opaque contexts require an understanding that terms do not simply map on to the referent of a linguistic expression, but instead highlight the point that the choice of referring expression indicates how that object is to be represented or is represented in the mind of the speaker and/or the addressee.

I argue that it is important to consider these kinds of knowledge discrepancies in addition to the cases involving physical co-presence because the former might be more challenging for
comprehenders to manage (see Apperly & Butterfill, 2009)—a point discussed in greater detail in Chapter 3. Furthermore, as recognized by a number of researchers (e.g., Apperly, 2011; Apperly & Butterfill; 2009; Baron-Cohen, Tager-Flusberg, & Lombardo, 2013; Frisson & Wakefield, 2012; Low & Watts, 2013; Schober, 1998), in many natural contexts, knowledge discrepancies involve these kinds of perspective differences. Thus, considering these cases is an important step in ensuring that psycholinguistic studies exploring effects of common ground capture the range of knowledge discrepancies that are a part of daily communicative interactions.

One goal of this dissertation is to first establish whether and how listeners use knowledge about the true identity of visually misleading objects during real-time referential processing even though this knowledge contradicts the outward visual features of the objects. The next step is to then explore whether and when listeners are sensitive to another person’s lack of knowledge about the true identity of objects being referred to, in situations where the listeners are nonetheless aware of the objects’ identity. In addition, these cases will be directly compared to the more well-studied experimental situation where a knowledge discrepancy stems from a physical co-presence manipulation. Finally, the dissertation will investigate listeners’ sensitivity to knowledge about the identity of entities in narrative comprehension. As I will demonstrate, by exploring and contrasting cases that reflect qualitatively different types of knowledge discrepancies, it is possible to provide new insights into the timing and character of processing mechanisms, as well as the nature and richness of the knowledge representations used in real-time language comprehension.
Chapter 2: Appearance vs. Reality in Real-Time Referential Processing

In most cases, mapping a referring expression to its intended referent seems relatively straightforward. However, in some cases this process can involve challenges for the language processing system. One of such instances is where listeners have to deal with references to objects that reflect a mismatch in terms of what they look like and what they really are. Consider, for example, a donkey that looks like a zebra because someone painted white stripes on its back. This situation creates two ways to conceptualize and subsequently refer to the animal. On the one hand, the animal could be called “the zebra”, if the conceptualization of the animal is based on its appearance. On the other hand, it could be referred to as “the donkey”, if the conceptualization is based on the idea that changes in appearance do not change the underlying identity of the animal. In an offline categorization study, Keil (1989) asked his participants to read the stories where an animal (e.g., a donkey) is transformed into another (e.g., a zebra) by changing its visual features (e.g., by painting white stripes on its back). Keil found that when asked what kind of animal it really is, participants were more likely to classify it based on what it was before the transformation (e.g., a donkey) rather than based on what it looks like after the transformation (e.g., a zebra). This demonstrates the idea of psychological essentialism, which in this case refers to the idea that people believe that living entities possess an innate essence that cannot be changed by superficial alterations (e.g., Gelman & Wellman, 1991; Medin & Ortony, 1989; Sloman & Malt, 2003).

Frisson and Wakefield (2012) tested whether adults use psychological essentialist reasoning automatically in a task that does not require an explicit evaluation of a superficially altered entity’s category or a decision about which word fits better to describe it. They asked participants to read stories such as the following while monitoring their eye movements:
(3) a. A zookeeper took a lion and cut off its entire mane. He then used special paint to put stripes all over its body so that it looked like a tiger. He stood back and admired his work.

b. Soon after, the [lion \ tiger] was put in a cage.

c. A visitor walked past and said, “I think that [lion \ tiger] looks rather sick.”

Critically, in the “psychological essentialism” test sentence (e.g., “Soon after, the [lion \ tiger] was put in a cage.”) the animal was referred to either by a word that coincided with what the animal was before the transformation (e.g., “lion”) or what it was after the transformation (e.g., “tiger”). The participants read the word that coincided with the essentialist reasoning (e.g., “lion”) faster than the word that did not (e.g., “tiger”). Frisson and Wakefield took these results as evidence that readers experience processing difficulty when a referring expression clashes with the essentialist interpretation of the referent. This suggests that psychological essentialist reasoning about natural kinds is rather automatic and predominates in real-time referential processing in narrative comprehension.

In contrast to natural kinds, findings regarding psychological essentialist reasoning about artifacts have been less consistent. For example, Rips (1989) asked participants whether an object created to be a lampshade but which looked like an umbrella was more likely to be an umbrella or a lampshade. Most participants judged it was more likely to be a lampshade. This would be consistent with psychological essentialist reasoning because the object was categorized based on its function or essence rather than its appearance. In contrast, Keil (1989) found that when an artifact undergoes transformation, it is usually classified based on its new appearance rather than on what it was initially. Again, Frisson and Wakefield (2012) tested whether psychological essentialist reasoning applies to artifacts in real-time narrative comprehension. Consider one of their artifact transformation stories:
a. Susannah took a curtain and cut it down. She then sewed it up and added some buttons so that it looked like a jacket. She looked at her work and beamed with pride.

b. After that, she placed the [curtain \ jacket] on top of the bed.

c. Her friend walked into the bedroom and said, “I think that [curtain \ jacket] looks really colourful.”

Frisson and Wakefield found that, in the artifact transformation stories such as (5), participants read the word that coincided with the essentialist reasoning (e.g., “curtain”) slower than the target word that did not (e.g., “jacket”) in the “psychological essentialism” test sentence (e.g., “After that, she placed the [curtain \ jacket] on top of the bed.”). This finding in this case supports the idea that psychological essentialist reasoning does not apply in the same way for artifacts. However, in the above example, a change in the appearance of the artifact is confounded with the change of its function. For example, after curtain is transformed into a jacket, it can no longer serve its initial function of a curtain, while it is perfectly fit to be a jacket.

In contrast, a critical feature of Rips’s (1989) stories was that, although the artifact looked like one object (e.g., an umbrella), it still could perform the intended function of another object (e.g., a lampshade). Thus, it remains unclear whether psychological essentialist reasoning has any consequences on real-time referential processing when an artifact’s appearance is in conflict with its function (e.g., an object that looks like an umbrella but can be used only as a lampshade). In addition, it is unclear to what degree referential mapping might be affected when the experimental paradigm allows direct interaction with real visually misleading objects, rather than stories about objects as in Frisson and Wakefield (2012). This is important because direct interaction can remove any ambiguity in terms of what function(s) an object can/cannot perform after the transformation. On the other hand, it is possible that when the appearance of a transformed object is readily available, as in the case of visually-situated experimental
paradigms, salient visual information might dominate over memory-based knowledge about object’s function in real-time referential processing. These issues are addressed in Experiment 1.

2.1 Experiment 1
The goal of Experiment 1 was to evaluate whether knowledge about object “identity” that is in conflict with an object's outward features (e.g., knowing that the object that looks like a lightbulb is a candle) can produce the pattern of real-time competition that is associated with overlapping object names. As noted in Chapter 1, this lexical competition process (e.g., implicitly considering “candle” as the sounds in “candy” unfold in time) is regularly used to explore the lexical-conceptual representations that are active as language input unfolds in time.

In Experiment 1, eye movements were monitored as participants followed recorded instructions such as “Take the candy and move it one square down.” In one condition, the critical displays contained what I refer to as a regular phonological competitor for the target noun (candy) whose identity was apparent on the basis of its appearance (e.g., a regular candle, see Figure 2.1, Panel A for example). A second condition used what I refer to as a knowledge–based phonological competitor (e.g., a candle that looks like a lightbulb, see Panel B). In this case, the potential for the visually misleading object to attract fixations as the target noun (e.g., candy) is heard depends on the participant's knowledge. From its outward appearance alone, the visually misleading object should be identified as a lightbulb, thereby eliminating the overlap in sounds with the target name, in turn making fixations to this object unlikely. Only when the listener has foreknowledge about the visually misleading object, as when it has been explicitly described as a candle and/or the visual features that make it a candle (e.g., a wick) are revealed, should this aspect of the object's identity be clear to the participant. In a third condition, the competitor was a phonologically unrelated visual control object whose actual identity was the same as the "apparent" nature of the corresponding knowledge-based phonological competitor (e.g., a
genuine lightbulb, Panel C). In all conditions, the identity of objects was revealed/described to participants as objects were placed on the shelf display.

The core experimental question concerns the pattern of results observed in the condition with the knowledge-based phonological competitor. Specifically, can knowledge about the object’s “true” identity (that is in conflict with an object's outward features) produce characteristic patterns of real-time competition as speech sounds are mapped to potential referents? If so, participants should look to the knowledge-based phonological competitor (e.g., candle that looks like a light bulb) and the regular phonological competitor (e.g., a regular candle) to the same extent while hearing the target noun (e.g., candy). Furthermore, comparatively few fixations should be made to competitors in the condition where they are simply the visual control objects. This pattern would be consistent with a view in which the “essence” of objects outweighs their appearance, even in an experimental paradigm where visual appearance is particularly salient. Alternatively, it is possible that visual features predominate in early stages of referential processing while conceptual knowledge about the object is integrated “later”. If this is the case, the early moments of processing should reflect a pattern where participants should look to the knowledge-based phonological competitor (e.g., candle that looks like a lightbulb) less than to the regular phonological competitor (e.g., a regular candle) and to the same extent as to visual control object (e.g., a regular lightbulb) while hearing the target noun (e.g., candy). Such an outcome would be consistent with a view that an object’s appearance outweighs its “essence” in real-time referential processing within the visual world eye-tracking paradigm. A third possibility is that knowledge about the object’s “true” identity and its outward features both serve as contextual cues when speech sounds are mapped to potential referents. If so, participants would look to the knowledge-based phonological competitor (e.g., a candle that looks like a lightbulb) less than to the regular phonological competitor (e.g., a regular candle)
and more than to visual control object (e.g., a regular lightbulb) while hearing the target noun (e.g., candy). This outcome would support the notion that visual cues and knowledge-based information about object identity have roughly equal influence as language is mapped to referential entities.

2.1.1 Method

2.1.1.1 Participants

The participants were 24 undergraduate students recruited at the University of Toronto Mississauga. All participants were either native speakers of English or started to learn English in North America from the age of 7 or earlier. An additional three participants were excluded, either because they strategically maintained their gaze towards the center of the display throughout the entire experiment \( (n = 1) \) or because of equipment problems \( (n = 2) \). Participants were paid $5 per half hour or received course credit for their participation. Participants were tested individually in one session. Each session lasted approximately half an hour.

2.1.1.2 Materials and design

*Visual stimuli.* There were 9 critical displays (see Table 2.1). Each critical display contained three items: the target object, the competitor object, and the unrelated object (see Figure 2.2 for an example display). The type of competitor object was manipulated across conditions. In the regular phonological competitor condition, the competitor was an object whose name shared onset sounds with the target name (e.g., a candle for the target candy; a yo-yo for the target yogurt). In the knowledge-based phonological competitor condition, a visually misleading object was used as the competitor, and the potential for phonological overlap with the target name depended on knowing the actual identity of the competitor object (e.g., a candle that looks like a lightbulb, a yo-yo that looks like a baseball). Importantly, the name corresponding to the outward
appearance of the object did not share onset sounds with the target noun (i.e., *lightbulb*, for the target noun *candy, baseball*, for the target noun *yogurt*). The knowledge-based phonological competitors were acquired from various novelty stores and thus represent objects that people may encounter in real life rather than being constructed solely for the purposes of this experiment. Finally, in the visual control condition, the competitor was an actual instance of the item that the knowledge-based competitor resembled (e.g., a typical lightbulb, a typical baseball). For a given object array, the target object and the unrelated object remained unchanged across conditions. To facilitate the coding of eye movements, target and competitor objects were always placed either in the top-left, bottom-left, top-right or bottom-right compartment of the open shelf and were located on the opposite sides of the display. The unrelated objects were placed in one of the four middle open shelf compartments such that they were not directly beside either the target or the competitor. The targets and competitors were approximately 85 cm apart.

![Figure 2.1](image)

*Figure 2.1.* A regular candle (Panel A), a candle that looks like a lightbulb (Panel B), and a regular lightbulb (Panel C), illustrating the type of competitor manipulation common to Experiments 1, 2, 3, and 4.
Table 2.1. List of target objects and the three types of competitors used in critical trials in Experiments 1, 3 and 4.

<table>
<thead>
<tr>
<th>Target</th>
<th>Regular phonological competitor</th>
<th>Knowledge-based phonological competitor (looks like)</th>
<th>Visual control object</th>
</tr>
</thead>
<tbody>
<tr>
<td>Popcorn</td>
<td>Potholder</td>
<td>Potholder (record)</td>
<td>Record</td>
</tr>
<tr>
<td>Crane</td>
<td>Crayons</td>
<td>Crayons (Lego)</td>
<td>Lego</td>
</tr>
<tr>
<td>Paper bag</td>
<td>Paperclips</td>
<td>Paperclips (can of sardines)</td>
<td>Can of sardines</td>
</tr>
<tr>
<td>Gnome</td>
<td>Note block</td>
<td>Note block (Rubik’s cube)</td>
<td>Rubik’s cube</td>
</tr>
<tr>
<td>Scrub brush</td>
<td>Screwdriver</td>
<td>Screwdriver (pen)</td>
<td>Pen</td>
</tr>
<tr>
<td>Candy</td>
<td>Candle</td>
<td>Candle (light bulb)</td>
<td>Light bulb</td>
</tr>
<tr>
<td>Yogurt</td>
<td>Yo-yo</td>
<td>Yo-yo (baseball)</td>
<td>Baseball</td>
</tr>
<tr>
<td>Tiger</td>
<td>Timer</td>
<td>Timer (pepper grinder)</td>
<td>Pepper grinder</td>
</tr>
<tr>
<td>Pennies</td>
<td>Pencils</td>
<td>Pencils (drumsticks)</td>
<td>Drumsticks</td>
</tr>
</tbody>
</table>

Critical instruction given the example display:

"Take the candy and move it one square down"

Figure 2.2. Illustration of a critical display with target (candy), competitor (*candle that looks like a light bulb), and unrelated object (tennis ball) in Experiment 1.
In addition to the nine critical trials, there were two practice trials and 15 filler trials. Thirteen fillers had displays that contained four objects and two fillers had displays that contained three objects. Finally, five fillers had displays containing visually misleading objects (e.g., a laser pointer that resembled an arrowhead) and three fillers had displays containing hard-to-name objects (e.g., several constructor pieces attached to each other). The location of the items referred to in the instructions was varied in the filler trials so as to break any expectations based on the patterns in the critical trials. On 12 of the trials (6 critical and 6 fillers), the first-to-move objects occurred on the left side of the display and on the other 12 (3 critical and 9 fillers) on the right side of the display. The pairing of object arrays to conditions was cycled across participants through the use of three stimulus lists. Each participant saw one of the three versions of each critical display.

**Auditory stimuli.** Critical instructions were of the form “Take the [target noun] and move it [location]” (for an example, see Figure 2.2). The instructions were pre-recorded by a female speaker using normal descending intonation. They never referred to the competitor objects. The same sound files were used for all participants. On average, the onset of the critical word occurred 1028 ms ($SD = 99$) after the start of the sound file. The average duration of the target word was 728 ms ($SD = 183$). The disambiguation point within each word (i.e., relative to the competitor name, e.g., where the final vowel in “candy” distinguishes it from the word “candle”) occurred at an average latency of 326 ms ($SD = 119$) from word onset.

On critical trials, the first instruction referred to the target object, and the second referred to the unrelated object. In five filler trials, the instructions required participants to move two objects. In the remaining 10 fillers, they required participants to move three objects. The instructions never referred to the visually misleading objects that appeared in the filler trials. For all participants, objects were named while putting them on a display (e.g., *In this trial we have a*
candle [in conditions involving regular and knowledge-based phonological competitors], a lightbulb [in conditions containing the visual control object], a tennis ball and a candy). In addition to naming, the feature of any visually misleading object that would give away its identity was also demonstrated (e.g., the wick of a candle that looks like a lightbulb, the string of a yo-yo that looks like a baseball). Thus, the true identity of the knowledge-based phonological competitors was always introduced explicitly. Three of the nine visually misleading objects had visual features that could give away their identity to the listener (e.g., the wick of the candle that looks like a lightbulb, the tip of the screwdriver that looks like a pen, and the tips of the pencils that look like drumsticks) and potentially could be visible to the listener based on their orientation on the shelf display. The other six visually misleading objects’ visual features that could give away their identity were completely concealed by the orientation of the objects on the shelf display (e.g., a string of the yo-yo that looks like a baseball) or the objects did not have such features (e.g., crayons that look like Lego blocks).

2.1.1.3 Procedure

Each participant filled out a language questionnaire prior to the experiment. The participant was then seated in front of and approximately 60 cm away from the tabletop shelf display. For each trial, the experimenter named and manually placed the objects on the display, and then initiated the audio instructions (played via computer and using speakers located to either side of the participant). Two practice trials were presented at the beginning of the experiment to familiarize participants with the procedure. The participant’s face was recorded using a miniature SONY 1/3” CCD Camera attached to the center of the display, focused on the participant's eye region. Video output from the camera was recorded directly onto the hard disk of a desktop computer using Final Cut Express video editing software.
2.1.2 Results and Discussion

2.1.2.1 Coding

Eye movements were analyzed from the video recordings using playback with synchronized audio and video channels. Eye movements were coded in 33 ms intervals (1 video frame) within an interval beginning 200 ms after and ending 700 ms after the onset of the target noun (an interval corresponding to the first two-thirds of the target word, on average). Eye position was coded as reflecting a fixation to one of the four corner positions (where the objects of interest were located on critical trials) or "other".

Of primary interest is the extent to which the competitor object is fixated as the temporarily ambiguous target name (e.g., candy) unfolds in time. In view of this, the data analyses are based on a likelihood measure calculated as the proportion of saccades to the competitor out of all trials within the specified time interval\(^1\). This measure includes saccades made within a 500 ms interval (200-700 ms after target name onset) corresponding to the first two-thirds of the target word, on average. This measure was chosen over fixation measures because it more cleanly reflects the looking behaviour triggered by the relevant language stimulus rather than other factors such as visual salience of the objects. An additional 0.5% of the trials were discarded due to poor recording quality. Trials where participants made a gaze shift to the competitor were coded as 1 and trials where participants did not make a gaze shift to the competitor were coded as 0.

2.1.2.2 Statistical modeling.

The competitor consideration measure was examined in a logistic mixed effects model (Jaeger,

\(^1\) When the likelihood measure was calculated as the proportion of saccades to the competitor using only those trials on which a saccade was made to either the competitor or the target within the specified time interval (a measure that reflects the relative consideration of the target and competitor but which involves only a subset of all critical trials), an identical pattern of results was found.
2008) with participants and items as crossed, independent, random effects, implemented in package lme4 of the statistical software R 2.15.2 (Bates, Maechler, & Bolker 2012; R Core Team, 2012). The model that properly converged included random intercept for participants as well as random intercept and competitor type slope for items.

Main effects and interactions were evaluated by performing likelihood ratio tests, in which the deviance (-2LL) of a model containing the fixed effect is compared to another model without the effect in question but is otherwise identical in random effects structure. If removing the fixed effect caused the comparison model to fail to converge, the model was simplified following the “best path” procedures (Barr, Levy, Scheepers, & Tily, 2013). For the fixed effects, I report \( \chi^2 \) as well as \( p \) values associated with the corresponding likelihood ratio test. For the follow-up pairwise comparisons, I report the \( \beta \), \( SE \), \( z \), and \( p \) values associated with a particular contrast.

2.1.2.3 Eye movement data.

Figure 2.3 plots likelihood of a saccade to the competitor\(^2\) as a function of competitor type. The proportion of fixations to the competitor and target as a function of time and competitor type are shown in Appendix A. The analysis revealed that there was a main effect of competitor type\(^3\), \( \chi^2 (2) = 13.02, p = .001 \). Follow up pairwise comparisons revealed that competition scores for the regular phonological competitor (likelihood = 0.25) and the knowledge-based phonological competitor (likelihood = 0.21) were higher than for the visual control object (likelihood = 0.04), \( \beta = 2.31, SE = 0.92, z = 2.52, p = 0.012 \) and \( \beta = 2.04, SE = 0.93, z = 2.20, p = 0.028 \), respectively. However, the likelihood of a saccade to the competitor in conditions with the regular phonological competitor (0.25) and the knowledge-based phonological competitor (0.21) respectively.

---

\(^2\) Although the analyses were conducted in logit space, raw likelihood values are reported for easier interpretation.

\(^3\) When the three visually misleading objects whose visual features giving away their true identity could be recognized by the listener were excluded from the analysis, the pattern of descriptive and inferential statistics remained the same.
was not significantly different across these conditions, $p = 0.528$.

The results regarding the regular phonological competitors are consistent with previous findings by Allopenna et al. (1998), showing greater visual consideration of phonological competitors (e.g., candle) compared to unrelated objects (e.g., tennis ball) as the target noun (e.g., candy) unfolds. Critically, the results showed a similar pattern for knowledge-based phonological competitors and for regular phonological competitors: When participants knew the true identity of the knowledge-based phonological competitors, these competitors functioned just as regular phonological competitors, despite their visually misleading appearance. These findings are consistent with a psychological essentialist reasoning account, where the “essence” of objects easily outweighs their appearance. In some ways the strength of this effect is surprising because in contrast to the knowledge about the “true” identity of the objects, the outward features that contradicted this knowledge were particularly salient in this paradigm.

Figure 2.3. Likelihood of a saccade to the competitor as a function of competitor type, Experiment 1. Note: Although the analyses were conducted in logit space, raw likelihood values are reported for easier interpretation.
2.2 Summary
In this chapter I asked how knowledge about object identity that is in conflict with an object's outward features (e.g., knowing that the object that looks like a lightbulb is a candle) is used in incremental referential processing. Specifically, to what extent is a description for an object that is based on its essence or function available as the noun is unfolding in real-time, when the object’s visual features are at odds with this information? The findings revealed that visual consideration of objects whose outward features were in conflict with their “true” identity (e.g., candle that looks like a lightbulb) was similar to objects whose outward features were consistent with their true identity (e.g., a regular candle) when a phonologically overlapping target name (e.g., candy) was heard. For both types of objects, visual consideration was higher compared to the visual control objects (e.g., a regular lightbulb). This suggests that when participants know the true identity of a visually misleading object, the label that corresponds with this identity is activated to a similar extent as when the object has visual features that match its identity. This outcome is important in that the salience of objects’ visual information in the experimental methodology used here could arguably boost the status of “appearance” information, in turn damping the effect of knowledge about objects’ essence.

These findings are in some ways inconsistent with the results obtained by Frisson and Wakefield (2012). As noted earlier, the authors found that while reading “transformation” narratives, participants took longer to process a sentence referring to an artifact that used an untransformed term (e.g., “curtain”) than when they used a transformed term (e.g., “jacket”). Frisson and Wakefield argued that psychological essentialist reasoning does not apply to transformed artifacts as it does to transformed natural kinds. However, the findings from Experiment 1 (in which all the stimuli were artifacts) are incompatible with this view. One explanation for the diverging results is that when the task involves a real and co-present visually
misleading object and when there is no ambiguity in terms of what function(s) an object can/cannot perform, the label that corresponds with an object’s function is more accessible in real-time referential processing. But regardless of whether this is the correct explanation for the contrasting results, the current findings are, in fact, consistent with earlier studies using off-line categorization tasks, showing that, when the function of an artifact is changed due to a transformation, the artifact is still categorized based on its function rather than based on its appearance (e.g., Rips, 1989).

In this chapter it was established that when participants know the true identity of a visually misleading object, the label that corresponds with this identity is activated to a similar extent as when the object has visual features that match its identity. This effect provides a crucial element of the logic behind the next two experiments which used a paradigm that “scales up” the communicative situation to involve conversational interaction and perspective differences across conversational partners. Critically, when listening to spoken instructions produced by another individual, the potential for the knowledge-based competitor’s name to overlap with the target’s name should depend on the way in which knowledge is attributed to the speaker, and not just a listener's own knowledge. Consider, for example, the case of a candle that actually looks like a lightbulb. In the absence of foreknowledge about the object's true identity, the speaker could not plausibly refer to this object using a term such as *candle*. Thus, an addressee who is efficiently considering the knowledge state of the speaker should be unlikely to consider this object during real-time comprehension as the speaker utters the target word (*candy*), even when the addressee possesses this knowledge. Such an outcome would require the addressee to suppress or bypass her or his privileged knowledge about the “true” identity of an object. The following chapters explore this possibility and also provide a comparison with the
more well-studied case involving knowledge differences regarding the existence or presence of objects in the situational context.
Chapter 3: Privileged Knowledge about Object Existence vs. Object Identity in Conversation

Numerous psycholinguistic studies have shown that speakers take shared knowledge into account when constructing utterances, and that addressees use shared knowledge to contour their understanding of utterances (e.g., Clark et al., 1983; Fussell & Krauss, 1989; Gerrig & Littman, 1990; Gibbs et al., 1988). For a given conversational participant, determining what information is likely shared and what is privileged is itself a complex task that draws on multiple sources of information such as community membership as well as the situational and linguistic context (Clark & Marshall, 1981) and, accordingly, entails various cognitive demands.

As noted earlier, one question in contemporary work on real-time referential processing involves the extent to which addressees are able to efficiently coordinate shared and privileged knowledge in the earliest moments of comprehension. As outlined in Chapter 1, some studies have suggested that the integration of common ground occurs relatively late in the course of processing, and that an addressee’s egocentric perspective predominates in the earliest moments (e.g., Barr, 2008; Keysar et al., 1998; Keysar et al., 2000; Keysar et al., 2003; Kronmuller & Barr, 2007; Pickering & Garrod, 2004). However, other studies of referential interpretation have challenged the egocentric-first view by showing that, although addressees do not inhibit their egocentric perspective completely, shared knowledge does have a measurable influence on referential mappings from the earliest moments of interpretation in a variety of circumstances involving a number of different types of referring expressions (e.g., Brown-Schmidt, 2012; Brown-Schmidt et al., 2008; Hanna et al., 2003; Hanna & Tanenhaus, 2004; Heller et al., 2008).

The review in Chapter 1 also highlighted the fact that, in most experimental studies addressing the time course of common ground integration in conversation, the status of knowledge as privileged versus shared is manipulated by varying the physical co-presence of the
entities that are available for reference. For example, an object may be visually occluded from the perspective of one conversational partner, thereby preventing that partner from knowing whether there is an object behind the occluder (and what it is). By incorporating this individual’s perspective, the other partner should, in turn, reduce any expectation for that object to be referred to. Although this approach has provided an important methodological template for numerous studies, including work on cognitive development (e.g., Nadig & Sedivy, 2002) as well as work on individual and cultural differences in perspective taking (e.g., Brown-Schmidt, 2009; Wu & Keysar, 2007), it provides only one means of exploring the coordination of shared versus privileged knowledge. In fact, in many natural contexts, knowledge discrepancies do not involve different perspectives regarding the existence of referents. Consider, for example, Donnellan's (1966) illustration in which a person who poses the question “Who is the man drinking a martini?” is unaware of the fact that the referent's glass contains only water—a fact that may, however, be known to the addressee. In this case, there is a knowledge mismatch regarding the speaker's and the addressee's understanding of the identity of the liquid in the glass, although the physical presence of the liquid in the context and the liquid’s availability for shared reference is clearly established for both. Similar kinds of knowledge differences involving object identity could arise with the visually misleading objects used in Experiments 1 (e.g., a candle that looks like a lightbulb). For example, one interlocutor might know both the appearance and true identity of a visually misleading object, while another might know only the appearance. These kinds of knowledge mismatches across conversational partners, which I refer to as discrepancies involving object “identity”, are the focus of this chapter.

Knowledge discrepancies involving object identity are important in view of their link with issues in formal semantics such as the distinction between speaker reference (i.e., situation-specific meaning) versus truth-theoretic (i.e., context-invariant) meaning, as well as phenomena
such as referential opacity (discussed in greater detail in Chapter 4, e.g., Quine, 1960; Roberts, 1993). In addition, in the theory of mind literature, a number of studies of cognitive development have suggested that perspective differences involving object identity are very challenging for children to overcome compared to cases where perspective differences involve keeping track of which objects are occluded from another's view. For example, at around 14 months of age, children show evidence of understanding that an adult should not see an object if it is blocked by an opaque screen (Sodian, Thoermer, & Metz, 2007), or when individuals are wearing a blindfold or having their eyes closed (Brooks & Meltzoff, 2002; Caron, Butler, & Brooks, 2002), even though the object is visible to the child. In addition, Moll and Tomasello (2006) showed that 2-year-olds understand and adjust communicative behaviors accordingly when another person cannot see an object that is occluded from her/his view, even though the 2-year olds themselves could see that object.

In contrast, the representational skills necessary to deal with cases involving discrepancies in object identity appear to emerge later, only between 3 and 4 years of age (e.g., Flavell, 1988; Sapp, Lee, & Muir, 2000). In a typical task used to assess this ability, children are first shown a visually misleading object (e.g., a pencil sharpener that looks like a ball) and are asked what it looks like. Then, the visually misleading object’s true identity is revealed. Finally, children are asked questions about the visually misleading object such as What does it look like? Does it look like a ball or does it look like a pencil sharpener? (appearance question) and What is it really really? Is it really really a ball or is it really really a pencil sharpener? (true identity question). Flavell showed that in such tasks 3-year-olds typically make either a "phenomenism" error (e.g., respond that the object looks like and really is a ball) or an "intellectual realism" error (e.g., respond that it both looks like and really is a pencil sharpener). Only at 4 years of age do children begin to respond correctly by noting that although the object looks like a ball, it is really
a pencil sharpener. Finally, in other work, Apperly and Robinson (1998) showed that only at around age 5 could children correctly gauge when another person would think that a visually misleading object (e.g., a pencil sharpener that looks like a ball) is what it looks like (e.g., a ball) rather than what it actually is (e.g., a pencil sharpener). In sum, then, developmental work suggests that the ability to coordinate knowledge differences related to the physical co-presence of objects develops well in advance of the ability to manage the distinction between appearance and true identity, perhaps because these two aspects of theory of mind may rely on different cognitive mechanisms (e.g., Apperly & Butterfill, 2009). Of particular interest here is whether the different patterns associated with the latter type of knowledge discrepancy in theory of mind development are evident in real-time referential processing in adults.

To date, experimental attempts to tap into knowledge discrepancies about object identity have used a type of explicit misinformation paradigm for varying the understood identity of hidden objects (Hanna et al., 2003; Keysar et al., 2003). For example, in Hanna et al., Experiment 2, although two jars visible to the addressee but not to the speaker were empty, the experimenter misinformed the speaker that one of them was filled with olives. The task then required the speaker to describe the objects she or he believed the addressee was viewing from left to right, just before the speaker provided an instruction pertaining to one of the objects. This procedure was meant to reinforce the addressee’s awareness of the speaker’s false belief about the characteristics of the critical objects. In this context, the authors found that the addressees were able to overcome their private knowledge about the actual properties of display objects. Thus upon hearing empty in take the empty jar..., addressees more strongly considered the jar that the speaker falsely believed was the (only) empty jar (as compared to a condition where the speaker was correctly informed that both jars were empty). Nevertheless, in a situation where the second jar did actually contain olives, addressees were overall fastest at identifying the target
referent. The authors take these results as evidence that, although addressees do not completely inhibit their egocentric perspective, they are able to accommodate to the speaker’s false belief about an object’s identity during incremental processing.

In Keysar et al. (2003), the experimenter placed a competitor object (e.g., a roll of adhesive tape) in an opaque paper bag, covered it from the speaker’s view with an occluder panel, and also mislabeled it to the speaker both verbally and with a picture (e.g., in the presence of the addressee attached a picture depicting a small leather ball on the part of the occluder that was visible to the speaker). The target was a mutually visible object (e.g., a cassette tape) whose name overlapped with the name of the actual competitor item from the listener’s perspective. Consistent with Hanna et al., the authors observed that addressees considered the competitor object (e.g., an adhesive tape) when hearing the tape more than in a control condition where the competitor in the bag was replaced with an unrelated object (e.g., a toy truck). However, unlike Hanna et al., Keysar et al.’s design did not include a condition where the speaker’s belief would match the actual identity of a competitor object. Thus, it remains an open question whether a competitor in the Keysar et al. study (e.g., adhesive tape) would actually be considered even more when the speaker’s belief is known by the listener to match its actual identity than when the speaker has a false belief about its identity. Finally, and perhaps most important, the above-mentioned studies do not allow a direct comparison as to whether the ability to manage identity discrepancies is different from the ability to manage existence discrepancies during incremental referential interpretation, and what the precise profile of differences might be.

The present chapter address three questions: First, are addressees effective at managing shared versus privileged knowledge about object identity as initial speech sounds of referential expressions are being processed? Second, are they more likely to consider privileged information in this situation compared to the well-studied situations where privileged knowledge involves
objects that are not known to be physically present to the speaker? Finally, are addressees less likely to consider privileged knowledge about object identity when the language used in prior discourse explicitly signals a speaker’s perspective regarding a visually misleading object’s identity? Instead of the kind of misinformation paradigm used in previous studies examining identity discrepancies (Hanna et al., 2003; Keysar et al., 2003), the current study builds on the methodology used in Experiment 1 (involving visually misleading objects) in order to create a situation where addressees could infer how the incorrect assessment of an object’s identity would arise as a result of the objects’ intrinsic visual features, making a misinformation intervention by the experimenter unnecessary. Furthermore, the knowledge discrepancies in the current study involved the entire semantic category to which an object belonged, and not just a misattributed feature of an object. This brings the current study closer in line with work on theory of mind as well as the kinds of cases of interest in formal semantics.

To address these questions, three experiments were conducted. Experiment 2 was a validation study conducted to assess how the visually misleading objects used in the subsequent experiments would be referred to without foreknowledge about their actual identity, when viewed from the speaker’s perspective. To foreshadow the results, the experiment confirmed that in the absence of foreknowledge about the objects’ true identity, participants did not refer to these objects using a term reflecting their true identity (e.g., “the candle” for a candle that looks like a lightbulb) and instead selected a term reflecting the objects’ appearance. In Experiments 3 and 4, eye movements were recorded as participants (henceforth addressees) participated in a referential communication task with a confederate speaker (henceforth the speaker). In this task, the ability to manage privileged and shared information was measured in the same way as Experiment 1, namely by examining addressees' consideration of a competitor object whose name shared initial speech sounds with the name of a target referent (e.g., candle, given the
target word *candy*, see Barr, 2008, for a similar approach). As was seen in Experiment 1 where prerecorded instructions were used, competitor objects should attract fixations as the initial speech sounds in the target word unfold in time. In the current experiments, however, the potential for the knowledge-based competitor’s name to overlap with the target depends on the way in which knowledge is attributed to a conversational partner, and not just the participant’s own knowledge. Consider, for example, the case of a candle that actually looks like a lightbulb. In the absence of foreknowledge about the object’s true identity, the speaker could not plausibly refer to this object using a term such as *candle*. Thus, an addressee who is efficiently considering the knowledge state of the speaker should be unlikely to consider this object during real-time comprehension as the speaker utters the target word (*candy*)—an outcome that requires the addressee to suppress or bypass her or his privileged knowledge about object identity. The core question is whether this pattern is in fact observed (suggesting identity discrepancies are managed effectively during incremental interpretation, as is often shown for discrepancies involving object existence) or whether addressees exhibit difficulty in managing their privileged knowledge about object identity.

Experiment 3 directly compares addressees’ abilities to manage knowledge discrepancies involving the identity of objects with the situation in which discrepancies involve objects’ existence. Experiment 4 is designed to investigate whether prior descriptions on the part of the speaker that reinforce her lack of knowledge about an object’s true identity can improve addressees’ ability to coordinate shared vs. privileged knowledge about object identity in the earliest moments of language processing.

### 3.1 Experiment 2

The goal of Experiment 2 was simply to confirm that, when the true identity of visually
misleading objects is not known, they will be identified and named in terms of their outward appearance. This is an important methodological step to ensure that the addressees (participants) in subsequent experiments should not expect that the speaker might be able to infer the “true” identity of the visually misleading objects in situations where this identity had not been explicitly stated for the speaker. If it turned out that the “true” identity of the visually misleading objects could be recognized from their outward features, it would be impossible to create a situation where objects’ identity can be safely understood to belong to addressees’ privileged knowledge.

The visually misleading objects were the same as in Experiment 1. As in Experiment 1, each visually misleading object (e.g., a candle that looks like a lightbulb) was varied with two other objects, creating a triad of three substitutable objects whose identity could be judged when presented separately to participants. One of the other objects looked like and actually was a member of the category that the visually misleading object resembled (e.g., a genuine lightbulb). The final one belonged to the same "actual" category as the visually misleading object, but was not visually misleading (e.g., a candle that in fact looked like a prototypical candle). On each trial, participants were asked to simply name a solitary object that was placed on a shelf compartment in front of them. As noted above, the goal was to ascertain that the label that participants provided for visually misleading objects (e.g., a candle that looked like a lightbulb) should reflect the objects' perceptual nature (e.g., "lightbulb"), and as such the pattern of results should be the same as when participants were shown a non-Visually misleading object with the same outward appearance (a genuine lightbulb). In contrast, when the object was a non-Visually misleading object belonging to the actual category as the visually misleading object (a candle), a different category label should be provided.
3.1.1 Method

3.1.1.1 Participants

Participants were 30 undergraduate students recruited from the University of Toronto Mississauga community. All participants were either native speakers of English or started to learn English in North America from the age of 7 or earlier. An additional two participants were excluded because of experimenter error during the presentation of stimuli. Participants were paid $5 per half hour or received course credit for their participation. Participants were tested individually in one session. Each session lasted approximately 30 minutes.

3.1.1.2 Materials and design

On each trial, an object was placed alone in a compartment of the shelf display also used for Experiment 1, and the participant was asked to name the object. There were nine critical trials, and the nature of the object that was presented was manipulated across three conditions. In the visually-misleading object condition, the object was the same as in the knowledge-based phonological competitor condition in Experiment 1 (e.g., a candle that looked like a lightbulb, a yo-yo that looked like a baseball). In the identity-matched object condition, the object was the same as in the regular phonological competitor condition in Experiment 1 (e.g., a prototypical candle or a prototypical yo-yo). In the appearance-matched object condition, the object was the same as in the visual control object condition in Experiment 1 (e.g., a genuine lightbulb or a genuine baseball). The pairing of objects to conditions was cycled across participants through the use of three presentation lists. Each participant saw only one of the three versions of each critical object, but across participants, each critical object occurred in all experimental conditions.

In addition to the nine critical trials, there were 15 filler trials. Ten fillers involved objects that were difficult to name (e.g., a component of a bicycle gear shifter) and five fillers involved comparatively easy-to-name objects (e.g., a glass). The goal of the filler trials was to make the
task more challenging and to distract participants from the true purpose of the experiment.

3.1.1.3 Procedure
The participant was seated approximately 60 cm away from the tabletop shelf display. For each trial, the experimenter manually placed an object in a shelf compartment and gave a signal to the participant to provide his or her response. As mentioned in Chapter 2, three of the nine visually misleading objects had visual features that could give away their true identity, if placed on a certain orientation (e.g., the wick of the candle that looks like a lightbulb, the tip of the screwdriver that looks like a pen, and the tips of the pencils that look like drumsticks). These three objects were placed in the orientation they would be seen by the speaker in subsequent experiments. Namely, the candle that looks like a lightbulb, the screwdriver that looks like a pen, and the pencils that look like drumsticks were placed so that the visual features that could give away their true identity were not visible to the participants. The participant’s responses were recorded by the experimenter.

3.1.1.4 Coding
The names provided for the critical objects by the participants were coded based on whether they reflected the outward appearance or actual identity of the objects. For example, given a candle that looked like a lightbulb, responses such as “the lightbulb”, “the toy lightbulb” or “the weird lightbulb” would be classified as reflecting the object’s outward appearance. Responses such as “the candle” or “the lightbulb candle” would be classified as reflecting the object’s actual identity.

3.1.2 Results and Discussion
As expected, in the identity-matched and appearance-matched conditions, the object names reflected a judgement of object identity based on the perceptual features of objects (100% for
each condition). Although objects' outward appearance and identity overlapped in these conditions, this result confirmed that the objects in these conditions were indeed easy to recognize. More important, participants overwhelmingly named the visually misleading objects based on their appearance (98%) rather than their identity (2%). For example, the yo-yo that looked like a baseball or the candle that looked like a lightbulb were called “the baseball” and “the lightbulb” rather than “the yo-yo” and “the candle”. This confirmed that, when the true identity of the visually misleading objects was not made apparent, this identity could not be discerned from their outward appearance.

3.2 Experiment 3
Experiment 3 involved a referential communication task in which participants took the role of addressees who followed spoken instructions such as Take the candy and move it one square down…. The critical part of the instructions was the same as in Experiment 1. However, in the current experiment, the instructions were provided by a confederate speaker in a different room, communicating via audiovisual conferencing software. Addressees’ eye movements were recorded as they followed the instructions provided by the speaker. As described earlier, the goal was to directly compare addressees' abilities to manage knowledge discrepancies involving the identity of objects with the situation in which discrepancies involve objects’ existence.

3.2.1 Method
3.2.1.1 Participants
The participants were 72 undergraduate students recruited from the University of Toronto Mississauga campus community. All participants were either native speakers of English or started to learn English in North America from the age of 7 or earlier. An additional three participants were excluded, either for strategically maintaining their gaze toward the center of the
display throughout the entire experiment (n = 1), or because they reported suspecting that the speaker was not another naïve participant during post-experiment debriefing (n = 2). Participants were paid $10 per hour or received course credit for their participation. Participants were tested individually in a single session that lasted approximately one hour.

3.2.1.2 Materials and design

Eight of nine of the critical displays were the same as in Experiment 1 (see Table 2.1). For one critical display, the candy target and the regular candle competitor were substituted with more prototypical examples of these categories. As in Experiment 1, the current experiment included a within-participant manipulation that varied the type of competitor object used in a given critical trial. Specifically, this object could be a regular phonological competitor (e.g., a regular candle) for the target name (e.g., candy), a knowledge-based phonological competitor (e.g., a candle that looks like a lightbulb), or a phonologically unrelated visual control object (e.g., a genuine lightbulb).

In addition to the competitor type manipulation, a second, between-participants manipulation varied the knowledge state that the addressee would attribute to the speaker. The knowledge state attributed to the speaker was manipulated across three conditions (see Table 3.1. for a summary). As noted earlier, the speaker and the addressee were located in different rooms and communicated via audiovisual conferencing software. One third of the participants (n = 24) were assigned to the shared-knowledge condition. In this condition, objects were placed on the display and named by the experimenter while the audiovisual conferencing system was on-line (i.e., the speaker could see and hear everything being said in the other room). This procedure therefore made it clear to the addressee that the speaker should know the identity of all objects, including the actual identity of knowledge-based phonological competitors.

Another 24 participants were assigned to the privileged-identity condition. In this
condition, the audiovisual conferencing system was off-line at the beginning of each trial, as the objects were being placed on the display and simultaneously named by the experimenter. The system was then turned on just before communication between the speaker and addressee began. This is called the privileged-identity condition in view of the fact that the addressee should consider the speaker to be unaware of the knowledge-based competitors’ true identity on critical trials (e.g., that the candle that looks like a lightbulb is actually a candle or that the yo-yo that looks like a baseball is actually a yo-yo). In this condition, the addressee should assume that the speaker should categorize these objects in terms of their appearance (e.g., as a lightbulb, a baseball), an outcome that was clearly established in Experiment 2. In other words, for a listener who effectively tracks the speaker’s knowledge state, the speaker’s assumed lack of knowledge should eliminate the overlap in the name of this competitor type with the target object’s name. Although in this condition addressees should infer that the speaker is unaware of the true identity of visually misleading objects, the true identity of any non-Visually misleading object (including non-Visually misleading competitors, see Figure 2.1, Panels A and C) should, of course, be correctly inferred by the speaker from their outward appearance.

The remaining 24 participants were assigned to the privileged-existence condition. In this condition, objects were also placed on the display and were named for the addressee while the audiovisual conferencing system was off-line between trials. Furthermore, certain objects were hidden from the speaker’s perspective by placing opaque covers over shelf compartments (the sides facing the speaker’s view) before the audiovisual conferencing system was turned on again. On critical trials, one cover hid the competitor object, and the other cover hid an empty compartment. In this condition, the addressee should believe that the speaker could not know either the identity of the competitor or even the fact that the competitor object is present in the display. This corresponds to the co-presence manipulation used in a number of past studies
investigating common ground effects on real-time interpretation (see Figure 3.1 for example displays).

Figure 3.1. Illustration of a critical display (from the addressee’s perspective) as a function of knowledge state (the display was identical for the shared-knowledge and privileged-identity conditions) and competitor type as well as an example of the critical part of the verbal instructions given by the confederate speaker. Notes: (1) Asterisk indicates visually misleading object; (2) the black background within a display compartment represents an occluder panel whereby an object was visible to the listener but not to the speaker.

The 15 filler trials were the same as in Experiment 1. As in Experiment 1, all objects, visually misleading or not, were named for the addressee by the experimenter as they were
placed on the display (e.g., *In this trial we have a candle* [in conditions involving regular and knowledge-based phonological competitors], *a lightbulb* [in conditions containing the visual control object], *a tennis ball and a candy*). In addition to naming, the feature of any visually misleading object that would give away its identity was also demonstrated (e.g., the wick of a candle that looks like a lightbulb, the string of a yo-yo that looks like a baseball). For the addressees, the orientation of objects was the same as for the listeners in Experiment 1, whereas the speakers viewed the objects from the opposite side (as they were seen by participants in Experiment 2). As noted earlier, Experiment 2 confirmed that visually misleading objects could be readily oriented in a way that did not give away their identity to the speaker.

*Table 3.1. Knowledge State manipulation in Experiment 3*

<table>
<thead>
<tr>
<th>Knowledge state condition</th>
<th>Visual availability of competitor</th>
<th>Awareness of knowledge-based competitor’s true identity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shared-knowledge</td>
<td>Both interlocutors</td>
<td>Both interlocutors</td>
</tr>
<tr>
<td>Privileged-existence</td>
<td>Addressee only</td>
<td>Addressee only</td>
</tr>
<tr>
<td>Privileged-identity</td>
<td>Both interlocutors</td>
<td>Addressee only</td>
</tr>
</tbody>
</table>

### 3.2.1.3 Procedure

Participants (addressees) were led to believe that the confederate speaker was another naïve participant. Upon arrival, each addressee and the speaker (a female undergraduate student who would always arrive a few minutes later) were seated in the same room and asked to complete consent forms and questionnaires regarding language background. While in the same room, they were told that the purpose of the experiment was to investigate how people cooperate on a collaborative task via audiovisual conferencing software, and they were assigned their roles. Thus, addressees did not know that the instructions provided by the speaker were partly scripted.

Addressees were seated facing the shelf display and a computer monitor with a built-in
webcam connected to audiovisual conferencing software (see Figure 3.2.). The confederate speaker was then taken to another room and was seated in front of another monitor/camera running audiovisual conferencing software. The shelf display was positioned between the addressee and the monitor/camera, creating a situation in which the shelf display was viewed on the speaker’s display as mirror-reversed and from the opposite side relative to the addressee. As in Experiment 1, a SONY 1/3” CCD camera was attached to the center of the shelf display and recorded the addressee’s face, centered on the eye region. The experimenter (in the same room as the addressee) then briefly demonstrated the view that was available to the speaker via the camera/monitor facing the addressee. After that, the main task began. The experimenter named objects for the addressee as they were placed in the shelf display. As noted earlier, the audiovisual conferencing program was paused during this phase to create the privileged-knowledge conditions. Addressees were explicitly aware of the pause phase because they could see the experimenter turn off the audiovisual conferencing program and could observe that the video window on the computer display disappeared and the audio stream from the other room ceased. The end of the pause phase was fully evident to addressees as well because the experimenter had to use the conferencing software to initiate a call to the speaker and the speaker had to answer the call to establish video and audio streams for subsequent communication. After the audiovisual link was reestablished (if it had been paused), the experimenter gave a cue for the speaker to start her instructions.

Unbeknownst to the addressee, the speaker followed a partly scripted list of instructions. The critical instructions were the same as in Experiment 1; however, the speaker was allowed to improvise somewhat on non-essential parts of the instructions when the fluency of communication required or allowed for such improvisation (e.g., when the addressee picked up an incorrect object, the speaker corrected the addressee or paused for a longer period of time if
the addressee took a notably long time to perform an action).

![Figure 3.2. Bird’s-eye view of the physical arrangement of the participant (addressee) and confederate speaker in Experiments 3 and 4.](image)

3.2.1.4 Predictions

A comparison between the shared-knowledge and privileged-existence conditions provide a test of whether addressees take into account the speaker’s perspective about object existence/visual co-presence during the processing of a noun. For example, if addressees readily integrate knowledge that the speaker cannot see an occluded object during real-time referential comprehension, visual consideration of the regular phonological competitor and visual control
object should be similar while hearing the temporarily ambiguous part of the target noun in the privileged-existence condition. However, there should be more visual consideration of the regular phonological competitor compared to the visual control object in the shared-knowledge condition. A number of previous studies have shown that the latter type of knowledge discrepancy can be readily taken into account in the earliest moments of language processing. In addition, a comparison between the shared-knowledge and privileged-identity conditions provides a direct test of whether addressees are able to take into account the speaker's lack of knowledge about object identity during the interpretation of a referring expression. For example, if addressees in the privileged identity condition show efficient use of knowledge about the speaker's perspective as they process referring expressions, they should show comparatively little visual consideration of the knowledge-based phonological competitor (e.g., a candle that looks like a lightbulb) compared to the regular phonological competitor (e.g., a regular candle) while hearing the temporarily ambiguous portion of the target noun (e.g., *candle*). However, when the actual identity of knowledge-based competitor is mutually known (shared-knowledge condition), addressees should show comparatively greater consideration of the competitor as the target noun is heard.

### 3.2.2 Results and Discussion

#### 3.2.2.1 Coding

Coding procedures were as in Experiment 1. A total of 0.2% of trials were excluded because of experimenter error while placing the objects on the display.

#### 3.2.2.2 Statistical modeling

Statistical modeling was conducted in the same way as in Experiment 1. The model that properly converged included random intercept for participant and intercept as well as competitor type
slopes for items.

### 3.2.2.3 Eye movement data

Figure 3.3 plots the likelihood of a saccade to the competitor\(^4\) across the knowledge state and competitor type conditions\(^5\). The proportion of fixations to the competitor and target as a function of time, the type of competitor and knowledge state are shown in Appendix B. The logistic mixed effects model revealed a main effect of knowledge state, \(\chi^2(2) = 13.42, p = 0.001\), reflecting the pattern observed in Figure 3.3 whereby consideration of the competitor (across all competitor types) was overall reliably lower in the privileged-existence condition (likelihood = 0.07) than in the privileged-identity condition (likelihood = 0.17), \(\beta = 0.97, SE = 0.36, z = 2.66, p = 0.008\), or than in the shared-knowledge condition (likelihood = 0.22), \(\beta = 1.29, SE = 0.36, z = 3.60, p < 0.001\). Further, consideration of the competitor did not differ in these latter two conditions, \(p = 0.307\). In other words, when the competitor objects were occluded from the speaker’s view, addressees were less likely to look at these objects than when the competitor objects were visible to both an addressee and the speaker.

The results also revealed a significant main effect of competitor type, \(\chi^2(2) = 8.79, p = 0.013\). Follow-up pairwise comparisons revealed that visual consideration of the competitor was lower when the competitor was a visual control object (likelihood = 0.08) compared to when it was a regular phonological competitor (likelihood = 0.21), \(\beta = 1.32, SE = 0.51, z = 2.59, p = 0.010\), or a knowledge-based phonological competitor (likelihood = 0.17), \(\beta = 1.25, SE = 0.52, z = 2.38, p = 0.017\). However, competitor consideration was not significantly different for regular and knowledge-based phonological competitors (0.21 and 0.17, respectively), \(p = 0.860\). Thus there is overall less visual consideration of the competitor when the competitor is a visual control

---

\(^4\) Although the analyses were conducted in logit space, raw likelihood values are reported for easier interpretation.

\(^5\) When the three visually misleading objects whose visual features giving away their true identity were oriented towards the addressee were excluded from the analysis, the model failed to converge; however, the pattern of descriptive statistics remained the same.
object. The small likelihood of eye movements to the competitor in the visual control condition relative to the regular phonological competitor condition is not unexpected because in the former condition there is no match between the sounds in the unfolding target object name and the name for this object. More interesting is the lack of any difference for the regular and knowledge-based phonological competitors in the context of this main effect. This latter outcome suggests that addressees were consistent in their use of knowledge about object identity across the knowledge state conditions.

![Graph showing likelihood of saccade to competitor as a function of competitor type and addressee’s understanding of the speaker’s knowledge state, Experiment 3.](image)

**Figure 3.3.** Likelihood of a saccade to the competitor as a function of competitor type and addressee’s understanding of the speaker’s knowledge state, Experiment 3.

Although the interaction between knowledge state and competitor type was not significant, \( p = 0.838 \), I conducted two planned pairwise comparisons to test whether the activation of regular phonological competitor was higher relative to the visual control object within shared-knowledge and privileged-existence conditions. In the shared-knowledge condition, there should be more visual consideration of the regular phonological competitor compared to the visual control object. However, if addressees readily integrate knowledge that
the speaker cannot see an occluded object, the same comparison conducted in the privileged existence condition should show a similarly low degree of visual consideration of the regular phonological competitor and visual control object while hearing the temporarily ambiguous part of the target noun. In the shared-knowledge condition, the likelihood of saccades was indeed significantly higher when the competitor was a regular phonological competitor (e.g., regular candle, likelihood = 0.31) compared to when the competitor was a visual control object (e.g., a genuine lightbulb, likelihood = 0.10), $\beta = 2.43$, $SE = 1.18$, $z = 2.05$, $p = 0.040$. In the privileged-existence condition, the likelihood of saccades was not significantly higher when the competitor was a regular phonological competitor (likelihood = 0.10) compared to when the competitor was a visual control object (likelihood = 0.06), $p = 0.704$. This outcome is consistent with previous findings showing that a listeners’ sensitivity to the shared versus privileged status of objects in terms of their existence/physical co-presence is evident in the earliest moments of referential comprehension (e.g., Brown-Schmidt et al., 2008; Hanna et al., 2003; Heller et al., 2008). In the current study, the contrast between the shared-knowledge and privileged-existence conditions in particular resembles the kind of physical co-presence manipulations used in those studies.

Of greatest interest is the fact that the pattern of visual consideration of the competitor objects in the privileged-identity condition was in fact most similar to the one in the shared-knowledge condition. Recall that knowledge-based competitor objects should not reflect phonological competition in this condition if it were the case that addressees were bearing in mind the knowledge that should be attributed to the speaker. The fact that these objects are considered by addressees to the extent observed here therefore suggests that addressees cannot help but consider privileged knowledge about object identity to a rather strong degree. This clearly contrasts with addressees’ demonstrated ability to manage their privileged knowledge about object existence (in situations involving physical co-presence), as illustrated in both the
current experiment and earlier studies (Brown-Schmidt et al., 2008; Hanna et al., 2003; Heller et al., 2008).

One possible explanation for addressees’ observed performance in the privileged-knowledge condition could be that the cues that addressees can use to maintain a mental representation of the speaker’s “ignorance” (i.e., regarding the identity of the visually misleading objects) were comparatively weak. By contrast, in the privileged-existence condition, the opaque covers placed on the shelf display provide a comparatively salient and high-quality “reminder” for addressees about the speaker’s unawareness of occluded objects. It is therefore possible that addressees’ performance when faced with identity discrepancies would be boosted if they had additional cues to reinforce the speaker’s knowledge state. One relevant cue explored in a number of previous studies involves the way in which objects whose existence has been established via linguistic mention are efficiently recognized as constituting “shared information” (in common ground) by addressees, regardless of whether they are present in the physical context or in a shared field of view (Brown-Schmidt et al. 2008; Brown-Schmidt, 2009; Brown-Schmidt, 2012; Hanna et al., 2003). For example, using eye movement measures, Brown-Schmidt (2012) showed that addressees were more likely to consider a game-piece hidden to the speaker as a possible referent if that particular game-piece had been mentioned during previous conversational interaction. Hence, in Experiment 4 a previous mention manipulation is used to examine whether stronger cues regarding the speaker’s knowledge state may improve addressee’s ability to manage privileged information about object identity.

3.3 Experiment 4
According to Clark and Marshall (1978, 1981), information mentioned in a discourse becomes part of the mutually shared information in the communicative context and, as noted above, the
effects of this information on real-time referential processing are well established. In the current experiment, it was examined whether previous descriptions produced by a speaker facilitate addressees’ ability to track the speaker’s knowledge state when the descriptions explicitly reinforce the speaker’s belief about an object’s identity.

Similar to Experiment 3, participants in Experiment 4 were assigned the role of addressee and followed instructions from a confederate speaker in a different room via an audiovisual conferencing program. The communicative scenario was always the same as in the privileged-identity condition from Experiment 3, namely where the identity of the objects (including knowledge-based phonological competitors) was revealed only to the addressee before each trial while the audiovisual conferencing system was turned off (the shared-knowledge and privileged-existence conditions in Experiment 3 were not used in Experiment 4). As before, the type of competitor that appeared in the critical trials was manipulated. The object could be a regular phonological competitor (e.g., a regular candle) for the target name (e.g., candy), a knowledge-based phonological competitor (e.g., a candle that looks like a lightbulb), or a phonologically unrelated visual control object (e.g., a genuine lightbulb). This manipulation was crossed with a linguistic context manipulation. The linguistic context manipulation varied whether the confederate speaker referred to the competitor object when it occurred in what I refer to as a “context trial” occurring two trials before the corresponding critical trial. In the prior-mention condition, the context trial involved the speaker naming the competitor object in the course of describing the intended location for a different display object (e.g., Pick up the flower and put it above the lightbulb). Critically, for the visually misleading objects used as knowledge-based phonological competitors, the confederate speaker provided a name based on the objects’ outward appearance (e.g., the lightbulb for a candle that looks like a lightbulb). Thus, the name reflected the speaker’s lack of knowledge about the true identity of the visually misleading
object. In the no-prior-mention condition, the speaker did not refer to the competitor object in the context trials, although the action being communicated was the same (e.g., *Pick up the flower and put it one square to the left*). The verbal instructions and object layouts for critical trials were the same as in Experiment 3.

If an earlier description used by the speaker can help support an addressee’s ability to consider the speaker’s perspective regarding the identity of an object, then at the later point where the target object is named (e.g., *candle*), addressees should show less visual consideration of a knowledge-based phonological competitor (e.g., a candle that looks like a lightbulb) when the speaker had previously referred to the visually misleading object based on its appearance (e.g., *the lightbulb* for a candle that looks like a lightbulb) than when speaker did not previously refer to the object. Furthermore, consideration of a regular phonological competitor (e.g., a regular candle) and visual control object (e.g., a genuine lightbulb) should show high and low amounts of visual consideration, respectively, regardless of previous mention because their true identity is congruent with their appearance. (In these conditions, the label provided by the speaker on context trials reflected the competitor object’s true identity.)

### 3.3.1 Method

#### 3.3.1.1 Participants

The participants were 48 undergraduate students recruited from the University of Toronto Mississauga campus community. All participants were either native speakers of English or started to use English by age 7. An additional nine participants were excluded, either for strategically maintaining their gaze toward the center of the display throughout the entire experiment (n = 2), equipment problems (n = 5), because they reported suspecting that the speaker was not another naïve participant during the post-experiment debriefing procedure (n = 1), or because they reported noticing that some of the objects had names with overlapping initial
sounds \((n = 1)\). Participants were paid $10 per hour or received course credit for their participation. Participants were tested individually in one session that lasted approximately one hour.

3.3.1.2 Materials and design

The competitor type manipulation was identical to the one used in Experiments 1 and 3. In the current experiment, knowledge about the true identity of the visually misleading object was always privileged to addressees (i.e., as in the privileged-identity condition in Experiment 3); that is, objects were always placed on the shelf display and named only to the addressees while the audiovisual conferencing system was off-line. Thus, the addressee should expect the speaker to be unaware of a knowledge-based competitor’s true identity (e.g., the addressee should assume that the speaker should not know that a candle that looks like a lightbulb is actually a candle). Furthermore, as described above, I included a linguistic context manipulation in which the competitor object occurred in a display two trials earlier than the corresponding critical trial (see Figure 3.4 for an example of the trial sequence). In the prior-mention condition, one of the instructions in these context trials involved reference to the eventual competitor object (including knowledge-based phonological competitors), using a description that reflected the object’s appearance. In the no-prior-mention condition, the speaker did not refer to the competitor objects in the context trials at all. The no-prior-mention condition therefore resembled the privileged-existence condition in Experiment 1, where the addressee infers the speaker’s knowledge state based only on the assumption that the speaker will categorize a visually misleading object based on its appearance. On critical trials, the instructions provided by the speaker were identical to those in Experiment 3.
In addition to the nine critical trials and nine context trials, there were two practice trials and 15 filler trials. The filler trials were constructed to avoid contingencies that might bias addressees towards the target in the critical trials. Out of 33 trials in total, 11 had all-new objects and 22 (nine critical and 13 filler/context trials) had objects that appeared in previous trials. Eleven of the objects that appeared in previous trials (all filler and context trials) were objects that had to be moved on their second occurrence, whereas 11 were not moved in either trial (nine on critical trials and two in filler/context trials). Furthermore, in 16 trials (six critical and 10 filler/context), the first object to be moved was on the left side of the display and in another 17 trials (three critical and 14 filler/context), it was on the right side of the display. In addition, 16 trials (nine critical and seven filler/context) had displays that contained three objects and 17 trials (all fillers/context trials) had displays that contained four objects. Finally, seven filler trials and
context trials had displays containing hard-to-name objects (e.g., several constructor-set pieces attached to each other).

3.3.1.3 Procedure
The procedure was similar to Experiment 1. However, the experimenter always turned the audiovisual conferencing system off before placing objects on the display and naming them to the addressee (as in the privileged-identity condition in Experiment 1).

3.3.2 Results and Discussion
3.3.2.1 Coding
Coding procedures were as in Experiments 1 and 3. A total of 2.5% of trials were excluded because of experimenter error while placing the objects on the display or equipment failure (no sound recorded).

3.3.2.2 Statistical modeling
Statistical modeling was conducted in the same way as in Experiments 1 and 3. The model that properly converged included random intercept and random slope of competitor type for participants as well as random intercept, linguistic context and competitor type slopes for items.

3.3.2.3 Eye movement data
Figure 3.5 plots the likelihood of a saccade to the competitor\(^6\) across the linguistic context and competitor type conditions in Experiment 2\(^7\). (The proportion of fixations to the competitor and target as a function of time, the type of competitor and linguistic context are shown in Appendix C.) The logistic mixed effects model conducted on the saccade data revealed a main effect of

---

\(^6\) Although the analyses were conducted in logit space, raw likelihood values are reported for easier interpretation.\n
\(^7\) When the three visually misleading objects whose visual features giving away their true identity were oriented towards the addressee were excluded from the analysis, the pattern of descriptive and inferential statistics remained the same.
competitor type, $\chi^2(2) = 10.37, p = 0.006$. Follow-up pairwise comparisons revealed that visual consideration of the competitor was lower when the competitor was a visual control object (likelihood = 0.11) compared to when it was a regular phonological competitor (likelihood = 0.24), $\beta = 1.07$, $SE = 0.47$, $z = 2.25$, $p = 0.024$, or a knowledge-based phonological competitor (likelihood = 0.25), $\beta = 1.32$, $SE = 0.48$, $z = 2.78$, $p = 0.005$. However, competitor consideration was not significantly different for regular and knowledge-based phonological competitors (0.24 and 0.25, respectively), $p = 0.472$. Furthermore, the lack of a significant main effect of linguistic context, $p = 0.858$, or an interaction between competitor type and linguistic context, $p = 0.860$, suggests that this pattern was consistent across the no-prior-mention and prior-mention conditions.

Figure 3.5. Likelihood of a saccade to the competitor as a function of competitor type and linguistic context, Experiment 4.

In sum, the results of Experiment 4 suggest that addressees strongly considered their own privileged knowledge about the competitor’s identity as the ambiguous part of the target object name was heard, even when they had supplementary linguistic cues highlighting that the speaker
was unaware of the knowledge-based competitor’s true identity. In other words, the addressees made a similar proportion of saccades to the knowledge-based competitor relative to the regular phonological competitors even when speaker previously referred to the visually misleading objects on a basis of their appearance (e.g., *the lightbulb* for the candle that looks like a lightbulb). Thus, supportive linguistic information that explicitly signals the speaker’s knowledge state regarding the identity of the visually misleading objects did not change addressees' tendency to consider their privileged knowledge about these objects' identity.

### 3.4 Summary

The goal of the present chapter was to investigate how participants taking the role of addressees in conversation manage privileged knowledge about the identity of objects in the earliest moments of language processing, and to directly compare this to cases involving knowledge about the existence/physical co-presence of objects. In Experiment 2, participants were asked to name visually misleading objects without foreknowledge about their true identity to confirm that the true identity of these objects could not be recognized from just their outward appearance. The naming pattern confirmed that participants referred to the visually misleading objects using a term reflecting their appearance (e.g., *lightbulb* for a candle that looks like a lightbulb), rather than a term suggesting participants managed to guess their true identity (e.g., *candle*). In the main experiments (Experiments 3 and 4), eye movements were monitored as participants (addressees) followed a confederate speaker’s instructions in a conversational context in which the visually misleading objects served as competitor items in displays where another object served as the target (e.g., *candy*). In Experiment 3, it was found that, when the competitor objects were occluded from the speaker’s view, addressees were less likely to visually consider these objects upon hearing the target name than when they were visible to both an addressee and the speaker. This outcome is consistent with previous findings showing that the shared versus
privileged status of objects in terms of their existence/physical co-presence is evident in the earliest moments of referential processing (e.g., Brown-Schmidt et al., 2008; Hanna et al., 2003; Heller et al., 2008). In addition to replicating earlier findings, this outcome also confirms that the particular “two rooms” video chat methodology used here does not hinder the ability to detect listeners’ use of common ground during language processing. Interestingly, this specific result stands in contrast to Barr (2008), who also used a lexical competition paradigm to explore co-presence effects. In a computer-based interaction task, Barr found that when the target name (e.g., “buckle”) was heard, the proportion of fixations to a phonological competitor (e.g., a bucket) initially increased at the same rate as target fixations regardless of whether the competitor was in privileged or common ground (which was signaled by the color of a background square accompanying the image of the competitor). In Experiment 3, however, the likelihood of saccades to the regular phonological competitors was significantly higher compared to visual control objects when they were visible to both interlocutors but not significant when they were covered from the speaker’s perspective. One possible explanation for the discrepancy between the current results and those of Barr (2008) involves certain differences in the methodology used in the two studies. Barr (2008) presented photo object stimuli on a computer screen whereas in the current study addressees interacted with real objects. In addition, Barr (2008) used pre-recorded instructions, whereas the current study involved a live speaker located in a different room. According to Brennan and Hanna (2009), the use of pre-recorded materials may remove any potential for rich interaction and weaken the potential to detect partner-specific effects because they reduce the utility of taking the speaker’s perspective.

Importantly, addressees’ ability to successfully manage privileged knowledge about the *existence* of entities in Experiment 3 contrasted with their performance when knowledge about entities’ *identity* was at issue. Indeed, the most striking finding from Experiment 3 was that
listeners clearly considered privileged knowledge about the identity of objects in the course of incrementally mapping nouns to referential candidates. Indeed, the extent to which privileged knowledge was considered did not differ from a situation where the identity knowledge was shared between conversational participants.

In Experiment 4, I investigated whether the confederate speaker’s prior use of descriptions that highlighted her lack of knowledge about the visually misleading objects’ true identities would reduce an addressee’s tendency to consider privileged knowledge in the early moments of processing. Surprisingly, the results showed a pattern that was similar to the one found in Experiment 3: Even when the speaker had previously referred to a visually misleading object on a basis of its outward appearance—thereby reinforcing her understanding of the object’s identity—addressees still showed evidence of using privileged knowledge about object identity during referential interpretation. It is important to point out, however, that the current linguistic context manipulation and data patterns do not suggest that the early moments of processing will always reflect the intrusion of privileged knowledge about object identity. As suggested by a number of researchers (Brennan & Clark, 1996; Clark & Brennan, 1989; Brown-Schmidt, 2012), in some cases, establishing common ground may require a substantial amount of collaboration and interaction between the conversational partners. For example, Brennan and Clark (1996) showed that it required several turns to establish mutually agreed upon conceptualizations of ambiguous figures and to use these conceptualizations consistently in subsequent communication. In addition, Brown-Schmidt (2012) suggested that listeners’ sensitivity to other’s knowledge was somewhat moderated by the way the existence of occluded objects was revealed and by the type of feedback provided by the speaker. Thus, it is possible that more positive and/or negative evidence regarding the knowledge state of the speaker is required to reduce listeners’ consideration of privileged knowledge about the identity of objects.
in the course of incrementally mapping nouns to referential candidates. Nevertheless, the results of Experiments 3 and 4 suggest that the bias to take into account privileged knowledge regarding objects’ identity is comparatively strong and is not easily shifted by factors that have been shown to affect perspective-taking abilities in situations where knowledge discrepancies involve the existence/physical co-presence of referential objects.
Chapter 4: Knowledge Discrepancies Involving Identity in Narrative Comprehension

Referential perspective-taking and common ground are not only essential for communication in the here-and-now but also for the imagined worlds described in narrative discourse. For example, to understand subtle intricacies of narrative characters, their relationships, and the conversations they have, we need to take their knowledge and beliefs about each other as well as about the situation in general into account. Indeed, without the ability to track perspectives of multiple characters, complex fictional prose would be incomprehensible. Finally, as noted by Weingartner and Klin (2009), a large part of the enjoyment of reading fictional prose itself is experiencing the described events through the eyes of the characters.

A number of studies of narrative processing have demonstrated that comprehension is highly influenced by the protagonist’s perspective. Specifically, readers are sensitive to protagonists’ movements through space, their thoughts, knowledge, perceptual viewpoints, and goals (e.g., Gerrig et al., 2001; Glenberg, Meyer, & Lindem, 1987; Graesser, Bowers, Bayen, & Hu, 2000; Horton & Rapp, 2003; Levine & Klin, 2001; Morrow, Bower, & Greenspan, 1989; Morrow, Greenspan, & Bower, 1987). For example, Graesser et al. asked participants to make explicit judgments about what was revealed in the direct discourse of multiple characters as well as the characters’ knowledge of the information mentioned in these speech acts. The materials involved short literary narratives (“Almost Her” by Bridget Mazur and “The Conscience of the Cop” by William Fay). Ratings of the likelihood that a character was aware of information mentioned in a speech act depended on the character’s status in a conversation. For example, the readers assumed (predictably) that the speakers had the most knowledge about the information revealed in a speech act, but more important, the amount of information was rated successively less for addressees, overhearers, friends, and, finally, nonfriends, who received the lowest
“knowledge” ratings. These findings suggest that, at a minimum, readers are sensitive to relatively subtle differences between characters’ knowledge states.

In a real-time comprehension study, Lea et al. (1998) varied whether both a protagonist and a friend-character or only the protagonist knew about a side-character. They measured how long participants took to read a target sentence in which the friend-character used a pronoun to refer to the side-character when talking with the protagonist. The reading times for the target sentence were shorter when both the protagonist and the friend-character knew about the side-character than when the friend-character was ignorant about the side-character. These results show more specific ways in which readers are sensitive to a story character’s knowledge when interpreting the character’s utterances in real-time. Nevertheless, because the analysis region used for reading times included the whole sentence, it is unclear whether integration of a character’s knowledge happened immediately when the pronoun was encountered or only later as the rest of the sentence was encountered.

A number of other studies have suggested that in some circumstances, readers make unwarranted inferences about fictional characters’ knowledge (Keysar, 1994; Moreno-Ríos, Rodríguez-Menchén, & Rodríguez-Gualda, 2011; Weingartner & Klin, 2005, 2009). For example, Moreno-Ríos et al. asked participants to read stories that included either a sarcastic or sincere message left as a note by one story-character to another. Critically, the knowledge to understand the message as sarcastic was privileged to the writer of the message (as well as to the participant reading the story). The off-line ratings showed that participants assumed that, despite being ignorant, the addressee understood the message as sarcastic. Interestingly, participants were less inclined to assume that the message was interpreted as sarcastic when accidentally encountered by an ignorant third character in the story. This suggests that, although readers are often sensitive to the subtle knowledge differences of multiple story characters, they sometimes
fail to take into account differences between what they know and what a character knows, as reflected even in off-line measures.

Of particular interest to the present dissertation is whether, during spoken narrative processing, listeners are sensitive to knowledge discrepancies regarding the identity of referential entities. As suggested by the results of Experiments 3 and 4 where conversational scenarios were used, privileged knowledge regarding the identity of entities may be relatively difficult for listeners to ignore compared to privileged knowledge about the presence/existence of entities in the context. (Note that in this context I use the term “shared” to refer to information known to the listener that should also be ascribed to the fictional protagonist whereas “privileged” refers to information that the protagonist should be ignorant of, from the listener’s perspective.)

Building on these findings, the current study has three goals. By examining how listeners manage privileged knowledge about identity versus presence/existence in narrative, it is first possible to provide a test of the generality of the findings in Experiments 3 and 4. Second, the use of a narrative-based methodology can provide a way to evaluate a specific question raised at the end of Chapter 3, namely whether listeners’ greater ability to suppress knowledge about object presence/existence in interactive conversation is related to the strength and availability of the visual cues that signal co-presence. In narrative comprehension, the shared versus privileged status of both identity and presence information depends solely on discourse representations stored in memory rather than differing visual contexts. By using a narrative methodology, identity and co-presence discrepancies can be put on a more equal footing. A third goal of the current study is to test whether participants consider privileged identity information when identity information is additive rather than in conflict with appearance information. Specifically, in the subsequent experiments, I manipulated whether participants possessed privileged information about the aspect of a side-character’s identity (his/her occupation) in a context in
which this knowledge “added to” but did not contradict information that another individual (the fictional protagonist) possessed regarding the side character. This feature of the current study addresses the possibility that the difficulties that listeners experienced in suppressing their privileged knowledge about the identity of entities in the previous two experiments arises from the large discrepancy between what the misleading objects looked like and what they actually were.

As mentioned previously, in certain circumstances, managing knowledge about object identity is closely related to the notion of referential opacity. A referentially opaque context can be defined as a situation where linguistic terms that objectively might refer to the same entity in a given scenario such as “watch” and “present” cannot be straightforwardly substituted without potentially creating a formally untrue statement (e.g., Davidson, 1984; Ludwig & Ray, 1998; Ray, 1980; Roberts, 1993). For example, the statement “Tom said: ‘the watch looked expensive’” cannot necessarily be substituted with “Tom said: ‘the present looked expensive’” in a circumstance where Tom may have no idea that the watch is also a present for someone. The correct linguistic encoding of this event by a speaker or writer requires a direct consideration of Tom’s knowledge when deciding on a relevant referring expression. In Experiments 5 and 6, referentially opaque contexts are used to explore whether listeners are sensitive to story-characters’ knowledge about the identity of an entity in real-time narrative comprehension.

The current chapter address two questions: First, do listeners successfully encode and recall story-characters’ knowledge about the identity of an entity to the same extent as story characters’ awareness about the presence of an entity. If so, do real-time comprehension measures show that listeners are nonetheless more likely to consider privileged information about the identity of an entity than when privileged knowledge involves awareness of an entity?
To address these questions, in Experiment 5, I asked participants to listen to short narratives such as (5), (6) and (7) and then recorded their responses to questions probing for the story-protagonist’s knowledge about a side-character.

(5)  a. Susan was enjoying a latte at her favourite café on a stormy day. Meanwhile, a ranger from one of the national parks was helping track a bear that had escaped from the local zoo. The bear at that moment was still eluding capture. However, Susan didn’t know anything about the bear or the ranger, and was calmly sipping her coffee in the café.

b. As she finished the last sip of her latte, Susan said to herself “I think the raincoat I brought today was a good choice given the weather”.

c. Then she ordered another latte and watched the storm outside turn the city streets into small creeks.

(6)  a. Susan was enjoying a latte at her favourite café on a stormy day. When she was there, Susan would often see a tall man, who happened to be a ranger at one of the national parks. He was at the café that afternoon as well. However, Susan didn’t know that he was a ranger, or anything else about him, because they had never talked.

b. As she finished the last sip of her latte, Susan said to herself “I think the raincoat I brought today was a good choice given the weather”.

c. Then she ordered another latte and watched the storm outside turn the city streets into small creeks.

(7)  a. Susan was enjoying a latte at her favourite café on a stormy day. When she was there, Susan would often see a tall man, who happened to be a ranger at one of the national parks. He was at the café that afternoon as well. However, all Susan
knew about him was that he was a ranger, and nothing else beyond that, because they rarely talked.

b. As she finished the last sip of her latte, Susan said to herself “I think the raincoat I brought today was a good choice given the weather”.

c. Then she ordered another latte and watched the storm outside turn the city streets into small creeks.

Comprehension questions after the narratives such as (5) probed for the participants’ awareness about the fact that the side-character was not in the same location as the protagonist (e.g., Was the ranger in the café with Susan?). The questions following the narratives such as (6) and (7) probed for story-protagonist’s knowledge about a side-character’s occupation (e.g., Was Susan aware that the man in the café was a ranger?). If participants perform equally well on the comprehension questions following the three types of the narratives, this would suggest that participants are effective at encoding and recalling story characters’ knowledge about the presence/existence (whether the side character was in the same location as the protagonist) as well as identity (whether the protagonist knew that the side-character was as ranger or not) of an entity. Such an outcome would provide a solid starting point for the subsequent experiment that tests whether listeners are sensitive to the story character’s knowledge about the identity of an entity during real-time referential processing. In other words, it would suggest that any observed difficulty in coordinating privileged knowledge about the identity of an entity during real-time referential processing is not because of an inability to encode and retain this information.

In Experiment 6, the same narratives were used in a look-and-listen paradigm that provided a way to track listeners’ on-line referential hypotheses as target descriptions were heard near the end of each narrative (e.g., raincoat in “Susan said to herself: ‘I think the raincoat I brought today was a good choice given the weather’”). Critical displays included pictures of a
protagonist (e.g., Susan), an unrelated object (e.g., coffee cup), a side-character (e.g., tall man/ranger), and a target object (e.g., raincoat). The target object’s name shared onset sounds with the name of the side-character’s occupation and was always mentioned in direct discourse of the protagonist. In this task, the ability to manage privileged and shared information was measured by examining participants’ consideration of an image of a side-character. This measure should reflect this ability because the potential for the protagonist to refer to the side-character by using his/her occupation name (e.g., ranger) depended on whether the protagonist knew the side-character’s occupation. For example, in the absence of such knowledge (e.g., narrative [6]), the protagonist could not plausibly refer to the side-character using a term such as ranger, and the side-character should therefore not attract fixations as the target object’s name (raincoat) is heard.

4.1 Experiment 5

4.1.1 Method

4.1.1.1 Participants.

The participants were 27 undergraduate students recruited from the University of Toronto Mississauga campus community. All participants were either native speakers of English or started to learn English in North America from the age of 7 or earlier. Participants were paid $10 per hour or received course credit for their participation. Participants were tested individually in a single session that lasted approximately one hour.

4.1.1.2 Materials and design.

The experimental materials consisted of nine short narratives. The narratives were created anew for the current study. In each experimental narrative, the first four to five sentences introduced the context of the situation, the protagonist (e.g., Susan), and the side-character (e.g., tall
man/ranger). The last sentence of the introduction explicitly stated the protagonist’s knowledge about the side-character; see Table 4.1 for a sample narrative or Appendix E for the full list of experimental narratives. In the control version of an introduction, the side-character was described as being in a different location (e.g., chasing a bear) than the protagonist (who was sitting in a coffee shop). In addition, it was stated that the protagonist was unaware of the side character (e.g., “[Susan didn’t know anything about the bear or the ranger, and was calmly sipping her coffee in the café”). In the referentially-transparent version, the side-character was in the same location as the protagonist (e.g., a café) and it was stated that the protagonist was aware of the side-character’s occupation (e.g., “[all Susan knew about him was that he was a ranger, and nothing else beyond that, because they rarely talked”). In the referentially-opaque version, the side-character was again in the same location as the protagonist; however, it was stated that the protagonist was not aware of the side-character’s occupation (e.g., “[Susan didn’t know that he was a ranger, or anything else about him, because they had never talked”). In each version, the introductory sentences mentioned the side-character using his/her occupation name (e.g., ranger) twice. In referentially-transparent and referentially-opaque versions of an introduction the side-character was also referred to by an alternative name (e.g., “a tall man”) once. The sentence stating the protagonist’s knowledge about the side-character or his/her occupation was followed by the test sentence. The test sentence included direct discourse by the protagonist that mentioned an object whose name shared onset sounds with the name of the side-character’s occupation (e.g., “As she finished the last sip of her latte, Susan said to herself: ‘I think the raincoat I brought today was a good choice given the weather’ ”). The last one to two sentences of a narrative concluded the story. The narratives were between 99 and 160 words long. Each of the three versions of the experimental narratives were assigned to one of three stimulus lists such that each list contained an equal number of passages in each experimental
condition and a given participant encountered a specific passage only once.

After participants heard each narrative, a “yes-no” comprehension question appeared on a computer display in front of them. In the control condition, the critical comprehension question probed for the listener’s awareness about the location of the side-character (e.g., “Was the ranger in the café with Susan?”). The correct answer in this condition was “no”. In the referentially-transparent and referentially-opaque conditions, the critical comprehension question probed for the awareness about the protagonist’s knowledge of the side-character’s occupation (e.g., “Was Susan aware that the man in the café was a ranger?”). In the referentially-transparent condition, the correct answer was “yes”, and, in the referentially-opaque condition, the correct answer was “no”. After participants answered the critical comprehension question, a second “yes-no” comprehension question probed for other information that was provided either at the beginning or at the end of a narrative. (Five of the second comprehension questions probed for the information that appeared in the beginning of a narrative and four probed for the information that appeared at the end.) The correct answer for five of the second comprehension questions was “yes” and for the other four the correct answer was “no”. To answer the questions, participants had to press a labeled “yes” or “no” button on a computer controller.

In addition to the experimental narratives, 18 filler narratives were included to disguise the experimental manipulations. These narratives were more varied in their structure (range: 72 – 135 words) and did not adhere to the template used to construct experimental narratives. For example, six fillers included neither direct nor indirect speech attributed to story characters. In addition, nine fillers introduced information about knowledge states and intentions in different ways than in the experimental narratives (e.g., by using verbs and phrases such as “wanted”, “decided”, and “was disappointed that”). Four fillers mentioned a character’s occupation but did not refer to other objects whose name would partially overlap with the occupation’s name. For
the fillers, the first “yes-no” comprehension question after a narrative probed for information that appeared either in the beginning (six questions), middle (six questions), or at the end of a narrative (six questions). The correct answer for half of the questions was “yes” and for another half the correct answer was “no”. The second set of the “yes-no” comprehension questions was designed following a similar pattern. For each list, the 18 filler narratives were randomly interspersed with the nine experimental narratives with the constraint that there were no two experimental passages in the same condition occurring successively and that no more than two experimental passages occurred in sequence. All narratives were pre-recorded by the same female speaker.
Table 4.1. Example of the three versions of experimental narratives used in Experiments 5 and 6 and corresponding critical comprehension questions used in Experiment 5.

Example Narrative

Control Introduction
James was having a quiet Sunday morning stroll at his favourite flea market. Meanwhile, a local teacher was raising funds in James’s building for her school. Although James would never say no to donating money to the school, he knew nothing about the fundraising campaign or the teacher and, at the moment, was simply enjoying browsing through some antiques.

Transparent Introduction
James was having a quiet Sunday morning stroll at his favourite flea market. When he was there, he would usually bump into an older lady who happened to live in his building and who, in fact, was a teacher at a local school. She was at the flea market that morning as well. Despite living in the same building as her for many years, all James knew about the woman was that she was a teacher and that she enjoyed browsing at the flea market.

Opaque Introduction
James was having a quiet Sunday morning stroll at his favourite flea market. When he was there, he would usually bump into an older lady who happened to live in his building and who, in fact, was a teacher at a local school. She was at the flea market that morning as well. Despite living in the same building as her for many years, James was not aware that she was a teacher and only knew that she enjoyed browsing at the flea market.

Test Sentence
After looking at a beautiful old armchair, James said to himself “I think the teapot I saw last week is still for sale!”

Conclusion
He approached the seller and started to negotiate a price. James was happy to go home with a great buy that Sunday.

Control Comprehension Question
Was the teacher browsing at the flea market with James?

Transparent Comprehension Question
Did James know that the woman at the flea market was a teacher?

Opaque Comprehension Question
Did James know that the woman at the flea market was a teacher?
4.1.1.3 Procedure
Participants were randomly assigned to one of the three stimulus files with the constraint that an equal number of participants were assigned to each of the three lists. Narratives were played one at a time. Participants were instructed to listen for comprehension and were told they would be presented with two “yes-no” comprehension questions after each narrative. After each narrative was over, the first “yes-no” comprehension question appeared on the screen. Participants indicated their response by pressing a button labeled “yes” or “no”, and the second question then appeared. There was one practice trial to familiarize participants with the procedure.

4.1.2 Results and Discussion
4.1.2.1 Statistical modeling
The responses to critical comprehension questions were examined in a logistic mixed effects model (Jaeger, 2008) with participants and items as crossed, independent, random effects, implemented in package lme4 of the statistical software R 2.15.2 (Bates, Maechler, & Bolker 2012; R Core Team, 2012). Correct answers to the critical comprehension questions were coded as “1” and incorrect as “0”. The model that properly converged included random intercept for participants as well as random intercept and narrative type slope for items. The main effect of narrative type and follow-up pairwise comparisons were evaluated following the same modeling procedures as in Experiments 1, 3, and 4.

4.1.2.2 Comprehension scores
The overall average of proportion of correct answers to critical comprehension questions was 0.89\(^8\). The analysis revealed that the main effect of narrative type was not significant, \(p = 0.515\). The average proportion of correct answers in control condition was 0.83, in the referentially-

---

\(^8\) Although the analyses were conducted in logit space, proportions are reported for easier interpretation.
transparent condition it was 0.90, and in the referentially-opaque condition it was 0.94\(^9\). The results are consistent with previous findings showing that participants are generally quite effective at encoding and recalling relatively subtle information about a protagonist’s knowledge (e.g., Graesser et al., 2000; Kamawar & Olson, 1999). More importantly, the findings of this experiment suggest that listeners successfully encode and recall story character’s knowledge about the identity of an entity. For example, they were able to correctly recall whether Susan knew that the tall man in the café was also a ranger or not equally well in referentially-transparent and in referentially-opaque contexts. This outcome provides a solid starting point for the subsequent experiment that tests whether listeners are sensitive to the story character’s knowledge about the identity of an entity during real-time referential processing. Specifically, if real-time measures show less success at suppressing privileged knowledge about identity relative to privileged knowledge of an entity’s presence in the situation, it would be implausible (on the basis of the current results) to attribute the effect to differences in the ability to encode the relevant knowledge-state information across conditions.

4.2 Experiment 6

The goal of Experiment 6 was to explore whether listeners are sensitive to protagonists’ knowledge about the identity of other narrative characters as they are processing unfolding linguistic input, rather than when they are asked explicitly about it at the end of a narrative. In Experiment 6, participants listened to the same narratives as in Experiment 5. In addition, a look-and-listen paradigm was used to track listeners’ on-line referential hypotheses as target descriptions (e.g., raincoat) were heard. Each critical narrative was paired with a display, which included pictures of a protagonist (e.g., Susan), an unrelated object (e.g., coffee cup), a side-

\(^9\) Although numerically the performance was poorest in the control condition, the pattern was driven by a single question that was unintentionally ambiguous. If this question is excluded from the analyses, the main effect becomes even less significant and the average proportion of correct answers in control condition increases to .88.
character (e.g., a tall man/a ranger), and a target object (e.g., raincoat). If listeners are sensitive to the protagonist’s knowledge about the identity of an entity during real-time referential processing, they should direct their attention (as measured by eye movements) to the picture of a side-character (e.g., a tall man/a ranger) during the processing of a target noun (e.g., raincoat) in referentially transparent narratives more than in referentially opaque narratives. This is because the protagonist should be able to refer to the side-character by using his/her occupation name only when the protagonist is stated as possessing this knowledge (referentially transparent narratives) and not when he/she is described as ignorant of the occupation information (referentially opaque narratives).

4.2.1 Method

4.2.1.1 Participants
The participants were 27 undergraduate students recruited from the University of Toronto Mississauga campus community. All participants were either native speakers of English or started to learn English in North America from the age of 7 or earlier. An additional four participants were excluded, either because they correctly guessed the critical manipulation (n = 1) or because they correctly answered less than 75% of the comprehension questions on filler trials (n = 3). Participants were paid $10 per hour or received course credit for their participation. Participants were tested individually in a single session that lasted approximately one hour.

4.2.1.2 Materials and procedure
The critical and filler narratives as well as the narrative type manipulation were the same as in Experiment 5. However, in Experiment 6, participants were asked only one comprehension question after each narrative. Further, for critical narratives the sole question was the one that was asked second in Experiment 5 (namely, the question that did not probe for the knowledge
state of the protagonist) and for fillers it was the question that was asked first in Experiment 5. Thus, because the comprehension questions did not encourage listeners to consider the knowledge states of the story protagonists with regard to their knowledge of other characters, the experiment can provide a cleaner test of the extent to which listeners encode and use this information naturally during on-line narrative comprehension.

In the current experiment each narrative was paired with a display containing four photographic images (see Figure 4.1 for an example). Critical displays included a picture of a protagonist (e.g., Susan), an unrelated object (e.g., coffee cup), a side-character (e.g., tall man/ranger), and a target object (e.g., raincoat). Importantly, the target object’s name shared onset sounds with the name of the side-character’s occupation and was always mentioned in direct discourse of the protagonist (e.g., “Susan said to herself: ‘I think the raincoat was a good idea given the weather’ ”). Participants were instructed that the pictures on the display may or may not be related to the narrative and that they were free to look at them as much as they preferred. As participants performed the task, their eye movements were monitored. Following the logic of Experiments 3 and 4, listeners’ ability to manage privileged information was measured by examining participants’ consideration of the picture depicting the side-character because the potential for the protagonist to refer to the side-character by using his/her occupation name (e.g., ranger) depended on whether the protagonist knew the side-character’s occupation or not. For example, in the absence of such knowledge, the protagonist could not plausibly refer to the side-character using a term such as ranger, and the side-character should not attract fixations as the target object’s name (raincoat) is heard as part of the protagonist’s direct discourse.

Participants’ eye movements were recorded using an EyeLink II eye tracking system (SR Research Ltd., Mississauga, Canada). The eye that gave the most precise measures during calibration was used to record gaze position for the reading task. The eye tracker was set to
detect saccades with an amplitude of 0.5° or greater, using an acceleration threshold of 9500°/s² and a velocity threshold of 30°/s. The eye position was sampled every 2 ms. The experimenter monitored performance and recalibrated the eye tracking system as necessary.

![Figure 4.1. Example critical display used in Experiment 6. Top-left: side-character; Top-right: target object; Bottom-left: unrelated object; Bottom-right: protagonist.](image)

### 4.2.2 Results and Discussion

#### 4.2.2.1 Coding

Coding procedures were the same as in Experiments 1, 3 and 4. The analysis window was a 400 ms interval beginning 150 ms after and ending 550 ms after the onset of the target noun, corresponding to the first 80% of the target word, on average. The average length of the target object name was 477 ms (range: 393 – 562 ms). As in Experiments 1, 3 and 4, the trials where a saccade to the competitor occurred were coded as “1” and “0” otherwise.

#### 4.2.2.2 Statistical modeling

Statistical modeling was conducted in the same way as in Experiments 1, 3 and 4. The model that properly converged included random intercepts for participants and for items.
4.2.2.3 Eye movement data

Figure 4.2 plots the likelihood of a saccade to the side-character\(^{10}\) across the narrative type condition. The proportion of fixations to the side-character and target object as a function of time and narrative type are shown in Appendix D. The logistic mixed effects model revealed a main effect of narrative type, \(\chi^2(2) = 10.38, p = .006\). Follow-up pairwise comparisons revealed that visual consideration of the side-character was lower in the control condition (likelihood = 0.02) compared with both the referentially-transparent condition (likelihood = 0.16), \(\beta = 2.02, SE = 0.78, z = 2.60, p = 0.009\), and the referentially-opaque condition (likelihood = 0.12), \(\beta = 1.72, SE = 0.79, z = 2.17, p = 0.030\). However, side-character consideration was not significantly different across the referentially-transparent and referentially-opaque conditions (0.16 and 0.12, respectively), \(p = 0.500\).

The pattern of results suggests that when listeners knew that the side-character (e.g., the ranger) was not in the same location as the protagonist (e.g., Susan) and the protagonist did not know anything about him/her, they were less likely to look at the side-character’s picture upon hearing a word sharing speech sounds than when the side-character was in the same location as the protagonist and the protagonist had some knowledge about him/her. The fact that the side-character is considered less in the control than in the referentially-transparent and referentially-opaque narratives suggests that listeners are effective in suppressing their privileged knowledge about the overall presence of an entity within the broader context, as shown in Experiments 3 and 4, as well as in previous studies (cf. Brown-Schmidt et al., 2008; Hanna et al., 2003; Heller et al., 2008; Lea et al., 1998). However, when the side-character was in the same location as the protagonist, there was no detectable effect of whether the protagonist knew the side-character’s occupation or not: Visual consideration of the side-character was elevated in both the

---

\(^{10}\) Although the analyses were conducted in logit space, raw likelihood values are reported for easier interpretation.
referentially transparent and referentially opaque conditions. This was despite the fact that, without the knowledge of the side-character’s occupation (i.e., the referentially-opaque condition), the protagonist should not be referring to the character in terms of his/her occupation (e.g., the ranger) thereby eliminating phonological competition with the target object name (e.g., the raincoat). The elevated and statistically equivalent consideration of the side-character in the referentially-transparent and referentially-opaque narratives therefore suggests that addressees tend to consider privileged knowledge about the identity of referential entities in the course of interpreting a referring expression in real time. Thus, consistent with Experiments 3 and 4, listeners’ ability to manage knowledge discrepancies about identity clearly contrasts with their ability to manage knowledge discrepancies involving object existence/presence, even in a situation where both types of knowledge discrepancies rely on discourse memory in a similar way.

![Figure 4.2](image_url)

**Figure 4.2.** Likelihood of a saccade to the side-character as a function of narrative type, Experiment 6.
4.3 Summary

The goal of the present chapter was to investigate whether language users are sensitive to narrative protagonists’ knowledge about the identity of side-characters as referring expressions unfold in real time, and to directly compare this to cases involving knowledge about the awareness of side-characters. In Experiment 5, an off-line comprehension test was used to confirm that participants were able to successfully encode and recall protagonists’ knowledge about the side-characters. The results indeed suggested that participants were extremely effective in recalling whether the protagonists were aware of the side-characters’ occupation or not. These findings are consistent with a number of studies suggesting that language users are able to recall rather subtle differences between the story characters’ beliefs (e.g., Gerrig et al., 2001; Graesser et al., 2000; Kamawar & Olson, 1999; Moreno-Rios, Rodriguez-Menchen, & Rodriguez-Gualda, 2011). In Experiment 6, the same stories were used in a methodology where eye movements were monitored as participants listened to the short narratives and looked at displays in which the pictures of side-characters (e.g., picture of a man) served as competitor items in displays where another object served as the target (e.g., raincoat). The results showed that, when the side-character was described as being in a different location (e.g., the ranger chasing a bear) than the protagonist (e.g., Susan sipping coffee at a café), and it was stated that the protagonist was unaware of the side-character, there was little visual consideration of the picture of the side-character when participants heard the target object name (e.g., raincoat) as part of the protagonist’s direct discourse (control condition). Interestingly, however, there was a comparatively high proportion of saccades to the picture of the side-character upon hearing the target object name independent of whether the protagonist knew the side-character’s occupation, when he/she was in the same location as the protagonist (referentially-transparent and referentially-opaque conditions). This outcome is consistent with Experiments 3 and 4 as well as
previous studies showing that the shared versus privileged status of objects in terms of their co-presence is evident in the earliest moments of referential processing (e.g., Brown-Schmidt et al., 2008; Hanna et al., 2003; Heller et al., 2008). Furthermore, the current findings are in line with Horton and Rapp (2003), who demonstrated that accessibility of information about objects mentioned in a narrative was lower when they were described as being occluded from narrative protagonists’ perspective than when they were not occluded. Again, in contrast to Barr (2008), the results suggest that language users’ sensitivity to another person’s perspective can influence comprehension as individual speech sounds are mapped to noun candidates, at least in situations involving privileged knowledge about the existence/presence of side characters.

Importantly, participants’ ability to successfully manage knowledge of this sort in Experiment 6 contrasted with their performance when knowledge about the side-characters’ identity was at issue. Namely, as in Experiments 3 and 4, participants considered their privileged knowledge about the side-characters’ occupation in the course of incrementally mapping nouns to referential candidates. Indeed, the lack of difference between the referentially-transparent and referentially-opaque conditions suggests that the extent to which privileged knowledge about the side-characters’ identity was considered did not differ from a situation where the identity information was known to the protagonist. Based on the results of the off-line comprehension task from Experiment 5, the lack of difference between the referentially-transparent and referentially-opaque conditions cannot be explained by a lack of encoding or ability to interpret the protagonist’s knowledge about the identity of the side-character. This is because listeners performed extremely well on comprehension questions querying the protagonist’s knowledge about the occupation of the side-character.

Together the results demonstrate that language users still show evidence of considering privileged knowledge about entities’ identity (but not privileged knowledge about their
existence/presence in the context) in the course of real-time referential processing even when knowledge about other’s beliefs regarding identity and existence/presence involves fictional characters and relies on discourse representations stored in memory rather than readily available perceptual features and the configuration of visible objects in the here-and-now. Furthermore, language users considered privileged knowledge about identity even though it involved an aspect of a side-character (his/her occupation) which did not contradict any other information about the side character and was simply additive (e.g., in the case of the “ranger”, the side-character’s picture on a display simply depicted a regular man whose occupation would constitute additional rather than contradictory information).
Chapter 5: Conclusions

A number of early studies exploring the effects of common ground have shown that language users can effectively coordinate privileged and shared knowledge in live conversation and in narrative comprehension (e.g., Clark & Wilkes-Gibbs, 1986; Fussell & Krauss, 1989; 1992; Gerrig & Littman, 1990; Gibbs, Mueller, & Cox, 1988; Lea et al., 1998). Furthermore, recent investigations of the detailed time course of common ground integration during real-time comprehension have provided evidence that perspective differences are taken into account in the processing of a range of referring expressions in the “early” moments of processing (Brown-Schmidt et al., 2008; Hanna et al., 2003; Heller et al., 2008; Nadig & Sedivy, 2002). However, as mentioned in Chapter 1, previous research also left several open questions including the extent to which the ability to bypass privileged information might depend on different types of knowledge discrepancies. In the current dissertation, the focus is on knowledge differences in terms of language users' understanding of the identity of entities and how it contrasts with the knowledge about the presence of entities. This specific focus is motivated in part by work suggesting that these knowledge discrepancies might involve different mechanisms (e.g., Apperly & Butterfill, 2009; Frisson & Wakefield, 2012; Low & Watts, 2013). Furthermore, as discussed in Chapter 1, in many natural contexts, knowledge discrepancies involve identity information. Thus, investigating and contrasting the cases that reflect different types of knowledge discrepancies is essential to more fully understand the properties of real-time processing mechanisms as well as the nature and richness of knowledge representations in language comprehension.
5.1 Is identity information readily available when it does not match visual features?

In Experiment 1, I examined how information about an object’s true identity or essence is used upon hearing a referring expression in cases where the object’s visual features were incongruent with this information. For example, is the label candle readily considered for a candle that is known to be a candle but clearly looks like a lightbulb? The findings of Experiment 1 revealed that visual consideration of the regular phonological competitors (e.g., a regular candle) and knowledge-based phonological competitors (e.g., candle that looks like a lightbulb) were higher than for visual control objects (e.g., a regular lightbulb) but not significantly different from each other when a target object name (e.g., candy) was unfolding in real time. In contrast to Frisson and Wakefield (2012), these findings suggest that when participants have the ability to interact with a real visually misleading object and there is no ambiguity in terms of what function(s) an object can/cannot perform, the label that corresponds with the object’s function (e.g., candle) is readily accessible in real-time referential processing. This is despite the fact that misleading visual features remain comparatively salient throughout the task. Furthermore, this outcome validated the use of the same visually misleading objects for the subsequent experiments exploring whether people can overcome their own knowledge about the identity of visually misleading objects when this information is not available to their conversational partners. Experiment 2 provided an additional methodological validation, showing that the objects’ actual identity cannot be guessed in the absence of the relevant foreknowledge.

5.2 Are language users sensitive to others’ knowledge about object identity in conversation?

In Experiments 3 and 4, the visually misleading objects used in the earlier experiments were utilized to investigate how addressees manage privileged knowledge about the identity of such
objects in a conversational context. Here, in contrast to Experiment 1, participants’ (addressees’) eye movements were monitored as they followed instructions provided by a live (confederate) speaker rather than prerecorded instructions. However, just like in Experiment 1, the visually misleading objects (e.g., candle that looks like a lightbulb) served as competitor items in displays where another object served as the target (e.g., candy). In Experiment 3, it was found that, when the competitor objects were occluded from the speaker’s view, addressees were less likely to visually consider these objects upon hearing the target name than when they were visible to both an addressee and the speaker. This specific finding extends previous studies (e.g., Brown-Schmidt et al., 2008; Hanna et al., 2003; Heller et al., 2008) by showing that language users are sensitive to the shared versus privileged status of objects in terms of their physical co-presence, even within unfolding words, as individual speech sounds are mapped to noun candidates. Further, along with those earlier studies, this result rules out the idea that common ground information (broadly construed) cannot be used in the early moments of language comprehension for reasons arising from the basic architecture of processing mechanisms.

Interestingly, addressees’ ability to successfully manage privileged knowledge about the existence of entities in Experiment 3 contrasted with their performance when knowledge about entities’ identity was at issue. Specifically, addressees clearly considered privileged knowledge about the identity of objects in the course of incrementally mapping nouns to referential candidates. Indeed, the extent to which privileged knowledge was considered did not differ from a situation where the identity knowledge was shared between conversational participants. Moreover, even when the confederate speaker had previously referred to a visually misleading object on a basis of its outward appearance in Experiment 4—thereby reinforcing her own understanding of the object’s identity—addressees still showed evidence of using privileged knowledge about object identity during referential interpretation. Thus, the results of
Experiments 3 and 4 suggest that the tendency to temporarily consider privileged knowledge regarding objects’ identity is comparatively strong and is not easily shifted by factors that have been shown to affect perspective-taking abilities in situations where knowledge discrepancies involve the existence/physical co-presence of referential objects.

5.3 Are language users sensitive to protagonists’ knowledge about identity in narrative comprehension?

Referential perspective-taking and common ground are not only essential in a conversation in the here-and-now but also for the imagined worlds described in narrative discourse. Indeed, most complex prose would be impossible to comprehend if we were not able to differentiate between what is known and what is not known to different narrative characters. However it is possible that, like in conversational contexts, real-time comprehension is accompanied by temporary difficulty in suppressing privileged information about the identity of entities that should not be attributed to the knowledge of a fictional character. A narrative situation can provide a way to replicate and refine the pattern of findings observed in Experiments 3 and 4, using an extremely different kind of language situation. Furthermore, in contrast to the conversational contexts used in most contemporary studies, where knowledge differences about the presence versus identity of objects arise from different kinds of perceptual and configurational characteristics of a physical scene, in narrative comprehension these differences depend on discourse representations in a highly similar way (namely, information stored in memory that is drawn from previous sentences). A narrative situation therefore puts identity versus co-presence knowledge on more equal footing.

Thus, in Chapter 4, I explored whether language users are sensitive to protagonists’ knowledge about the identity of side-characters during narrative comprehension. The results of Experiment 5 confirmed that participants were very effective in identifying whether the
protagonists were aware of the side-characters’ occupation or not as measured by performance on off-line comprehension questions. This outcome is consistent with a number of studies suggesting that language users are able to correctly recall what is versus what is not known to different narrative characters (e.g., Gerrig et al., 2001; Graesser et al., 2000; Kamawar & Olson, 1999; Moreno-Rios, Rodríguez-Menchen, & Rodríguez-Gualda, 2011). Experiment 6 then provided a test of whether this information is effectively used during real-time comprehension. Eye movements were recorded as participants listened to the same short narratives as in Experiment 5 and looked at displays in which the pictures of side-characters (e.g., a picture of a man) served as competitor items for another picture that served as the target (e.g., a raincoat). When the side-character (e.g., a ranger) was described as being in a different location than the protagonist and it was stated that the protagonist was unaware of the side-character, listeners showed little visual consideration of the picture of the side-character when participants heard the target object name (e.g., raincoat) mentioned as part of the protagonist’s direct discourse.

Importantly, there was a significantly higher proportion of saccades to the picture of the side-character upon hearing the target object when he/she was described as being in the same location as the protagonist than when he/she was described as being in a different location than the protagonist. Thus, consistent with Experiments 3 and 4 as well as previous studies showing that the shared versus privileged status of objects in terms of their co-presence is evident in the earliest moments of referential processing (e.g., Brown-Schmidt et al., 2008; Hanna et al., 2003; Heller et al., 2008), language users are sensitive to others’ knowledge about the presence of an entity during referential processing, even in the comprehension of fictional narratives. In addition to replicating earlier findings, Experiments 3, 4 and 6 showed that language users readily integrate common ground about object existence at the level of lexical processing (mapping speech sounds to lexical candidates). As mentioned in Chapter 3, the only other study that used a
lexical competition measure rather than temporary referential ambiguity to explore the timing of common ground integration was Barr (2008). However, Barr did not observe any evidence that language users integrate common ground at the level of lexical processing.

The key result of Experiment 6, however, was that visual consideration of the side-character’s picture was not reduced when the protagonist was described as being unaware of his/her occupation (thus it was participant-listener’s privileged information). In other words, as in Experiments 3 and 4, participants were unable to avoid considering privileged knowledge about the side-characters’ occupation in the course of incrementally mapping nouns to referential candidates. Thus, taken together the findings from Experiments 5 and 6 suggest that, even when knowledge about other’s beliefs regarding identity and presence relies on discourse representations rather than readily available perceptual features (and when knowledge states are judged correctly “off-line”) and when there is no conflict between appearance and identity information (identity information is additive), listeners still show evidence of using privileged knowledge about entities’ identity (but not about their existence or presence) as nouns unfold in real time.

5.4 Why do language users consistently consider privileged knowledge about the identity of an entity?

Taking Experiments 3 and 4 as a starting point, one explanation for why privileged knowledge about the identity of an entity seems to be so routinely considered may be related to the salient properties of the visually misleading objects used in those experiments. In particular, the novelty items used as stimuli in Experiments 3 and 4 are quite striking in terms of the gap between what they look like and what they really are. It is possible that, once an individual recognizes that a visually misleading object is something very different from what it appears to be, it becomes interesting at some level to think about the contrast with its true identity. Thus, the acquired
knowledge about an object’s true identity might not be effectively inhibited, even when the communicative context suggests that the conversational partner does not possess this knowledge. According to this explanation, it would follow that materials involving less striking discrepancies about identity knowledge might lead to a different pattern of performance on the part of addressees (e.g., a wrapped present that the speaker believes is a gift for person A whereas the addressee knows it is intended for person B). On the other hand, an explanation based on the idea that the visually misleading objects were particularly interesting or noteworthy cannot explain the finding that participants in Experiment 3 did not have any trouble inhibiting consideration of these objects when they were covered from the speaker’s view (i.e., in the privileged-existence condition). In addition, participants considered privileged identity information in Experiment 6, where there was no conflict between appearance and identity. Specifically, in Experiment 6, participants possessed privileged information about the aspect of a side-character’s identity (his/her occupation) in a context in which this knowledge did not contradict any other information known about the side character. Rather, this information was simply additive (e.g., given the “ranger”, the side-character’s picture on a display simply depicted a regular man whose occupation would constitute additional rather than contradictory information).

Another explanation relies on the notion that the tendency to consider privileged knowledge about an object’s true identity arises because of a general bias toward psychological essentialism. As noted in Chapter 2, psychological essentialism involves the idea that most ontological categories (e.g., ‘insect’) have some stable underlying reality (Gelman, 2004). Once this reality is assigned to an object, altering its superficial appearance (e.g., making an insect perceptually similar to a leaf) does not remove its essence (Gelman & Markman, 1986; Gelman & Wellman, 1991; Keil, 1989). Indeed, even children draw inferences about objects based on
their essence even when it competes with their perceptual features (Gelman & Markman, 1986; Gelman & Wellman, 1991). In other words, observable changes in features of an object are not sufficient to alter what are perceived as its essential characteristics (e.g., even when a candle looks like a lightbulb, it is still a candle, not a lightbulb). Thus, it is possible that participants in Experiments 3 and 4 maintained salient mental representations reflecting the true identity of visually misleading objects because this information indeed reflects the underlying reality of these objects, and remains highly accessible even during the interpretation of language produced by individuals not in possession of this information. This explanation is consistent with the reading time patterns reported by Frisson and Wakefield (2012). As discussed in Chapter 2, in their study, participants read a series of stories in which animals were superficially transformed into other animals (e.g., a donkey was transformed into a zebra by painting stripes on its back). When the animal after the transformation was referred to using a term reflecting its true identity by a narrator (e.g., the donkey), participants read it faster than when it was referred to using a term reflecting its appearance (e.g., the zebra). However, reading times for the two types of terms did not differ when they were embedded into a direct discourse of a story character that was naïve to the transformation. In other words, Frisson and Wakefield showed a bias towards psychological essentialism as well as a failure to integrate a naïve character’s perspective about animal’s identity in real-time narrative comprehension. Nonetheless, the psychological essentialism bias cannot explain the results obtained in Experiment 6, where the privileged knowledge did not involve side-characters’ core identity. Namely, in the latter experiment, the listeners did not seem to suppress the privileged information about the occupation of the side-character which is relevant to the side-character’s identity but, arguably, does not constitute his/her essence.

A third and most adequate explanation for the observed differences between the existence
and identity cases could involve differences in the representational demands required to resolve the respective types of knowledge discrepancies. Namely, in the existence cases, the addressee’s task is to simply suppress an entity from a set of possible referents. This may be a comparatively easy task in terms of mental representation because the occluded object or the side-character who is in a different location than the protagonist is recognized as not being “perceived” by the speaker (whether live or as a narrative character) and, in turn, can be straightforwardly indexed as irrelevant or outside the currently relevant referential domain of interpretation. However, in the identity cases studied here, the range of referential candidates is the same in each case, and the task for the addressee or the narrative comprehender is to make an inference about the way in which the speaker categorizes elements in the set. The category assigned to a given object or character, in turn, restricts only the relevance of particular labels for referential communication. In other words, in identity cases, the addressee or the narrative comprehender needs to take into account the possible ways in which a real or fictional speaker can talk about objects or individuals, rather than whether he or she can talk about them at all. Thus, the knowledge discrepancies in identity cases involve selected subtypes of information associated with the referents rather than the referents as a whole.

The cognitive development literature provides support for this explanation, as tasks that require recognition of knowledge discrepancies involving only a subset of the features of referents (e.g., where knowledge that there is a tie in the context is shared, whereas knowledge that the tie is also a gift is privileged) are comparatively difficult for children (Apperly & Robinson, 1998, 2003; Hulme, Mitchell, & Wood, 2003; Sprung, Perner, & Mitchell, 2007). For example, Apperly and Robinson (2003) showed that five and six year olds had more difficulty taking into account that the story character did not know that a tie was also a gift rather than taking into account that the story character did not know that there is Lego rather than biscuits in
a biscuit box. Apperly and Butterfill (2009) have suggested that similar performance differences are found in adult cognition as well (also see Low & Watts, 2013). In particular, they claim that the ability of an adult to track other adults’ beliefs about objects they have not seen or are unaware of is rather automatic and quite efficient. In contrast, reasoning about others’ beliefs that involve complex combinations of properties and modes of presentation is effortful and cognitively demanding. However, this explanation cannot account for Hanna et al.’s (2003) findings showing early integration of common ground involving features that were not central to an object’s identity (e.g., a speaker correctly believed that there are two hidden jars but falsely believed that one of them contained olives). Nevertheless, an important aspect of the Hanna et al., study was that a conversational partner’s false belief was reinforced by explicit misinformation. Thus, it remains an open question whether the comprehension patterns observed by Hanna et al. would generalize to cases where another’s false belief is simply inferred from the situational context (e.g., where a listener who knows that an object appearing to be a glass jar is made of plastic will assume that it will be perceived as such by an ignorant speaker) rather than through explicit misinformation (e.g., where a listener witnesses an experimenter explicitly misinforming the speaker that an empty jar is full of olives).

Finally, it is also relevant to ask whether full suppression of privileged knowledge about identity would be adaptive or possible, even in a conversational context where it would seem that ignoring one’s own knowledge would be the most efficient strategy. On this approach it may be incorrect to describe listeners' use of privileged information about identity as reflecting "difficulty" in perspective taking, as opposed to rational behaviour. Recall that the intrusions of privileged information observed in the current dissertation were only temporary and that listeners did not act on the basis of these fleeting intrusions. One relevant observation here is that it has been argued that efficient perspective-taking in linguistic communication cannot be reduced to a
heuristic of considering only the information shared with an interlocutor, but in fact involves a focus on privileged knowledge in some contexts (Brown-Schmidt et al., 2008) and, more generally, is likely to involve a sophisticated process of balancing and integrating privileged and shared information. For example, when a conversational partner asks a question, one needs to attend to both privileged and mutually known information in order to correctly comprehend the question and to produce a relevant answer. Furthermore, even when communicative interaction does not involve direct requests about privileged information, simply suppressing privileged information does not seem particularly advantageous because we often share our privileged knowledge with conversational partners to facilitate communication. For example, in a conversation, whoever is a comprehender at one conversational turn is usually a speaker on the very next one and might want to describe that the lightbulb looking object is really a candle to their naïve partner. In such situations, the information about both the object’s true identity and the potential for mistaking it as something else would benefit from being salient rather than suppressed in some way. Further exploration of this notion might also tell us that the methodological approaches and measures we have been using are only weakly adequate for testing what psycholinguists have been focused on in terms of perspective taking abilities in the broader sense, and that we may be mistaking various signs of rational cognizing as failures in perspective taking. I suggest that discrepancies in identity knowledge may provide a particularly useful testing ground for exploring these ideas in future research.
References


Bates, D. M., Maechler, M., & Bolker, B. (2012). lme4: Linear mixed-effects models using S4
classes. R package version 0.999999-0. http://CRAN.R-project.org/package=lme4


Flavell, J. (1988). The development of children's knowledge about the mind: From cognitive
connections to mental representations. In J. Astington, P. Harris, & D. Olson (Eds.), *Developing theories of mind* (pp. 244-267). Cambridge, England: Cambridge University Press.


Processes, 29, 137–159.


Appendix A

The proportion of fixations to the target and competitor objects over time, across conditions (Experiment 1). The box with dashed lines represents the analysis window (200-700 ms after target noun [e.g., candy] onset).
Appendix B

The proportion of fixations to the target and competitor objects over time, across conditions (Experiment 3). The box with dashed lines represents the analysis window (200-700 ms after target noun [e.g., candy] onset).
Appendix C

The proportion of fixations to the target and competitor objects over time, across conditions (Experiment 4). The box with dashed lines represents the analysis window (200-700 ms after target noun [e.g., candy] onset).
Appendix D

The proportion of fixations to the target and side-character over time, across conditions (Experiment 6). The box with dashed lines represents the analysis window (150-550 ms after target noun [e.g., *raincoat*] onset).
Appendix E

All critical narratives used in Experiments 5 and 6 and critical comprehension questions used in Experiment 5.

1

Control Introduction. Susan was enjoying a latte at her favourite café on a stormy day. Meanwhile, a ranger from one of the national parks was helping track a bear that had escaped from the local zoo. The bear at that moment was still eluding capture. However, Susan didn’t know anything about the bear or the ranger, and was calmly sipping her coffee in the café.

Transparent Introduction. Susan was enjoying a latte at her favourite café on a stormy day. When she was there, Susan would often see a tall man, who happened to be a ranger at one of the national parks. He was at the café that afternoon as well. However, all Susan knew about him was that he was a ranger, and nothing else beyond that, because they rarely talked.

Opaque Introduction. Susan was enjoying a latte at her favourite café on a stormy day. When she was there, Susan would often see a tall man, who happened to be a ranger at one of the national parks. He was at the café that afternoon as well. However, Susan didn’t know that he was a ranger, or anything else about him, because they had never talked.

Test Sentence. As she finished the last sip of her latte, Susan said to herself “I think the raincoat I brought today was a good choice given the weather”.

Conclusion. Then she ordered another latte and watched the storm outside turn the city streets into small creeks.

Control Comprehension Question: Was the ranger in the café with Susan?

Transparent Comprehension Question: Was Susan aware that the man in the café was a ranger?

Opaque Comprehension Question: Was Susan aware that the man in the café was a ranger?

2

Control Introduction. Nora was picking up some vegetables and fresh fruit for dinner from her regular grocery store. Unknown to her, a pipe had burst in her apartment. Luckily, someone had noticed, and a plumber was called in to take care of the situation. Nora, on the other hand, did not know anything about the incident or the plumber and was blissfully strolling along in the grocery store.

Transparent Introduction. Nora was picking up some vegetables and fresh fruit for dinner from her regular grocery store. On entering the store, Nora noticed a man, who happened to be a plumber from a neighboring town. Nora was aware that he was a plumber because they had chatted at the store several times before.

Opaque Introduction. Nora was picking up some vegetables and fresh fruit for dinner from her regular grocery store. On entering the store, Nora noticed a man, who happened to be a plumber from a neighboring town. Of course, Nora was not aware that he was a plumber because she had never seen him before.

Test Sentence. As she pushed her shopping cart along the produce aisle, Nora said to herself “I think the plums look especially good today!”

Conclusion. Nora could not believe that fresh plums were available so early in the season. Once again her grocery store had lived up to her expectations.

Control Comprehension Question: Was the plumber at the grocery store with Nora?

Transparent Comprehension Question: Did Nora know that the man at the grocery store was a plumber?

Opaque Comprehension Question: Did Nora know that the man at the grocery store was a plumber?
3

**Control Introduction.** It was a holiday potluck at Jeannette’s office that day. Jeannette was waiting at the bus stop to take a bus to work. Meanwhile, on the other side of town, a man, who was a pilot from the nearby military base, had lost consciousness on the bus and as a result the bus was falling behind schedule. However, Jeannette didn’t know anything about the incident or the pilot and was patiently waiting for her bus to arrive.

**Transparent Introduction.** It was a holiday potluck at Jeannette’s office that day. Jeannette was waiting at the bus stop to take a bus to work. Usually, a middle aged man from her neighborhood would also be waiting for the bus. He was a pilot from the nearby military base. He was there that morning as well. Even though Jeannette would see him there almost every morning, all she knew about him was that he was a pilot and that he took the same bus.

**Opaque Introduction.** It was a holiday potluck at Jeannette’s office that day. Jeannette was waiting at the bus stop to take a bus to work. Usually, a middle aged man from her neighborhood would also be waiting for the bus. He was a pilot from the nearby military base. He was there that morning as well. Jeannette would see him there almost every morning, however, she didn’t know that he was a pilot and only knew that he took the same bus.

**Test Sentence.** As she stood at the bus stop waiting, Jeannette said to herself “I think the pie I am bringing to the potluck may be not enough for everyone, and so I should get something else before the party.”

**Conclusion.** When Jeannette eventually made it downtown, she stopped at a nearby bakery and picked up a box of cupcakes. Jeannette now felt confident that she had enough treats for everyone in the office.

**Control Comprehension Question:** Was the pilot waiting at the bus stop with Jeannette?

**Transparent Comprehension Question:** Was Jeannette aware that the man in the bus stop was a pilot?

**Opaque Comprehension Question:** Was Jeannette aware that the man in the bus stop was a pilot?

---

4

**Control Introduction.** James was having a quiet Sunday morning stroll at his favourite flea market. Meanwhile, a local teacher was raising funds in James’s building for her school. Although James would never say no to donating money to the school, he knew nothing about the fundraising campaign or the teacher and, at the moment, was simply enjoying browsing through some antiques.

**Transparent Introduction.** James was having a quiet Sunday morning stroll at his favourite flea market. When he was there, he would usually bump into an older lady who happened to live in his building and who, in fact, was a teacher at a local school. She was at the flea market that morning as well. Despite living in the same building as her for many years, all James knew about the woman was that she was a teacher and that she enjoyed browsing at the flea market.

**Opaque Introduction.** James was having a quiet Sunday morning stroll at his favourite flea market. When he was there, he would usually bump into an older lady who happened to live in his building and who, in fact, was a teacher at a local school. She was at the flea market that morning as well. Despite living in the same building as her for many years, James was not aware that she was a teacher and only knew that she enjoyed browsing at the flea market.

**Test Sentence.** After looking at a beautiful old armchair, James said to himself “I think the teapot I saw last week is still for sale!”

**Conclusion.** He approached the seller and started to negotiate a price. James was happy to go home with a great buy that Sunday.

**Control Comprehension Question:** Was the teacher browsing at the flea market with James?

**Transparent Comprehension Question:** Did James know that the woman at the flea market was a teacher?

**Opaque Comprehension Question:** Did James know that the woman at the flea market was a teacher?
5

Control Introduction. Anna was about to take a flight to France. It was her longtime dream to visit Paris and the southern regions of the country. As she was boarding the plane, she noticed a pair of gloves lying on one of the chairs at the gate. A doctor who had been on his way to a scientific conference had left them there. Anna had no idea that it was a doctor who had left his gloves behind.

Transparent Introduction. Anna was about to take a flight to France. It was her longtime dream to visit Paris and the southern regions of the country. She even shared her excitement with the man in the seat next to her. He was a doctor flying to a scientific conference. Anna learned only that he was a doctor, and not much else because he didn’t seem to be interested in conversation.

Opaque Introduction. Anna was about to take a flight to France. It was her longtime dream to visit Paris and the southern regions of the country. She even shared her excitement with the man in the seat next to her. He was a doctor flying to a scientific conference. However, Anna had no idea that he was a doctor because he did not tell her anything about himself.

Test Sentence. As she looked through her carry-on luggage for her headphones, Anna said to herself “I think the documents I packed are missing from my bag!”

Conclusion. She was worried that someone might have taken her passport as well. But then, Anna found everything in her jacket and ordered some wine to calm her nerves.

Control Comprehension Question: Was the doctor on the same plane as Anna?
Transparent Comprehension Question: Did Anna know that the man sitting next to her was a doctor?
Opaque Comprehension Question: Did Anna know that the man sitting next to her was a doctor?

6

Control Introduction. Michelle was having a great holiday in the Bahamas. She was taking some time off after closing a multimillion-dollar deal. While she was enjoying a tropical drink on the beach, the hotel staff were turning away a disappointed surgeon who had not made a reservation for himself in advance. Michelle knew nothing about the surgeon or his disappointment and so didn’t fully appreciate how fortunate she was for having booked early.

Transparent Introduction. Michelle was having a great holiday in the Bahamas. She was taking some time off after closing a multimillion-dollar deal. She was enjoying a tropical drink on the beach when she saw a man she had met back at the hotel, who happened to be a surgeon. Despite having talked to him for only a few minutes, Michelle had quickly learned that he was a surgeon because he made sure to reveal that fact right away.

Opaque Introduction. Michelle was having a great holiday in the Bahamas. She was taking some time off after closing a multimillion-dollar deal. She was enjoying a tropical drink on the beach when she saw a man she had met back at the hotel, who happened to be a surgeon. However, Michelle had not yet learned that he was a surgeon because he hadn't revealed that fact to her.

Test Sentence. As a big wave washed up against her beach chair, Michelle said to herself “I think the surfboard I rented might float away, I should tie it up.”

Conclusion. Michelle laughed at herself for renting the surfboard in the first place because she hadn’t got out of her beach chair all day.

Control Comprehension Question: Was the surgeon at the beach with Michelle?
Transparent Comprehension Question: Did Michelle know that the man at the beach was a surgeon?
Opaque Comprehension Question: Did Michelle know that the man at the beach was a surgeon?
Control Introduction. Ellen was going door to door in a wealthy neighborhood soliciting on behalf of a local charity. The house she was about to visit had a very nice garden. The owner of the house, who was a carpenter, had moved in last year. Ellen didn’t know the owner or that he was a carpenter and, in fact, had never been to that neighbourhood before.

Transparent Introduction. Ellen was going door to door in a wealthy neighborhood soliciting on behalf of a local charity. The house she was about to visit had a very nice garden. The owner of the house, who was a carpenter, maintained the garden by himself. Despite fundraising in that part of town for years, all Ellen knew about the owner was that he was a carpenter and that he loved to chat about his garden.

Opaque Introduction. Ellen was going door to door in a wealthy neighborhood soliciting on behalf of a local charity. The house she was about to visit had a very nice garden. The owner of the house, who was a carpenter, maintained the garden by himself. Despite fundraising in that part of town for years, Ellen was not aware that the owner was a carpenter because they would only chat about the garden.

Test Sentence. As she peered through the window in the door, Ellen said to herself “I think the carpet in the hallway is really beautiful.”

Conclusion. Ellen then knocked at the door hoping that the owner would be at home and would contribute generously to the charity.

Control Comprehension Question: Was the carpenter at home when Ellen knocked on his door?

Transparent Comprehension Question: Did Ellen know that the owner of the house was a carpenter?

Opaque Comprehension Question: Did Ellen know that the owner of the house was a carpenter?

---

Control Introduction. Debra decided to spend her Saturday stocking up on some things for home. She always preferred to buy items in bulk to save some money. Just after Debra entered Costco, a man, who happened to be a professional boxer, bumped into her car while backing out of a parking spot. However, Debra knew nothing about the boxer or what he had done to her car, and at that moment was diligently comparing the prices of two brands of toothpaste.

Transparent Introduction. Debra decided to spend her Saturday stocking up on things for home. She always preferred to buy items in bulk to save some money. Just after Debra entered Costco she noticed a fit looking man, who happened to be a professional boxer, standing in line at a checkout. Debra knew that he was a boxer because she always followed local sports very closely on TV.

Opaque Introduction. Debra decided to spend her Saturday stocking up on things for home. She always preferred to buy items in bulk to save some money. Just after Debra entered Costco, she noticed a fit looking man, who happened to be a professional boxer, standing in line at a checkout. Although she noticed his physique, Debra had no idea that he was a boxer because she didn’t follow local sports very closely.

Test Sentence. As she took a shortcut through the cleaning supplies to get to the food section, Debra said to herself “I think the boxes of family size breakfast cereal should be down this aisle.”

Conclusion. Debra was really enjoying herself. She took great satisfaction in the money she was saving.

Control Comprehension Question: Was the boxer at the store with Debra?

Transparent Comprehension Question: Was Debra aware that the man at the store was a boxer?

Opaque Comprehension Question: Was Debra aware that the man at the store was a boxer?
Control Introduction. Adam was a graphic designer who ran a small business out of the main floor of his house. He had just locked up the office because he needed to go shopping for some office supplies. Meanwhile, a woman who was the principal of the local elementary school was knocking on the door of his office because she wanted to order printed programs for the school play. Adam would have loved to do the job, but unfortunately he was entirely unaware of the principal and her visit and at that very moment was walking into the office supplies store.

Transparent Introduction. Adam was a graphic designer who ran a small business out of the main floor of his house. He had just locked up the office because he needed to go shopping for some office supplies. As he was rushing through the aisles, he had to dodge past a woman slowly pushing a cart almost overflowing with notebooks and pencils. The woman was in fact, the principal of a local elementary school. Adam recognized that she was a principal because he had attended her school.

Opaque Introduction. Adam was a graphic designer who ran a small business out of the main floor of his house. He had just locked up the office because he needed to go shopping for some office supplies. As he was rushing through the aisles, he had to dodge past a woman slowly pushing a cart almost overflowing with notebooks and pencils. The woman was in fact, the principal of a local elementary school. However, Adam had no idea that she was a principal because he hadn’t attended her school.

Test Sentence. As he passed the rows of pens and markers, Adam said to himself “I think the printer next to my desk is running low on ink, I had almost forgotten about it.”

Conclusion. Adam was always outraged at how expensive the refills were for a color printer. He often wondered how he managed to make ends meet at the end of the month.

Control Comprehension Question: Was the principal at the office supply store with Adam?

Transparent Comprehension Question: Was Adam aware that the woman at the office supply store was a principal?

Opaque Comprehension Question: Was Adam aware that the woman at the office supply store was a principal?