Stony Brook University

Assessing and accommodating addressees' needs:
The role of speakers' prior expectations and addressees' feedback

A Dissertation Presented

by

Anna Katharina Kuhlen

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Experimental Psychology

Stony Brook University

August 2010

Stony Brook University

The Graduate School

Anna Katharina Kuhlen

We, the dissertation committee for the above candidate for the Doctor of Philosophy degree, hereby recommend acceptance of this dissertation.

Dr. Susan E. Brennan, Dissertation Advisor
Professor, Department of Psychology

Dr. Arthur G. Samuel, Chairperson of Defense
Professor, Department of Psychology

Dr. Hoi-Chung Leung
Associate Professor, Department of Psychology

Dr. Marie K. Huffman
Associate Professor, Department of Linguistics

This dissertation is accepted by the Graduate School

Lawrence Martin
Dean of the Graduate School

Abstract of the Dissertation

Assessing and accommodating addressees' needs:
The role of speakers' prior expectations and addressees' feedback

by

Anna Katharina Kuhlen

Doctor of Philosophy

in

Experimental Psychology

Stony Brook University

2010

In dialogue, a conversational partner's behavior is not solely the product of one individual mind, but instead reflects a process of mutual coordination between both conversational partners. This project investigates how speakers' and addressees' behavior is shaped by addressees' informational needs and by speakers' prior expectations of addressees' informational needs.

In two experimental settings pairs of speakers and addressees were observed while giving directions (Experiment 1), or narrating short stories (Experiments 2a & 2b). Addressees had either high informational needs (the information they were receiving was new to them) or low informational needs (they were already familiar with the information). Speakers' expectations of addressees' informational needs were informed through prior experiences with the addressee (Experiment 1), or explicit information about addressees' knowledge made available prior to the interaction (Experiments 2a & 2b).

Results show that speakers' behavior was shaped by addressees' behavior: Corresponding with how much feedback they received, speakers shifted how clearly and deliberately they articulated target referring expressions, how many details they provided, whether they introduced salient information as new or old, and where they positioned this information syntactically. This suggests that addressee feedback is one important cue through which speakers monitor addressees' needs, and supports the assumption that information about the conversational partner can influence different levels of linguistic processing.

As for addressees, their behavior was shaped not only by their informational needs, but also by speakers' expectations of addressees' needs. This suggests that speakers' expectations modulate addressees' expression of their actual needs, perhaps by feeding back into the opportunities addressees have to give feedback, or by contributing to an implicit agreement on how to accomplish the task at hand.

These findings have theoretical implications for understanding dialogue as a collaborative process in which conversational partners mutually shape each other, integrating bottom-up information available online in the conversational situation with top-down expectations that are brought into the conversation. There are also methodological implications for how to go about studying dialogue, in particular with respect to the practice of replacing naïve addressees with experimental confederates who typically have very limited informational needs, thereby changing the nature of the interaction and the behavior under study.

Table of Contents

List of Tables

List of Figures

Acknowledgements

Though only my name appears on the cover of this dissertation, many people have contributed to its production. I owe my deepest gratitude to all those people who have made this dissertation possible.

First and foremost, I would like to sincerely thank Dr. Susan Brennan who has been a truly amazing advisor. While guiding, supporting, and inspiring me throughout my graduate studies at Stony Brook University she also let me develop and pursue my ideas freely. I learned a tremendous amount. Thank you. I would also like to thank the members of my committee, Dr. Arthur Samuel, Dr. Hoi-Chung Leung, and Dr. Marie Huffman for their generous help in developing my thesis. I owe my gratitude to my lab members, Alexia Galati, Karla Batres, Greg Perryman, and the members of the Gesture Focus Group for stimulating discussions and companionship. Special thanks goes to Dr. Elisabeth Brauner for her mentorship and friendship throughout the years.

I am deeply indebted to a fantastic team of research assistants: Marwa Abdalla, Krystan Lenhard, Balwin Ratra, and Heather Rayburn. And most notably, Chris Pinzone, Elizabeth Cohen, Emine Ozelkan, Laura Longo, and Sandy Weng for assisting me in my data collection and analyses.

Special thanks goes to Stony Brook's Department of Psychology, the most helpful and welcoming place in the world of psychology departments! Thank you also to Calion Lockridge for sharing his knowledge and materials; Anna Floyd, Michael Greenstein, and Vinny Grande for helping me pilot my study; Julie Weisenberg for helping me recruit volunteers for my experiments; and Katharina Schuhman for teaching me how to use Praat. I received invaluable advice on formatting, editing, and storing video files from Paul St. Denis and Christian Striboll. Thanks also to the Max Planck Institute for Human Development in Berlin, most notably Dr. Lael Schooler and the ABC research team, for welcoming me to their institute during the last months of writing my thesis.

I would not have made this without my wonderful friends who walked with me through ups and downs of graduate school (and life, in general): Adela Apetroaia, Katerine Aceti, Kristina Biskup, Agnes Brandt, Verena Doerfler, Jadzia Jagiellowicz, Anna Floyd, Mareike Fuchs, Annmarie McNamara, and Irina Theissen. I am also particularly grateful to my Brooklyn roommates for friendship and tolerance, Larissa Simpson, Julie Weisenberg, and Parker Caddell. Thanks also to my aunt Diana Couper for generously sharing her lovely home during my last stays in Brooklyn. And thanks to skype, which let me stay in touch with everyone I love.

Finally, and most importantly, I would like to thank my wonderful partner, Wolfgang Ruber, for his unwavering support, encouragement, patience, and love during the past ten years of my life; 6380 kilometers and 6 hour time difference could not sever our connection. Last, but not least, I want to thank my parents, Betty and Rainer, and my brother Michael for believing in me and role-modeling academic life.

Introduction

In everyday conversation, people generally avoid being told things they already know. When our best friend, for the tenth time, tells us how her westie won "Best in Show" we may urge her to move on by letting her know that we already heard this story before. We can do this explicitly by saying something like, "Yeah, yeah, I know. You told me last week". But we can also be more implicit and, without even interrupting, let her know whether this story seems noteworthy. For example, addressees who are told a piece of news they already know are said to respond with something like "*mhm*" or "*I know*", while addressees who are told a piece of news they don't know may respond with "*oh*" (Heritage, 1984).

Speakers usually honor their addressees' informational needs. The hypothesis that speakers produce utterances with their addressees' needs in mind has been called the *audience design hypothesis* (Bell, 1984; Clark & Carlson, 1982; Clark & Murphy, 1983). For example, speakers structure their utterances such that they build upon information that is known to both speaker and addressee, and highlight information that is new to the addressee (e.g., Clark & Haviland, 1977; Haviland & Clark, 1974). New and known information can thereby be distinguished in various ways. For example, new information may be pronounced more clearly, accented, and packaged syntactically such that it is easy to evoke or access. And known information may be pronounced less clearly, deaccented, and packaged earlier in the utterance.

Although it is relatively undisputed that speakers adapt to their addressees, it is debated when and how information about addressees affects speakers' utterances. While some argue that certain language processes are too fast and automatic to take addressees

into account (e.g., Bard et al., 2000; Keysar, Barr, Balin, & Paek, 1998), others argue that potentially all language processes can incorporate information about the addressee (e.g., Brennan & Hanna, 2009; Hanna, Tanenhaus, & Trueswell, 2003).

If speakers tailor their utterances to their addressees' informational needs, they must have some way of knowing what their addressees know and what they don't know. There are at least two types of cues speakers can use to draw inferences about addressees' knowledge: cues that are based on "top-down" knowledge, and cues based on perceptual information that becomes available "bottom-up" (e.g., Brennan, Galati, & Kuhlen, 2010; Kuhlen & Brennan, in press; Kuhlen, Galati, & Brennan, under review). Top-down, speakers may base their assumptions about addressees' knowledge on previous experiences in interacting with a particular addressee (Metzing & Brennan, 2003), or on characteristics of the addressee such as being a human versus a computer (Brennan, 1991), male versus female (Fussell & Krauss, 1992), or child versus adult (Newman-Norlund et al., 2009). Such prior expectations about addressees may shape what speakers say and how they go about saying it.

The other force with potential to shape speakers' utterances consists of bottom-up cues. These cues are available in the immediate conversational situation, for example addressees' verbal or nonverbal reaction to what speakers are saying. Speakers can monitor addressees' behavior moment by moment and adjust their utterances accordingly (Brennan, 2005; Clark & Krych, 2004). Such feedback behavior can provide speakers with information on whether addressees are attending to and understanding what speakers are saying (e.g., Duncan, 1973; Goodwin, 1981; Schegloff, 1981; Yngve, 1970). And it

may also contain information about whether this information is unknown or known to addressees.

Often, bottom-up cues corroborate top-down cues. For example, if speakers recall that they already talked with their addressees about a certain piece of information it is very likely that addressees' feedback behavior will confirm this. However, speakers can also have inaccurate expectations of addressees' knowledge, for example if they base their assumptions on stereotypes (e.g., their addressees' gender, see Fussell & Krauss, 1992; Kuhlen, 2004). In this case addressees' feedback behavior in the interaction may not confirm speakers' prior expectations. Little is known about how speakers integrate their prior expectations of addressees' needs with addressees' actual needs as expressed during the conversation in situations when they don't converge.

A deeper understanding of the mechanisms through which conversational partners adapt to each other has important theoretical implications for our understanding of how dialogue works. But it also has important practical implications about how we go about studying dialogue. When studying dialogue, researchers often replace one conversational partner, the addressee or the speaker, with an experimental confederate. One concern with this practice is that a confederate's behavior may differ from a naïve participant's behavior, thereby systematically changing the nature of the interaction and the behavior under study (e.g., Kuhlen & Brennan, 2008). This concern is especially pronounced when confederates play the role of addressees: Since confederates typically go through the same experimental procedures multiple times they quickly accumulate more knowledge than their conversational role affords. Their behavior (which often is unscripted) may reflect this and prompt participants to adapt to a particular type of addressee, namely one

who has no real informational needs. The question of under which circumstances confederates can replace conversational partners therefore engages with the question of how conversational partners shape each other.

The goal of this dissertation is to investigate the types of cues speakers use to adapt their utterances to addressees' informational needs. More specifically, this project investigates whether addressees' feedback signals their informational needs, and whether speakers integrate this information with their prior expectations of addressees' needs. In the following I will review further the concept of audience design and different models that have been proposed to accommodate audience design. Building on previous findings on the impact of speakers' expectations and addressees' feedback, I will then consider how speakers may assess their addressees' informational needs, which I will test with two experiments designed to address the questions at stake. I will then apply my findings to the question of when confederate addressees may elicit different behavior from speakers than naïve addressees.

1 Theoretical Background: Addressees shape speaking

1.1 Audience Design

Speakers design their utterances with their addressees in mind. This is particularly evident in situations when adults address children using an exaggerated prosody or more accessible syntactic constructions (Fernald & Simon, 1984; Shatz & Gelman, 1973), or when multilingual speakers use the language they know their addressee will understand (Bortfeld & Brennan, 1997). But even in more subtle ways speakers adapt to their addressees' needs on the basis of what they think addressees know or believe.

For example, in a classic study by Isaac and Clark (1987) pairs of participants worked together to identify and arrange a set of pictures of New York City landmarks. Participants were either familiar or unfamiliar with New York. At the beginning of the interaction, speakers quickly assessed their addressees' familiarity with New York and subsequently adapted their referring expressions to match their addressees' level of expertise. Speakers who were familiar with New York used more proper names and definite references (e.g., "the Citicorp building") when giving instructions to someone who was also familiar with New York. But when they gave instructions to someone unfamiliar with New York, they used fewer proper names and instead referred to features presented on the picture (e.g., "the building with the slanted roof").

The proposal that speakers design utterances with their addressees in mind has far-reaching implications for our understanding of the processes underlying speech production[1]. Understanding how and when information about the conversational partner

[1] Audience design also has implications for our understanding of speech comprehension: Not only when speaking, also when listening, do conversational partners take each

influences speaking informs and limits the kind of cognitive models used to explain speech production. But even beyond architectural considerations, the idea that conversational partners influence speech production (a process traditionally assumed to depend solely on the speaker) implies that these processes are fundamentally shaped by the social context they are embedded in. Language production, and language use in more general terms, is therefore more accurately understood as a collaborative activity, as "joint action" (e.g., Clark, 1996; Schober & Brennan, 2003; Tanenhaus & Brown-Schmidt, 2007). Such a perspective makes conversational partners and their role in shaping speech production a focal research interest.

1.2 Audience design on different linguistic levels of speech production

Speakers can take their addressees into account on various linguistic levels of speech production. For example, as seen in Isaac and Clark's study reviewed earlier, speakers can adapt their lexical choices to their addressees' knowledge by choosing one referring expressions over another (see also e.g., Metzing & Brennan, 2003). Speakers can also decide to embellish their utterances with more or fewer details (e.g., Kuhlen & Brennan, in press), emphasize certain information by choosing one syntactic framing over another (e.g., Lockridge & Brennan, 2002), articulate more or less clearly (e.g., Galati & Brennan, 2010; Samuel & Troicki, 1998), or illustrate their ideas with a hand gesture (e.g., Kuhlen, Galati, & Brennan, under review).

---

other's knowledge into account. For example, addressees interpret speakers' referring expressions based on what they know speakers can see, and hence could be referring to (Hanna, Tanenhaus, & Trueswell, 2003). For reasons of brevity, the current review will however focus on how addressees shape processes of speech production.

Different architectural models have been proposed to accommodate these adaptations. These models differ in their assumptions about *when* these adaptations occur in the process of speech production, and whether partner-specific information finds entry to *all* levels of speech production.

Dual process models

Some conceptualize language production as having a cascading architectural structure, in which speakers first plan the message of their utterance, next encode this message into a syntactic structure, then into a phonological representation, and then into a motor plan specifying its articulatory execution (Bock & Levelt, 1994; Garrett, 1980; Levelt, 1989). Based on the idea of such a cascading architecture, it has been proposed that information about the conversational partner influences only *some* levels of language production (e.g., Bard et al., 2000; Bard & Aylett, 2001). While slow and "inferential" processes (such as the decision to encode certain information or not) can be shaped by partner-specific information, fast-acting and automatic processes (such as articulation) are thought to be impervious to information about the conversational partner. According to this perspective, certain aspects of speakers' utterances are not adapted to addressees' needs. Instead, speakers construct utterances based on their own knowledge or perspective (see also Barr & Keysar, 2002, 2005; Keysar, Barr, Balin, & Paek, 1998; Keysar, Barr, Balin, & Brauner, 2000; Kronmüller & Barr, 2007).

In support of this perspective, Bard and colleagues (2000; see also Bard & Aylett, 2001) show that speakers adapt their articulation only in response to their own needs. Easing their own productions costs, speakers articulate referring expressions less clearly upon mentioning them the second time. Crucially, Bard and colleagues argue that

speakers do not adapt their articulation in response to addressees' needs, failing to articulate referring expressions more clearly when this information is new to addressees. Drawing upon a series of corpus-based studies, their observations are based on interactions in which one participant (the director) gives another (the matcher) directions on how to follow a route marked into a map. Both directors and matchers have maps depicting landmarks that can be referenced when describing the route. But directors' and matchers' maps are not identical: Some landmarks appear only on one participant's map. One of the challenges of this task is therefore to find out which landmarks are not mutually shared. To measure directors' articulation, their referring expressions to landmarks were analyzed with respect to their duration and their rated clarity.

In one of these studies (Experiment 2), Bard et al. explore whether directors adjust their articulation of landmarks in response to matchers' feedback. The rationale is that, upon mentioning this landmark for the second time, directors should attenuate their referring expression to a lesser degree when matchers previously indicate that the referenced landmark does not appear on their map. A comparison of those instances in which matchers explicitly stated the landmark was missing with those instances in which matchers failed to indicate so revealed that directors were equally likely to attenuate their referring expression, independent of matchers' feedback. Bard et al. conclude that directors were either not monitoring matchers' feedback, or they were not adjusting their referring expression in response. Together with other studies indicating directors *do* adapt to their matchers' needs with regard to the syntactic form of their referring expression, Bard and Aylett (2001) propose that partner-specific information (for example as

communicated through their feedback) can influence only slower processes of speech production, but has no influence on fast processes, such as articulation.

However, aspects of the experimental design may compromise the conclusion that articulation is indeed impervious to addressees' informational needs. One concern is that Bard and her colleagues should have compared their data with a more decisive control group (see Galati & Brennan, 2010). In the reviewed study, such a control group would have been those instances in which directors refer to landmarks that matchers had previously confirmed to have marked on their map. Following a similar rationale Bard et al. developed, directors should attenuate more strongly upon second mention those landmarks that appear on both director's and matcher's maps than those landmarks that appear only on the director's map. In other words, the current study only explores the role of negative feedback ("I can't see the landmark"), but not the role of positive feedback ("I see this landmark") in assessing matchers' needs.

Perhaps more crucially, it is doubtful whether matchers' feedback upon hearing the landmark referred to for the first time should be considered informative of matchers' informational needs upon hearing the landmark for the second time: Since landmarks were introduced upon first reference, it is likely that thereafter directors considered them part of their mutually shared knowledge, despite the fact that matchers' maps lacked its picture (see Galati, 2010 on how verbal reference can inform assessment of common ground, even without shared visual evidence). Matchers may not have had the need for a more deliberate articulation of these landmarks, and directors therefore had no reason to accommodate their articulation. Due to these methodological concerns, further empirical studies are needed to either validate or revise Bard and her colleagues' proposal.

Constraint based approach

Alternative to a modular view, others suggest that language processing is better characterized as a constraint based process, in which syntax, semantics, and phonology are separate, but equal aspects of speakers' utterances (e.g., Kaplan & Bresnan, 1982; Pollard & Sag, 1994). According to this view, information about the local discourse context (e.g., knowledge of the addressee) is simultaneously integrated in a probabilistic fashion with other information, such as syntactic or lexical biases based on global frequencies. Information about the conversational partner therefore has the potential to influence language production in early moments of processing and on *all* levels of linguistic representation (e.g., Brennan & Hanna, 2009; Hanna, Tanenhaus, & Trueswell, 2003; Hanna & Tanenhaus, 2004). Whether and when information about the conversational partner finds entry to speech production processes depends on how soon this information becomes available.

In support of this proposal, Galati and Brennan (2010) find that speakers can take their addressees' knowledge into account at the level of articulation. Their study follows up on one of Bard and her colleagues' studies (Bard et al., 2000, Experiment 1), which found speakers attenuate referring expressions upon second mention even when they are interacting with a new addressee who hears the reference for the first time. Galati and Brennan add to their design a crucial control: Unlike Bard et al., they not only compare the first referring expression to one addressee with the second referring expression to a new addressee, but also to the second referring expression of speakers who interact with the same addressee. Such a comparison suggests that speakers attenuate repeated references *less* when interacting with a new addressee compared to when interacting with

an old addressee. Galati and Brennan propose that speakers' articulation can at least in part be influenced by addressees' informational needs- as long as addressees' needs are easily computable by speakers.

1.3 How speakers know what addressees know

In order to adapt to their addressees' informational needs, speakers need to know what their addressees' needs are. Different sources of information have been proposed to flow into this assessment. One such possible source are top-down cues, prior beliefs about addressees' needs based on conceptual information about the addressee. Another possible source of information relies on bottom-up cues, which are available in the conversational situation, such as their addressees' behavior.

Top-down cues: The role of speakers' expectations

Prior to speaking, speakers commonly have at least some information about their addressees. Even with very little information (e.g., the addressee's gender or university affiliation) speakers can estimate quite accurately someone else's knowledge, and they use this assessment to adjust how they refer to objects or people (Fussell & Krauss, 1992). Speakers' prior expectations of their addressees' knowledge can therefore guide at least their initial utterance formulation.

Speakers' expectations of addressees' needs can have such a strong influence that they ignore evidence to the contrary (Russell & Schober, 1999). In this referential communication study speakers were led to expect their addressees to have either high or low informational needs: While some speakers expected addressees to need very detailed information to successfully identify a referent, other speakers expected their addressees to need less detailed information. In the interaction, speakers encountered addressees who

had informational needs either consistent or inconsistent with their expectations. Speakers who expected addressees with low informational needs initiated referring expressions in one single conversational turn, while those who thought addressees needed more detailed information initiated multiple-exchange contributions. Crucially, even though addressees' feedback could have informed speakers about addressees' real informational needs, speakers did not revise their initial expectations. Instead, they seemed to attribute their addressees' inconsistent behavior to particular personality traits, making the "fundamental attribution error" (Heider, 1946; Ross, Greene, & House, 1977).

Such a strong influence of speakers' expectations is supported by studies in social psychology that suggest that the expectations with which speakers enter an interaction can shape the course of the interaction. This phenomenon has been called *behavioral confirmation* and has been demonstrated in variety of contexts (see e.g., Snyder & Klein, 2005; Snyder & Stukas, 1999; Snyder, 1984; Rosenthal & Jacobson, 1968, 1992). For example, in a study by Snyder, Tanke, and Berscheid (1977), men were led to believe they would be interacting with either a physically attractive or an unattractive woman over the telephone. When men expected to be interacting with an attractive woman, independent judges perceived the woman's behavior as more sociable, poised, humorous and socially adept than when the woman was supposedly unattractive, thus confirming the stereotype that attractive people also have more agreeable personalities. Snyder and colleagues concluded that the men's expectations of their female partner motivated them to behave in a way that caused the woman to act according to their expectations.

But many studies suggest that speakers' prior expectations shape their behavior only initially. Speakers then update their expectations in light of the behavior addressees display in the interaction. For example, speakers expecting to interact with a natural-language-processing computer initially designed utterances differently than speakers who expected to be interacting with a remotely located human partner (Brennan, 1991). Initial utterances addressed to presumed computers were short and telegraphic, whereas initial utterances addressed to human partners were more polite, contained more first and second person pronouns, and were more likely to be full grammatical sentences. Over the course of the interaction, however, speakers adapted their utterances to match their partner's response style, responding to short, telegraphic answers with short, telegraphic questions, and to sentential answers with sentential questions.

Along similar lines, when adult speakers expected to play an interactive computer game with a child they spent more time communicating with the presumed child, marking crucial information through extended pauses, than when they thought they were interacting with another adult (Newman- Norlund et al., 2009). These adaptations initially occurred irrespective of the behavior addressees displayed; a confederate acted as both child and adult partner, blind to the experimental condition. However, over time the performance of addressees in the game gained more weight than speakers' initial expectations and the magnitude of communicative adaptations to the presumed child decreased. These results suggest that utterances may initially be shaped by speakers' expectations of their addressees, but are then adapted to the type of behavior they experience over the course of the interaction.

Bottom-up cues: The role of addressee feedback

Another possible source of information are therefore cues available in the conversational situation. Instead of maintaining more or less elaborate representations of their addressees, the advantage of bottom-up cues is that these computational costs can be offloaded into the immediate conversational environment. For example, when interpreting speakers' utterances, addressees have been shown to take into account what speakers can see (Hanna, Tanenhaus, & Trueswell, 2003), are currently looking at (Hanna & Brennan, 2007; Brennan, Chen, Dickinson, Neider, & Zelinsky, 2008), or can reach (Hanna & Tanenhaus, 2004). For speakers, one of the most prominent cues available in the conversational situation is addressees' feedback.

When listening to speakers' utterances, addressees are anything but passive recipients of information. Addressees may interject speakers' utterances with a short "yeah", "right", or "uh-hm", change their facial expressions, nod their head, smile, or laugh. These kinds of behavior have been termed "backchannel responses" (Yngve, 1970), or addressee feedback. Addressee feedback is thereby commonly understood as any kind of verbal or nonverbal behavior on the part of addressees that indicates they are attending, following, appreciating, or reacting to what speakers are saying (e.g., Bavelas, Coates, & Johnson, 2000; Duncan, 1973; Goodwin, 1981; Schegloff, 1981; Yngve, 1970). Feedback may be given in the brief pauses between speakers' utterances or may overlap with speakers' utterances (Goodwin, 1986).

The amount of feedback addressees give has been shown to affect speakers in different ways. Without addressee feedback speakers' utterances become less efficient. In classic referential communication studies by Krauss and his colleagues, speakers

instructed addressees via intercom to select a target referent among a set of abstract pictures. When confederate addressees were instructed to refrain from producing feedback, speakers used more words to refer to the pictures, presumably because speakers had no evidence that their addressees had understood (Krauss & Weinheimer, 1966). Such an effect of addressee feedback (or lack thereof) suggests that feedback indicates addressees' current understanding. Addressee feedback is therefore instrumental in the process of grounding (Clark & Brennan, 1991). By giving feedback, addressees accept or reject speakers' current utterance. By monitoring addressees' feedback, speakers are able to make their utterances more efficient without running the risk of becoming incomprehensible.

A different set of studies suggests that addressees' feedback enables speakers to speak more fluently, more vividly, and with more detail. In a study by Bavelas, Coates, and Johnson (2000) speakers told addressees an autobiographical story. Addressees' level of attention to the stories was manipulated by instructing half of the addressees to count the number of words speakers produced that begin with the letter /t/. This caused addressees to give less feedback. Speakers who received less feedback told their stories less well. In particular, the endings of the narratives, which were intended to be suspenseful, fell flat. A similar study by Pasupathi, Stallworth, and Murdoch (1998) found that speakers who received little feedback from their distracted addressees were less fluent and tended to speak for less time. And a third study confirmed that speakers narrate with less vivid details when distracted addressees give less feedback (Kuhlen & Brennan, in press). This pattern could be interpreted to mean that feedback indicates addressees' level of engagement, thereby motivating speakers to elaborate.

Although addressees' feedback clearly impacts speakers' utterances, it is unclear whether and how addressees use their feedback to communicate their informational needs. Addressee feedback is usually studied by manipulating it directly, for example by instructing addressees not to give feedback, or by preventing addressees from giving feedback by distracting them. By testing how this affects speakers' behavior, hypotheses about the function of addressee feedback can only be inferred retrospectively. Direct investigations of the factors influencing addressee feedback are rare. Such an investigation could manipulate addressees' informational needs and then measure to what extent addressee feedback reflects their informational needs.

Addressees may convey their informational needs not only by how frequently they give feedback, but also by when and how they give feedback. Addressee feedback that is delayed with respect to speakers' utterances has been shown to have a detrimental effect on speakers. In another referential communication study by Krauss and Bricker (1967; see also Krauss, Garlock, Bricker, & McMahon, 1977) addressees' feedback was systematically deferred by inserting a delay loop into the circuit. A delay of 1.8 seconds was sufficient to disrupt an efficient communication.

In support of this view, studies have shown that the delay in answering a question is indicative of the respondent's knowledge (Smith & Clark, 1993), and is also used by the person asking the question as a cue to infer the respondent's knowledge (Brennan & Williams, 1995). It is therefore plausible that the timing of addressees' feedback, specifically how closely their response is coordinated with speakers' utterances, may be informative of addressees' knowledge. Feedback that closely follows speakers' utterances may be interpreted as a sign that information is relevant and addressees are attending.

Feedback that is less-well timed, without regard to when speakers' utterances start or end, may signal the opposite.

As mentioned in the introduction, observational studies report that addressees may say "oh" when information is new, and "mhm" or "I know" when information is old (Heritage, 1984). Along similar lines, it has been suggested that during storytelling addressees give vocal feedback (such as "mhm" or "uh-huh") to signal that they are following the information provided in the telling, and they nod to signal that they agree and affiliate with the storyteller's viewpoint (Stivers, 2008). Bavelas and her colleagues (2000) propose the distinction between 'generic' and 'specific' feedback: While generic feedback is not explicitly related to what speakers are saying and may be appropriate in different contexts (e.g., saying uh-hm or nodding), specific feedback is a direct reaction to what speakers said (e.g., facial displays of concern or amusement). Possibly, generic feedback conveys addressees' general engagement, and specific feedback conveys addressees' informational needs, for example whether the information is new or old to addressees. Addressees' feedback responses are therefore not all equivalent; different types of responses have different meanings.

Alternatively, addressees' feedback may not signal qualitative information per se, but instead is open to multiple interpretations (Fussell & Krauss, 1992). For example, a response like "uh-huh" may be interpreted as a sign of attention, a sign of understanding, or a sign of agreement. Speakers rely on other information in the conversational situation to interpret these signals.

Speakers' prior expectations of addressees may provide such an interpretative framework. Using a conversational task as introduced by Bavelas, Coates, and Johnson

(2000), Kuhlen and Brennan (in press) explicitly teased apart speakers' expectations from addressees' feedback. Speakers learned jokes in the form of brief stories and told them to addressees. Speakers were led to expect either attentive addressees (who would have to retell the jokes later), or distracted addressees (who are working on a second task while listening to the jokes). And addressees were either attentive or distracted by a second task. Half of the times speakers' expectations were not matched with addressees' actual level of attentiveness. Attentive addressees gave more feedback than distracted addressees. Speakers who received more feedback told the jokes with more vivid detail, but only when they had also expected attentive addressees. Speakers who received less feedback put less time into the task, but only when they had expected distracted addressees to be attentive (when their initial expectation did not match the unfolding evidence). These results suggest that feedback cues are interpreted against prior expectations or attributions about a partner.

A similar pattern of partner-specific adaptations was found in speakers' speech-accompanying gestures (Kuhlen, Galati, & Brennan, under review). Independent of adjustments made in speaking, speakers gestured more frequently when their expectations were consistent with addressees' feedback behavior. Moreover, speakers used more gestures that were produced in the body's periphery when narrating to attentive addressees whom they had also expected to be attentive, supporting the idea that attentive feedback is only interpreted as a sign of attention when speakers' expectation substantiates this interpretation.

Speakers' expectations may also play an active role in *eliciting* feedback from addressees. Addressees' feedback behavior is closely coordinated with speakers'

behavior, most notably speakers' gaze. Although addressees typically look more often at speakers than the reverse (Kendon, 1967), speakers typically look at addressees during key points in their utterances, thereby creating a brief period of mutual gaze (Goodwin, 1981). During these "gaze windows" addressees are more likely to give feedback (Bavelas, Coates, & Johnson, 2002). Speakers' gaze may therefore be instrumental in eliciting a response from addressees. In this way speakers may be actively creating (or inhibiting) opportunities for addressees to give feedback, thereby providing a possible mechanism underlying phenomenon like behavioral confirmation, which lead to a confirmation of speakers' expectations.

Integrating top-down with bottom-up cues

In summary, both speakers' prior expectations and addressees' feedback behavior are likely to influence how speakers assess addressees' informational needs. But little is known about how these two sources of information are integrated. One possibility is that speakers primarily rely on their expectations about addressees' needs and either ignore or misinterpret evidence to the contrary. Another possibility is that speakers quickly discard their prior expectations and primarily adapt their utterances on the basis of addressees' feedback. Based on previous literature, it is likely that these two cues are integrated in a dynamic process. Such an integration could take different forms: Speakers may gradually update their initial expectations based on incoming feedback cues. But speakers may even shape addressees' feedback to match their expectations.

Many studies investigating the role of addressee feedback do not explicitly assess speakers' prior expectations. Presumably speakers either don't have prior explicit expectations about their addressees, or any expectations are assumed to converge with

addressees' actual behavior. Likewise, other studies manipulate speakers' expectations of addressees, but don't assess to what extend addressees' feedback confirms or contradicts these expectations. This makes it difficult to assess whether partner-specific adjustments are due to prior expectations about the partner (top-down cues), or due to the feedback addressees provide during the interaction (bottom-up cues). In order to distinguish possibly distinct contributions of these two sources of information an experimental design needs to tease apart information that is available prior to the interaction from information that becomes available during the interaction.

2 Implications: Confederates may shape speaking, too

Understanding how conversational partners adapt to each other is interesting not only for a theoretical understanding of how dialogue works. It also informs and shapes how researchers go about studying dialogue. The experimental situation that researchers create to test their hypotheses is often guided, implicitly or explicitly, by their understanding of what the essence of dialogue is. Without intending it, this may affect the behavior the study elicits. The use of confederates in studies of dialogue exemplifies this process.

Confederates as conversational partners

In experimental studies of dialogue, conversational partners are often not naïve participants but instead experimental confederates. For example, naïve research participants may think they are interacting via intercom with a fellow participant in another room, when in fact, they are responding to prerecorded utterances of an experimental confederate (e.g., Barr & Keysar, 2002). Or naïve research participants may think they are instructing a fellow participant, when, in fact, the other is a confederate who, for example, records whether participants' instructions match a target sentence (e.g., Bernolet, Hartsuiker, & Pickering, 2007).

Confederates serve as "accomplices" of the experimenter and can play the role of speaker or addressee (or both) to the naïve participant in the experiment. Confederates may receive detailed instructions on how to behave in the experimental setting and their verbal behavior may not be spontaneous, but predetermined by scripts. The same confederate may be used in all experimental sessions, which means that the confederate goes through the same experimental procedures repeatedly. The point of using

confederates is to elicit a certain behavior in the naïve participant by providing the appropriate conversational context, thereby keeping their own behavior standardized and comparable across experimental sessions. In this way, confederates serve as a sort of stimulus or independent variable with the naïve participant as the target or dependent variable.

There are different reasons why researchers decide to use a confederate (for a more complete discussion see Kuhlen & Brennan, 2008). For some, using confederates as conversational partners in studies of dialogue is an attempt to improve upon studying language processing entirely out of a social context. The presence of a conversational partner promises a more life-like experimental setting. The tacit assumption seems to be that addressees mainly serve the function of providing a projection space for the speakers' utterances and do not actively contribute to the interaction. From this perspective, confederates or experimenters themselves can easily fill this role by just being present along with the participant in the experimental situation.

The introduction of confederates to studies of dialogue helps address a long-standing challenge psycholinguists face when studying dialogue: the challenge of experimental control. Just as Bock (1996) identified the problem of "exuberant responsing" in studies of language production, studies of dialogue face the problem of contingent responding: Dialogue consists of a sequence of behavior that is not readily predictable from one person alone. In other words, if speakers' behavior depends on addressees' behavior, and vice versa, how can we gain experimental control of the conversation? By keeping one participants' behavior constant and comparable across

experimental conditions, confederates seem like an ideal way to establish a certain level of experimental control while maintaining a dialogue setting.

Not all researchers agree that confederates always do a good job of replacing conversational partners. In fact, doubts have been raised as to the validity of studies that use confederates instead of naïve conversational partners (e.g., Bavelas, Gerwing, Sutton, & Prevost, 2008; Kuhlen & Brennan, 2008; Lockridge & Brennan, 2002; Tanenhaus & Brown-Schmidt, 2007). The main concern is that confederates' behavior may differ from naïve participants' behavior, thereby systematically changing the nature of the interaction and the behavior under study.

As I argued earlier, one core characteristic of dialogue is that conversational partners assess each other's informational needs and adapt accordingly. Confederates, however, typically have an informational advantage. This may result from their insights into the experimental procedures, or simply from repeated experience with the experimental task. When confederates' knowledge does not match the knowledge associated with their conversational role, confederates may elicit behavior from their conversational partners that is not representative of behavior displayed under more natural circumstances. This is especially problematic when confederates are addressees, as this conversational role typically affords receiving information from a speaker who is more knowledgeable[2].

Based on how speakers assess addressees' informational needs, there are two main concerns with using confederates as addressees: (1) Speakers' expectations about

[2] Not in all conversational settings are speakers more knowledgeable than addressees. For example in a question-answering scenario, the person answering is commonly expected to be more knowledgeable than the person asking.

their addressees' status as confederate may guide how they assess addressees' informational needs. This subsequently may inform how speakers behave in the interaction, or how they interpret addressees' behavior in the interaction. (2) Confederates (who may have no real informational needs in the experimental task) may systematically differ from naïve participants (who have informational needs) with regard to feedback behavior. Speakers may subsequently perceive informational needs of confederate addressees differently than those of naïve addressees', and adapt speaking accordingly.

Top-down cues: When speakers expect their addressees to be confederates

One common concern with the use of confederates is that the success of an experimental manipulation through the confederate seems less likely if the naïve participant is aware of the confederate's role as accomplice of the experimenter (e.g., Martin, 1970, 1973). In fact, research participants who become skeptical about the confederate's true role in the experiment may behave differently from research participants who are undeceived or do not know of the deception (Stricker, Messick, & Jackson, 1969). Participants who know they are interacting with a confederate might become apprehensive about being evaluated (Rosenberg, 1965), or behave in accordance with what they think is expected of them (Bruehl, 1970; Orne, 1962). For these reasons, experimenters often lead participants to believe they are interacting with another naïve partner.

In studies of dialogue, the confederate's true role is sometimes concealed at great lengths. Confederates are often chosen to be from a similar population as the participants in terms of their age (e.g., Schoonbaert, Hartsuiker, Pickering, 2007), and they are trained

to pretend to be regular participants. Before the actual experiment, elaborate pre-experimental encounters with confederate and naïve participants are staged to ensure that participants believe the confederate is a naïve participant. For example, confederates will deliberately arrive late to the experimental session (e.g., Barr & Keysar, 2002), or the experimenter will pretend to make an explicit effort to learn the confederate's name (e.g., Branigan, Pickering, McLean, & Cleland, 2007). Although the research question usually predetermines the roles confederate and naïve participant play in the experimental session (e.g., being either the speaker or the addressee) the experimenter may pretend to assign these roles randomly (e.g., Keysar, Barr, Balin, & Paek, 1998; Keysar, Barr, Balin, & Brauner, 2000). In order not to appear unusually knowledgeable about the experimental task, the confederate may ask for clarification on the instructions (e.g., Branigan, Pickering, McLean, & Cleland, 2007), or display signs of uncertainty during the experimental session, deliberately interjecting utterances with hesitations and even making occasional errors (e.g., Branigan, Pickering, McLean, & Cleland, 2007; Keysar, Barr, Balin, & Brauner, 2000).

But keeping a confederate's status hidden should not be a goal in itself. In fact, it has been suggested that naïve participants behave no differently when they think they are interacting with confederates or naïve participants (e.g., Barr & Keysar, 2002). When studying dialogue, and especially when studying how speakers adapt to their addressees, confederates may become credible not by keeping their confederate status hidden, but more accurately, by ensuring that participants attribute the right kind of knowledge to them. Instead of asking whether speakers change their behavior when they know they are interacting with confederates, the more precise question therefore is whether speakers

will change their behavior based on their expectations to be interacting with a conversational partner who went through the experimental procedure multiple times, and consequently has no real informational needs. And a related question is, whether speakers, who are not aware that they are interacting with confederates, will behave as if they are interacting with naïve addressees, or whether they will pick up on cues in confederates' behavior that reveal their lack of informational needs.

Bottom-up cues: When confederates' feedback reveals their needs

If addressees express their informational needs in feedback, confederate addresses' behavior may differ systematically from naïve addressees' behavior. The type of feedback confederates give may not be under voluntary control. In fact, experimenters rarely seem to make an effort to specify confederate addressee's behavior in any way. Studies using confederate addressees almost never report what kind of instructions or training they gave confederates in preparation for their role. Presumably, their task is merely to "sit and listen"[3].

Consider a study by Brown and Dell (1987) that investigated whether speakers' syntactic choices are influenced by addressees' particular needs. Participants told confederate addressees a series of short stories in which a main character uses either a typical or an atypical instrument to perform a target action (e.g., using a knife or an ice pick to stab someone). While listening to the stories some of the addressees were able to follow the narrations with the help of illustrations; others did not have access to these

[3] If confederate addressees' behavior is controlled, scripts tend to be based on a simplified notion of feedback, such as responding "yes" or "no" to naive participants' instructions (e.g., Horton & Keysar, 1996). Scripting addressees' behavior in more detail may be difficult because it is often nonverbal and to a large degree contingent on what speakers say.

illustrations. Hence, addressees were either aware or unaware of the nature of the instrument. Brown and Dell were interested in whether storytellers would put less emphasis on atypical instruments when addressees knew about the instruments through the illustrations. They found that storytellers mentioned atypical instruments more often and earlier in the sentence than typical instruments. However, whether or not addressees could see pictures of the instrument did not affect how storytellers mentioned atypical instruments. Brown and Dell concluded that storytellers design their utterances depending on what is easiest for themselves and helpful for a generic audience (e.g., mentioning unusual information early): Addressees' specific informational needs (e.g., whether addressees have visual access to certain information) are considered only in a secondary process, as a kind of repair or afterthought.

However, confederate addressees in this study had much more knowledge about the stories than their conversational role would have justified. In fact, throughout the different experimental trials confederates had heard the stories over and over again and most certainly knew them better than storytellers themselves. Throughout the course of the interaction, confederates might have implicitly revealed this through their feedback. Therefore, storytellers might not have adapted to their addressees' needs because, rightly so, they did not perceive them to have any needs.

A study by Lockridge and Brennan (2002), who replicated Brown and Dell's study with naïve addressees, suggests that this may be the case. Storytellers in this study showed a different behavioral pattern: When *naïve* addressees did not have access to information about the instruments through the illustrations, storytellers were more likely to mention atypical instruments, they mentioned them earlier in the sentence, and they

marked them as indefinite. When addressees had access to the illustrations, storytellers were less likely to mention atypical instruments, they mentioned them later in the sentence, or they marked them as definite. In contrast to Brown and Dell, Lockridge and Brennan therefore concluded that speakers do adjust to their addressees' needs early in utterance planning, if their addressees have actual needs.

In summary, speakers' utterances are shaped by their assessment of addressees' informational needs, perhaps even on the level of articulation. This assessment may be based on speakers' own expectations of addressees' needs (top-down cues), addressees' feedback behavior in the interaction (bottom-up cues), or a combination of both. The same kind of assessment process is likely to occur when speakers are interacting with addressees who are experimental confederates. Since confederate addressees often have more knowledge about the conversational task than their role affords, they may elicit adaptations in speakers that are not representative of adaptations elicited by naïve addressees.

3 Research Questions

The goal of this dissertation is to further understand the processes underlying speakers' adaptations to addressees' informational needs. More specifically, the area of inquiry is threefold: (1) Do speakers adapt their utterances to their addressees' needs on different linguistic levels of speech production, (2) Do addressees indicate their informational needs through their feedback, and (3) How do speakers integrate cues from addressees' feedback (*bottom-up cues)* with their own expectations of addressees' needs (*top-down cues*)? The research project therefore unifies an investigation of addressees' behavior on one side and speakers' behavior on the other side. Two different experimental paradigms address these questions, allowing generalization to different conversational situations and different levels of language production.

Do speakers adapt to addressees' needs on different levels of speech production?

If information is perceived as relevant for the addressee, speakers are expected to make this information prominent and accessible. And vice versa, if information is perceived to be irrelevant to addressees, speakers are expected to attenuate this information. Experiment 1 follows up on studies by Bard et al. (2000), in which speakers give addressees directions on how to follow a route marked on a map, and assesses whether speakers adapt their articulation in response to addressees' feedback. Dual process models predict that fast and automatic processes, such as articulation, are impervious to adaptations to the conversational partner (e.g., Bard et al., 2000; Bard & Aylett, 2001). According to this perspective, speakers should not adapt their articulation in response to addressees' needs, as for example communicated through their feedback. In contrast, constraint based models allow partner-specific information to enter all levels

of speech production (e.g., Galati & Brennan, 2010). Under this assumption, speakers are expected to pronounce information more clearly when it appears to be relevant to addressees, and less clearly when it appears irrelevant.

Experiments 2a and 2b follow up on studies by Brown and Dell (1987), and Lockridge and Brennan (2002), in which speakers tell addressees a set of short stories involving typical and atypical instruments. These experiments assess to what extent speakers adapt their syntactic choice to the specific needs of their addressees and whether addressees who are overly familiar with stories elicit similar reactions in speakers as confederate addressees. The two previous studies predict different outcomes: While one study suggests that speakers adapt only to a generic addressee's needs (those that any addressee would have), but fail to take into account addressees' specific needs (Dell & Brown, 1987), the other study proposes that speakers take their addressees' specific informational needs into account (Lockridge & Brennan, 2002). If speakers adapt their utterances to addressees' specific needs, they are expected to package information early in the sentence when it appears to be relevant to their addressee, and vice versa, to package this information later in the sentence when it appears to be irrelevant to their addressee.

Do addressees indicate their informational needs through their feedback?

Addressees' informational needs were manipulated by making the information that speakers give them either relevant (high needs) or irrelevant (low needs). In Experiment 1 this was achieved by having addressees depend on speakers' instructions to follow a route (high needs), or by giving them visual access to the routes speakers were describing (low needs). Experiment 2 manipulated addressees' needs by having

addressees listen to the same set of stories once (high needs), twice (low needs, Experiment 2a), or four times (very low needs, Experiment 2b). Both experiments manipulated addressees' informational needs within subject, so all addressees were observed in situations when they had high needs, and when they had low needs.

If addressees' feedback is indicative of their informational needs, addressees are expected to adapt their feedback behavior when their needs change. One way in which addressees may adapt is how frequently they give feedback. Two possible outcomes would imply a slightly different function of feedback: Feedback may serve as a confirmation that addressees have sufficiently understood what speakers are saying. In this case addressees give more feedback when they have low informational needs, because, from the addressees' perspective, everything speakers say is already sufficiently established (*Feedback for Grounding Hypothesis)*. Moreover, how much feedback addressees give may also depend on the conversational task: When the task requires a close coordination between conversational partners, for example, when giving and receiving directions, addressees will need to give more feedback than when the task requires less coordination, for example when telling and listening to stories. From a grounding perspective, addressee feedback enables speakers to tailor their utterances more efficiently and more specifically to addressees' needs.

Alternatively, feedback may serve as an indicator for addressees' level of engagement. Higher informational needs would therefore be associated with more feedback (*Feedback for Motivation Hypothesis)*. From this perspective, addressee feedback signals that they are appreciating what speakers are saying, which in return would motivate speakers to give more information.

Other ways in which addressees may adapt feedback to express their informational needs can be found in qualitative aspects of their response. Experiments 2a and 2b therefore explore when and what kind of feedback addressees give. Addressees may acknowledge certain information by responding directly in reaction to this information, while their response may be less well timed if the information is not noteworthy. And similarly, feedback responses may be more specific reactions to speakers' utterances when information is noteworthy (e.g., such as laughing or making a funny face), while feedback response may be more generic when this information is already known (e.g., nodding).

How do speakers integrate top-down and bottom-up cues?

If addressees' feedback is indicative of their informational needs, speakers may use this cue and adapt what they say in response. Aside from addressee feedback speakers may base their assessment of addressees' needs on their expectations about addressees' needs based on information available prior to the interaction. Without taking speakers' expectations into account, it is difficult to tease apart whether speakers' behavior is a reaction to addressees' feedback, or a reaction to their own expectations about what addressees' needs are.

Therefore, in addition to manipulating addressees' actual needs, both Experiment 1 and Experiment 2 also manipulate speakers' expectations of addressees' needs. Experiment 1 manipulated speakers' expectations through previous experiences interacting with the same conversational partner. The assumption is that speakers' initial experience with addressees' informational needs will guide their expectations of addressees' needs in later interactions. Experiment 2a and 2b shape speakers'

expectations more explicitly by giving speakers specific information about whether addressees have heard story before. In summary, when assessing addressees' needs, speakers in both experimental paradigms have two different cues available: Cues guided top-down based on speakers' prior expectations about addressees' needs and cues guided bottom-up based on addressees' feedback in the conversational situation.

If speakers base the assessment of their addressees' needs primarily on their *expectations* about addressees' needs, they will ignore incoming cues based on addressee feedback or interpret addressee feedback based on their expectations (*Expectation Only Hypothesis)*. However, if speakers base their assessment of addressees' needs primarily on cues available in addressees' feedback, speakers are expected to adapt what they say in response to addressees' feedback (*Feedback Only Hypothesis)*. This effect should hold even for those conditions in which speakers initially expect addressees to have different informational needs than the addressee actually have.

Previous studies suggest that neither speakers' expectation alone nor addressees' feedback alone are sufficient to determine speakers' adaptations (e.g., Brennan, 1991; Kuhlen & Brennan, in press; Norlund et al., 2009). Instead, speakers appear to integrate top-down and bottom-up cues. There are several ways for such integration. Speakers may revise their expectations on the basis of incoming feedback (*Feedback Modifies Expectation Hypothesis)*. In this case, speakers may initially design their utterances based on what they expected addressees' needs to be, while later in the interaction they design their utterances based on addressees' feedback. A second possibility is that speakers' prior expectations impact addressees' feedback behavior, prompting addressees to behave in accordance with speakers' expectations (*Expectation Modifies Feedback Hypothesis)*.

This would be in line with studies in social psychology showing that a person's expectations are instrumental in eliciting behavior that confirms these expectations (see e.g., Snyder & Klein, 2005; Snyder & Stukas, 1999; Snyder, 1984; Rosenthal & Jacobson, 1968, 1992).

Next to gaining a better theoretical understanding about how speakers assess and adapt to addressees' perceived informational needs, this research also has implications for the use of confederate addressees in studies of dialogue. Since confederates typically go through the same experimental procedure again and again, their informational needs are presumably very low. To make things worse, confederate addressee's behavior is rarely scripted, well- defined, or characterized by published reports. It is therefore plausible that confederates' feedback reveals their low informational needs and triggers speakers to adapt their utterances accordingly.

In summary, this dissertation addresses questions concerning speakers' behavior on the one side, and addressees' behavior on the other side. Of focal interest are therefore how seemingly individual processes of speech production and speech comprehension are closely coordinated in a process that constitutes dialogue.

4 Experiment 1: Addressees shape speakers' articulation

Experiment 1 investigates whether speakers adapt the clarity of their referring expressions based on their addressees' informational needs as communicated through their verbal and nonverbal feedback behavior. The general research questions as outlined in Chapter 3 are addressed with particular focus on, (a) establishing whether addressees express their informational needs through the amount of feedback they give, (b) investigating whether speakers adapt how clearly they articulate referring expressions on the basis of this feedback, and (c) testing how flexibly speakers adapt to a change in addressees' feedback behavior, thereby integrating previous experiences (top-down cues) with current experiences (bottom-up cues) with the same conversational partner.

These questions were addressed building upon the experimental task used by Bard and colleagues (Bard et al., 2000; Bard & Aylett, 2001; see also e.g., Brown, Anderson, Yule, & Shillcock, 1984) in which one participant, the director, gives another participant, the matcher, directions on how to follow a route marked into a map. Maps depict labeled landmarks alongside the route (e.g., "camera shop", "white mountain", etc.), which are referenced when describing the route. Since landmarks are labeled, directors' referring expressions are typically lexically identical, allowing comparisons across conversational dyads. In addition, landmarks (and thus referring expressions) can be carefully selected based on particular phonological characteristics. The Map Task therefore elicits referring expressions that are ideal for assessing articulatory measures, such as word duration or word clarity (e.g. Anderson, Bard, Sotillo, Newlands, & Doherty-Sneddon, 1997; Bard et al., 2000; Bard & Aylett, 2001; Howarth & Anderson, 2009). The Map Task is also conducive for investigating addressee feedback responses. Presumably matchers give

confirmation that they are following the route on a regular basis (most likely at every landmark mentioned by the director).

Matchers' informational needs were manipulated by giving matchers two types of maps: Some maps only depicted the landmarks and matchers had to draw the route into their map based on directors' instructions (high informational needs condition). In others, unbeknownst to the directors, the route was already marked in the map and matchers only had to re-trace it (low informational needs condition). Thus, for some maps matchers really needed the information directors were providing in order to complete the task; and for others, directors' instructions were superfluous. To investigate how previous experiences with matchers' needs shaped directors' assessment of matchers' needs in later interactions, the order of these two conditions was counterbalanced: Half of the matchers initially had high informational needs and then, half-way through the experiment, switched to having low informational needs ("high needs first" group). And the other half of the matchers initially had low informational needs and then switched to having high informational needs ("low needs first" group). Directors' adaptations to matchers' needs were assessed by two articulatory measures: word duration and a rating of word clarity of references to target landmarks.

Hypotheses

Matchers were expected to mark their informational needs through the frequency of their feedback responses. If feedback indicates that something has been understood sufficiently (*Feedback for Grounding Hypothesis)*, matchers should give more feedback when they already have the route marked into their map (low needs), and less feedback when they need to draw the route (high needs). However, if feedback serves the purpose

of motivating speakers to give more information (*Feedback for Motivation Hypothesis*), matchers should give more feedback when they actually have to draw the route (high needs), and less feedback when they already have the route marked into their map (low needs).

If directors use matchers' feedback to assess their informational needs and adjust their articulation accordingly, directors are expected to reduce landmark references in length and clarity when matchers appear to have low informational needs. This would support a *constraint-based model* of speech production that allows partner-specific information to have impact at any point in processing. If directors however do not take matchers' feedback into account, or directors do not adapt articulation on the basis of this feedback, landmark references are expected to be comparable in length and clarity across all experimental conditions. This would match predictions of a *dual-process model* of speech production, which predicts that speakers do not take addressees into account in their articulation.

Since directors are unaware of the two types of maps manipulating matchers' informational needs, directors are assumed to expect matchers' needs to remain stable throughout the experiment. However, contrary to this expectation, matchers' needs change half-way through the experiment. If directors' behavior is guided primarily by their expectations of matchers' needs (*Expectation Only Hypothesis)* their articulation of landmarks will not systematically vary over the course of the experiment. In this case, directors will assess matchers' needs during initial trials and then base their behavior in later trials on this prior experience: Directors who initially encounter matchers with high needs ("high needs first") will therefore articulate landmarks longer and more clearly

even when matchers' needs change. And directors who initially encounter matchers with low needs ("low needs first") will, despite a change in matchers' needs, articulate less clearly throughout all trials.

If, however, directors rapidly discard their prior expectations and adapt to addressees' needs as expressed through their feedback (*Feedback Only Hypothesis)*, directors should modify the clarity of their referring expressions based on incoming feedback cues: When matchers appear to have high needs landmark references will be longer and more clear; when matchers' feedback changes (presumably because their needs change) speakers should flexibly adapt their articulation, independent of their previous experience.

Finally, directors may integrate their prior expectation with matchers' current feedback. If directors revise their expectations of matchers' needs based on currently available feedback about matchers' actual needs (*Feedback Modifies Expectation Hypothesis*), directors' overall behavior should represent a combination between their expectations and matcher's feedback. However, if directors' expectations shape matchers' feedback behavior (*Expectations Modifies Feedback*), matchers should continue to give a comparable amount of feedback, even if their informational needs have changed.

Method

Design

Directors guided matchers through two practice maps, followed by five additional maps. The practice maps were intended to familiarize pairs with the task, and reflected the (high) informational needs that are typical in real-world task-oriented conversation.

Then, for the next two maps in the sequence, half of the matchers had low informational needs, and the other half had high informational needs. The order of matcher's level of informational needs was counterbalanced, so after the practice maps, matchers experienced either (a) two maps with high informational needs, followed by two maps with low informational needs, or (b) two maps with low informational needs, followed by two maps with high informational needs. In order to end the experimental session with realistic social interaction (in which directors' instructions were actually relevant for matchers), matchers always had high informational needs during the last (seventh) map.

In order to test for effects of high and low informational needs as well as for directors' sensitivity to matchers' changing needs, data from two of the maps in this sequence were of particular interest. To allow participants time to detect and adapt to a change of informational needs, only the second map following a change of informational needs was analyzed. Therefore, the analysis focused on conversations from the second and fourth maps after the practice trials (Target Map 1 and Target Map 2, respectively).

The experimental session was videotaped by two miniDV ZR960 Canon camcorders (one recording the matcher, the other the director). In addition, director's speech was recorded in digitial audio format with a Marantz Professional PMD660 Portable Solid State Recorder connected to a Shure dynamic microphone mounted on a stand to the side of the director.

Participants

Thirty-four native speakers of English[4] participated in the study, forming 17 dyads. One experimental session was replaced because the matcher revealed the nature of the experimental manipulation to the director. Of the remaining 16 dyads, eight experienced the experimental conditions in the order high needs then low needs ("high needs first"); the other eight dyads experienced the reversed order ("low needs first"). Gender was equally distributed across conversational roles, yielding eight male and eight female directors, and eight male and eight female matchers. In all dyads, participants were unacquainted prior to the study. Participants were compensated with either research credit or $9 payment.

Materials

Seven maps were created on the basis of the maps provided by the HCRC Map Task Corpus (Anderson et al., 1991). Original maps were adapted such that each map contained ten main landmarks labeled with a specific term of reference (e.g., "gurgling brook"). All main landmarks were compound words and appeared only once in the entire set of maps. In addition, each map had five extra landmarks without a label that could be referenced more freely (e.g., the picture of a noose was referenced as "the noose", "the gallows", or "the hangman thing"). While main landmarks were essential for describing the route, extra landmarks were optional in that they provided additional but not crucial information on how to follow the route (see Appendix 1 for Target Maps).

The two target maps' main landmarks were selected on the basis of phonological characteristics aimed to optimize the two articulatory measures, word duration and

[4] All participants indicated in a prescreening questionnaire that they were fluent and native speakers of English. Some participants were fluent in more than one language.

ratings of word clarity. To make word duration measurements more precise, each target map contained six landmarks that started and ended with a stop consonant (e.g., "camera shop") that yielded a distinct pattern of speech on- and offsets, identifiable through a sudden, sharp burst of energy in the spectrogram, or an abrupt decrease in energy, respectively.

For rating word clarity, seven landmarks in each target map were selected for their potential to be sensitive to variability in clarity of expression. These landmarks either had the possibility for /t/ or /d/ reduction (e.g., in "cobbled street" the /d/ can be unreleased or even dropped), or the possibility for place assimilation (e.g., in "white mountain" the coronal or alveolar /t/ in "white" can be assimilated to the subsequent bilabial /m/ in "mountain" resulting in a bilabial /p/ sound).

The two target maps were made as similar to each other as possible with respect to the phonological characteristics of the landmarks[5]. Table 1 provides a list of all main landmarks, specifying their characteristics and whether they contributed to word duration or word clarity ratings.

Procedure

Matchers were scheduled to arrive 15 min prior to directors. Upon arrival, they were told that this experiment investigated how people give and receive directions and that their partner (another volunteer from the subject pool who would arrive shortly) would be giving them directions on how to follow a route marked on the director's map.

[5] Target Map 1 and Target Map 2 did not systematically differ in the average duration of target landmark references, the number of words directors used to explain the route, the duration of the interaction, or the amount of feedback they elicited in matchers. After excluding one outlier, Target Map 1 and Target Map 2 also did not differ in average word clarity rating (see upcoming material for details).

They were shown example maps (both the director's and the matcher's version) and they were instructed that their task would be to draw the route into their own map according to their partner's instructions. Matchers were then informed that out of the six maps two would serve as control trials, measuring their baseline behavior. These maps came with the route already marked. They were asked not to reveal this to their partners. Whenever they encountered such a baseline map they were instructed to simply re-trace the route following their partner's instructions.

When directors arrived shortly thereafter they were greeted and instructed separately from matchers. Parallel to matchers' instructions, directors were informed that this experiment explored how people give and receive directions. Viewing the example maps, directors were told that their task was to give their partners instructions on how to get from the start point to the finish point, as marked on their maps. The experimenter then gave an example of possible instructions by describing the first part of the route marked on the example map.

Before the actual experimental task participants became acquainted with each other by playing a round of Taboo, a popular word guessing game. In the presence of the experimenter, participants then practiced the experimental task with the example map they had received instructions on. The experimenter then left the room and the participants worked through six maps on their own. After participants had completed the last map, they were given a questionnaire probing their observations and possible hypotheses they may have formed about the experiment. Participants were then fully debriefed about the experiment's true nature and hypotheses.

Measures

Matcher feedback. Matchers' feedback responses during Target Map 1 and Target Map 2 were transcribed by one coder (a research assistant) with help of the video recordings. Feedback was defined as verbal and/or nonverbal responses on the part of matchers that indicated they were attending, following, or reacting to what directors were saying. Included in the analysis were verbal contributions to the interaction, as well as head nods and vocalizations such as *yeah*, *mhm*, *uh*, laughter, and gazes to the director[6]. Different types of feedback responses that occurred at the same time and appeared to convey the same meaning (e.g., head nodding while saying *yeah*) were coded as one feedback response. The total number of feedback responses was normalized by the total number of words used by the director, specifying the number of feedback responses that occurred per 100 director words.

Reliability of the coder in detecting feedback responses was established while coding a subset of the data obtained from Experiment 2a (see Experiment 2a for more details). This coder agreed with the author's coding decisions in 98% of the maximum number of 452 detected feedback responses, and agreed with a third coder 83% of the time.

Directors' instructions (general measures). Directors' instructions for both target maps were transcribed from the audio recording. The duration of each route description (one map) was recorded in seconds. The beginning marked directors' first utterance, the

[6] Matchers mainly studied their maps, but occasionally glanced up at directors. In line with Anderson, Bard, Sotillo, Newlands, & Doherty-Sneddon (1997), directors were assumed to interpret matchers' gaze similarly as they would interpret overt feedback.

end matchers' acceptance of the final landmark. In addition, the number of words directors used for describing each route was counted. This word count excluded words that were interrupted and followed by speech repairs as well as fillers such as "uh" or "um". Contracted forms such as *don't* and *gonna* counted as one word. To estimate the amount of extra details directors' instructions provided, the number of extra landmarks that were mentioned was recorded. Multiple references to the same landmark counted as one reference; the maximum number of possible references was five.

Word duration of target landmarks. For all 16 dyads and both target maps, the duration of the first reference to target landmarks was measured in milliseconds. In each of the two target maps six target landmarks were designated for measuring duration (see Table 1), yielding 192 possible observations. Excluded from the analysis were 11 references that did not realize all the words of the full citation form, and seven references that contained speech disfluencies or repairs. In addition, four references were excluded because they were longer than three standard deviations from the mean duration of the respective landmark due to extended pauses within the compound words. Finally, three references were excluded because the matcher had mentioned them prior to the director's first reference. These criteria yielded a set of 167 tokens for analysis.

Selected target landmarks were excerpted from digitally recorded materials using Praat (Boersma & Weenink, 2006), a software tool for acoustic analysis. The initial word boundary was placed on the onset of the burst (plosive) of the first stop consonant (e.g., "c" in camera shop); the final word boundary was placed at the release of the burst of the final stop consonant (e.g., "p" in camera shop).

Word clarity rating of target landmarks. As a second articulatory measure, clarity of referring expressions was assessed through a separate word clarity rating task. For this task 11 native speakers of English (4 males, 7 females) who had not participated in the main study rated directors' references to landmarks for clarity of expression. Raters received research credit for their participation.

Two lexically identical tokens of each of the 14 target landmarks designated for rating clarity of expression (see Table 1) were excised from the original interactions blind to the experimental condition. One stemmed from the low needs condition, the other from the high needs condition. Each director contributed two tokens in total, one from each target map (and consequently, one from each informational needs condition). Since there were more directors than tokens needed, two directors, both male, were randomly excluded (one from the group "High Needs First", and one from the group "Low Needs First"). Although one director contributed two tokens, the lexically identical counterparts they were compared to stemmed from two different directors (e.g., director A's token 1 was compared to director B's token 1; director A's token 2 was compared to director C's token 2).

Selected tokens consisted of speakers' first references to the target landmark and contained no speech disfluencies, overlapping speech, or elongated pauses in between compound words. In addition, selected tokens had been produced in a comparable phonological environment: The landmark reference was preceded by the definite article "the" and followed by a pause (e.g., "the gurgling brook ((pause))"). Tokens were excised from their original environment with a pause of approximately .05 seconds preceding and following each expression.

Excised tokens were distributed to two different rating lists of 14 word pairs each (28 tokens per list in total). The two lists varied the order in which tokens were presented (both within a word pair and between word pairs). Five raters listened to one list, and six raters listened to the other[7].

Word lists were presented to raters through an iTunes playlist. Raters were instructed to click and play each token in the order it was numbered. After listening to one word pair, they rated on a spreadsheet the clarity of the two words on a scale 1-5 (1= very unclear/ unintelligible, 5= very clear/ intelligible). If raters thought the two tokens were equal in clarity, they could give them the same value. Raters could play the tokens as often as they needed to make their rating decision. After rating a word pair, they typed out the word they had heard. After doing one practice trial with the Experimenter raters went through the rest of the speech files on their own and at their own pace.

Initial analysis of the clarity ratings yielded a significant difference between the two target maps across raters that was independent of any influence of matchers' informational needs (or the order in which needs changed), $F\,(1, 10) = 15.70$, $p < .01$. This indicates that, overall, landmarks in Target Map 1 were rated less clearly articulated than landmarks in Target Map 2. A comparison of mean ratings for each landmark individually revealed that one particular landmark in Target Map 1 was driving this difference (see Figure 1): The landmark "lost steps" was rated on average 2.50 (SD = 1.01), deviating more than one standard deviation from the average rating 3.73 (SD = 1.07). While landmarks in the set were correctly named, on average by 10.6 out of 11 raters, this landmark was correctly named by only five out of 11 raters. It was

[7] Ratings did not differ significantly across the two lists, t(9)= .59, p= .57, so the list factor was combined for all subsequent analyses.

subsequently treated as an outlier and excluded from the final analysis of the clarity rating measure, making the two target maps comparable in rated word clarity, F (1, 10) = .13, p = .73.

Statistical analyses

First, data from Target Map 1 and Target Map 2 were analyzed separately. Target Map 1 represents a point at which half of the matchers continued to have the usual high informational needs, while the other half have had to adjust to a change to low informational needs. Target Map 2 represents a point at which all pairs had been exposed to the matchers' changing needs. For each of these two analyses, situations in which matchers had high informational needs were compared to those in which matchers had low informational using independent sample t-tests for the variables of: frequency of matcher feedback, word duration of target landmarks, number of extra landmarks mentioned, duration of interaction, and total number of words used by director. Word clarity ratings for target landmarks originating from the high needs condition were compared to those originating from the low needs condition, using paired sample t-tests.

Second, Target Maps 1 and 2 were entered jointly into a 2x2 ANOVA in order to test for order effects. $F1$ or $t1$ is the analysis by-subjects, which averages values across all sixteen dyads (word duration), or eleven raters (word clarity rating), and $F2$ or $t2$ is the analysis by-items, which averages values across all twelve landmarks (word duration), or thirteen landmarks (word clarity).

Results

Impact of matchers' informational needs (Target Map 1)

Matchers gave marginally more feedback when they had high informational needs than when they had low informational needs, t (14) = 1.98, p = .07 (see Figure 2). And directors adapted their articulation of target landmarks accordingly: When interacting with matchers who had high needs (who were more likely to give feedback), target landmarks tended to be longer in duration[8], $t1$ (14) = -2.11, p = .05; $t2$ (10) = -1.33, p = .21 (see Figure 3), and they were also rated as being more clearly articulated, t (10)= 3.85, $p < .01$; $t2$ (4) = 1.18, p = .31 (see Figure 4). Directors mentioned marginally fewer extra landmarks when interacting with matchers who gave more feedback, t (14) = 1.67, p = .12. In the discussion, I will argue that this made directors' instructions more efficient. Interactions with different levels of matcher needs did not differ with respect to their duration, t (14) = -1.28, p = .22, or the number of words directors used, t (14) = -.68, p = .51 (see Table 2 for means and standard deviations of last three measures).

Adapting to a change in matchers' informational needs (Target Map 2)

The effect of informational needs on Target Map 2 was quite different: During Target Map 2, matchers' did not give more feedback when they had high informational needs, t (14) = -.71, p = .49. Instead, judging by mean values only, matchers tended to give *more* feedback when they had low needs and *less* feedback when they had high

[8] Not all effects that reached statistical significance across participants (by subject) also reached significance across landmarks (by item); and vice versa. Since landmarks always appeared in the same location of the map, this may be attributed to ordering effects (e.g., some landmarks having greater or lesser relevance in relation to the route), or variability within the landmarks (e.g., familiarity or frequency of referring expression). For the current purpose, this section will interpret those effects that are significant across all participants.

needs (see Figure 2). Although statistically not significant, this pattern converges with measures pertaining to directors' behavior: Unlike the pattern found for Target Map 1, during Target Map 2 directors' referring expressions were marginally *longer* in duration when matchers had low needs, $t1$ (14) = 1.8, p = .09; $t2$ (10) = .33, p = .75 (see Figure 3), and they mentioned *fewer* additional landmarks, t (14) = -2.35, p = .03 (see Table 2). Matchers' informational needs had no influence on the word clarity ratings, $t1$ (10) = .42, p = .68; $t2$ (6) = .24, p = .82 (see Figure 4), the duration of the interaction, t (14)= .38, p = .71, or the number of words directors used, t (14) = -.02, p = .99. Note that although directors' behavior does not seem to respond to matchers' actual informational needs (directors did not consistently articulate more clearly and deliberately when matchers had high needs), it does correspond closely to matchers' feedback behavior (directors consistently articulated more clearly and deliberately when matchers gave more feedback).

The combined analysis of Target Maps 1 and 2 shows that, overall, matchers gave marginally more feedback when they had high informational needs, F (1, 14) = 4.08, p = .06. Correspondingly, target landmarks were rated as more clearly articulated, $F1$ (1, 10) = 13.29, $p < .01$; $F2$(1, 11) = .07, p = .79, and the conversations tended to last longer overall, F (1, 14) = 3.27, p = .09. This main effect of matchers' informational needs did not hold for word duration of target landmarks, $F1$ (1, 14) = .1, p = .76; $F2$ (1, 10) = 4.5, p = .06, or number of extra landmarks mentioned, F (1, 14) = .28, p = .60, or number of words used by directors, F (1, 14) = 1.89, p = .19.

However, the overall data pattern suggests that matchers' and directors' behavior was not primarily influenced by matchers' informational needs. Instead, what seems to

have greater impact is whether matchers initially had high or low needs (during Target Map 1). The order in which matchers' needs changed yielded significant effects for directors' duration of referring expression, $F1$ (1, 14) = 6.19, p = .03; $F2$ (1, 10) = 21.27, p < .01, rated word clarity, $F1$ (1,10) = 8.69, p = .02; $F2$ (1, 11) = .02, p = .89, and number of extra landmarks mentioned, F (1, 14) = 5.64, p = .03. This order effect did not reach significance for matchers' feedback, F (1, 14) = 2.01, p = .18. There were no significant interaction effects between matchers' informational needs and the order in which informational needs changed.

In summary, when matchers had high informational needs during Target Map 1, they seem to continue to give more feedback during Target Map 2 (compared to matchers with low needs in Target Map 2), even though their informational needs had changed. Correspondingly, these directors' referring expressions continued to be longer and more clear, and their instructions contained fewer extra details. When matchers had low informational needs during Target Map 1 they continued to give less feedback during Target Map 2, although they now had high informational needs. Directors followed suit, keeping their referring expressions short, less clear, and instructions contained more extra details.

Discussion

Directors adapted their articulation corresponding to matchers' feedback. This pattern is particularly clear for Target Map 1: When matchers gave more feedback, referring expressions were longer in duration and rated as more clearly articulated. Even during Target Map 2, directors' behavior seems to align with matchers' feedback. Results stand in contrast to a dual-process account that proposes partner-specific information

only finds entry to certain processes of speech production. Such an account suggests that processes of articulation are impervious to influence by the conversational partner in general, and specifically to addressees' feedback behavior (e.g., Bard et al. 2000, Bard & Aylett, 2001). Instead, results better fit constraint-based approaches to language production that assume speakers can in principle adapt their utterances on all levels of speech production (see e.g., Brennan & Hanna, 2009; Galati & Brennan, 2010; Hanna, Tanenhaus, & Trueswell, 2003; Hanna & Tanenhaus, 2004). In support of studies suggesting that speakers adapt their articulation to their conversational partners' needs (e.g., Galati & Brennan, 2010; Samuel & Troicki, 1998), these results indicate that addressee feedback is instrumental in shaping speakers' behavior at the articulatory stage.

At first matchers' feedback seemed to be indicative of their informational needs: Matchers who needed to draw the route into the map (and hence needed the information directors were giving) gave more feedback than matchers who merely had to re-trace the route that was already marked into their map. This result seems to support the *Feedback for Motivation Hypothesis*, which proposes that addressee feedback serves the purpose of motivating speakers to give more information. Such an interpretation is in line with findings showing that attentive addressees (who presumably have informational needs) give more feedback than distracted addressees (who presumably don't have informational needs), which in return can prompt speakers to give more information (Bavelas, Coates, Johnson, 2000; Kuhlen & Brennan, in press).

However, in this case directors did not give matchers more information when they received more feedback. To the contrary, when directors received more feedback they mentioned *fewer* additional landmarks. Although this may initially seem mal-adaptive,

the *Feedback for Grounding Hypothesis* can explain this behavior. Additional landmarks were actually superfluous to the task[9]. By mentioning only those landmarks that are necessary for describing the route directors' instructions became more focused and efficient. In line with this interpretation are previous studies showing that, with little or no addressee feedback, speakers use more words to convey the same meaning (Krauss & Weinheimer, 1966; Kraut, Lewis, & Swezey, 1982). Similarly, when directors in the present experiment received little feedback they may have had difficulties finding an appropriate level of information. The functionality of giving additional details may thereby be specific to the conversational task: When giving directions, too much information may be harmful, because it can confuse or overload addressees. However, when telling stories additional information may go hand in hand with a more vivid style of narrating (Bavelas, Coates, & Johnson, 2000; Kuhlen & Brennan, in press).

Matchers' feedback was not *only* influenced by their informational needs. During Target Map 2, matchers' feedback did not reflect their informational needs. Instead, matchers who previously had high needs during Target Map 1 seemed to continue giving more feedback, even though their informational needs had changed. And matchers who had low needs during Target Map 1 continued to give little feedback, even though they now had high needs. Matchers' previous informational needs seemed to be influencing their behavior in later interactions.

Possibly, directors' expectations of matchers' needs shaped how much feedback matchers gave. Such would be predicted by the *Expectations Modify Feedback Hypothesis*, which proposes that a speaker's expectation of their conversational partner

[9] Recall that these landmarks had been positioned so that they were not essential to describing the route.

can have an effect on their partner's behavior. In the current situation, directors' experience of interacting with matchers who initially gave a lot of feedback may have caused them to give matchers more opportunities to give feedback in later interactions. One way speakers can prompt addressees to give feedback is through seeking eye contact (Bavelas, Coates, & Johnson, 2002). A follow-up analysis investigated whether directors gazed at matchers more frequently or longer during Target Map 2 when matchers previously had high informational needs than when matchers previously had low informational needs. However, directors gazed at matchers equally often, t (14) = 1.02, p = 32, and their gazes lasted equally long, t (14) = 1.18, p = .26. In fact, directors seemed to be consistently gazing at matchers more often and longer when they had high informational needs, independent of their previous experiences.

Another explanation why matchers' feedback during Target Map 2 was not indicative of their needs could be that directors and matchers jointly established a certain "modus operandi" on how to accomplish the task. Such a modus operandi may include what level of accuracy is sufficient for the current purpose ("grounding criterion", Clark & Schaefer, 1989; Clark & Wilkes-Gibbs, 1986). Following this rationale, those matchers who started off with low informational needs may have established a lower criterion of accuracy for doing the task and maintained this criterion (and the feedback behavior associated with it) for the following interactions. Subsequently, the accuracy with which these matchers drew the route into the map would have suffered.

Route accuracy can be assessed by superimposing both matchers' and directors' routes on a grid and counting the number of deviating cells (Anderson, Clark, & Mullin, 1991). Greater deviations between matchers' and directors' routes result in a higher

number of cells. If matchers and directors established a certain criterion for accuracy in initial interactions, matchers who initially had low needs (presumably establishing a low criterion of accuracy) should be less accurate in drawing the route on Target Map 2 compared to matchers who had high informational needs in Target Map 1 (presumably establishing a high criterion of accuracy). Indeed, matchers with initial low needs deviated on average 785.88 (*SD* = 268.57) cells from director's route, while matchers with initial high needs deviated on average only 675.75 (*SD* = 170.44) cells, suggesting that matchers with low needs may indeed have established a lower criterion for accuracy. However, this difference did not reach statistical significance, $t\,(14) = -.98$, $p = .34$. Note that, since route accuracy can only be established when matchers actually have to draw a route (hence only in the high needs condition), route accuracy for Target Map 1 (matchers with initial high needs) was compared with route accuracy for Target Map 2 (matchers with initial low needs), confounding possible effects due to different degrees of difficulty between the two maps, practice in following route instructions, or fatigue.

In summary, although directors seemed to adapt to matchers' feedback, directors did not consistently adapt to matchers' informational needs, because matchers' feedback was not consistently indicative of their needs. This pattern of behavior makes it difficult to estimate to what degree directors' assessment of matchers' needs was due to their expectations of addressees' needs (*Expectation Only Hypothesis*), or due to addressees' feedback (*Feedback Only Hypothesis*). Results seem to align best with the idea that directors initially assess matchers' needs based on their feedback behavior, and then later base their adaptations on previous experiences. As the *Expectations Modifies Feedback Hypothesis* would predict, it is possible that directors may be shaping matchers' behavior.

In any case, it is safe to assume that *both* directors' and matchers' behavior are shaped by experiences made in prior interactions.

It is unclear how persistent previous experiences shape the expression and assessment of matchers' informational needs. Note that all dyads initially experienced two maps in which matchers had high informational needs (practice map and filler map 1). Nevertheless, directors who then experienced matchers with low informational needs in the following two maps (filler map 2 and Target Map 1) re-assessed matchers' needs, presumably based on their feedback. In principle, matchers and directors therefore seem to be able to adjust to changes in informational needs. However, their assessment for Target Map 1 was then not re-evaluated during the next two maps (filler map 3 and Target Map 2). Possibly, there is a crucial period in which informational needs are established (e.g., during the first couple of maps), or, sufficient evidence needs to accumulate before informational needs are re-assessed. Follow-up experiments could address these questions by employing a paradigm that allows a more online investigation of how conversational partners incorporate previous experience with incoming evidence.

Concluding, this experiment suggests that addressee feedback provides a strong enough cue to shape speakers' speech production at the articulatory stage. However, addressee feedback does not only seem to be guided by their informational needs. Instead, previous experiences seem to influence both speakers' and addressees' behavior. The role of such top-down information available prior to the conversation will be investigated more explicitly in Experiment 2a and 2b.

5 Experiment 2: Addressees shape speakers' syntactic choices

Experiment 2 addresses the three central research questions laid out in Chapter 3 with the following concentrations: (1) What are the mechanisms by which addressee feedback signals addressees' specific informational needs, (2) Do speakers adapt their syntactic choices to accommodate their addressees' specific needs, (3) How do speakers integrate their initial expectations of addressees' needs with incoming feedback about addressees' actual needs. Experiment 2 extends Experiment 1's investigation of *quantitative* differences in addressee feedback by investigating *qualitative* aspects of addressee feedback. In addition, Experiment 2 investigates speakers' adaptations to addressees on a different level of speech production, namely syntactic constructions. Lastly, Experiment 2 investigates more specifically how speakers' prior expectations may shape their adaptations to a conversational partner's needs.

These questions were addressed building upon an experimental task used by Brown and Dell (1987), and Lockridge and Brennan (2002), in which speakers told addressees short stories that involve a main character performing a certain action using either a typical or atypical instrument (e.g., writing a term paper with a computer or a typewriter). The general finding was that storytellers are more likely to mention, and to put more emphasis on, atypical than typical instruments, introducing them early on in the narrative. However, the two studies had contradictory findings about to what extent speakers take their addressees' knowledge about the instruments into account. Brown and Dell's study found that speakers adapt to addressees in a generic way (in this case, always emphasizing atypical instruments), but fail to adapt to their addressees' specific informational needs (only emphasizing atypical instruments if this information is new to

addressees). In contrast, Lockridge and Brennan found that speakers do adapt to addressees' specific informational needs, stressing atypical instruments when addressees have no other way of knowing about these instruments. Lockridge and Brennan suggested that the discrepancy between the two studies was due to the fact that addressees in Brown and Dell's study were confederates who had heard the stories multiple times and therefore had no real informational needs (and in fact knew the stories better than the speakers telling the stories). This they may have conveyed through their feedback.

Experiment 2 follows up on this idea and investigates whether addressees' feedback indeed reveals their familiarity with the nature of the instruments. In addition, Experiment 2 investigates to what extent storytellers' expectations about addressees' needs contribute to their assessment of addressees' needs. Although Brown and Dell's participants were not informed that their addressee was a confederate (and hence was presumably familiar with the stories), it is relevant to understand to what extent a speaker's expectation about a conversational partner's informational needs plays into their adaptations to those informational needs.

Simulating the experiences of confederate addressees in Brown and Dell's study, the experimental design built upon a setting in which one addressee listens to the same set of stories multiple times, told by different storytellers. Three variables were manipulated: Typicality of instrument, informational needs of addressees, and storytellers' expectations of addressees' informational needs. Typicality of instrument was manipulated within subjects: Addressees were familiarized to one of two sets of stories (List A, List B) featuring a combination of stories with typical and atypical

instruments. Addressees' informational needs were manipulated in three steps: Addressees were exposed to the same set of stories once (high needs, Experiment 2a & 2b), a second time (low needs, Experiment 2a), and a fourth time (very low needs, Experiment 2b).

Storytellers' prior expectations of their addressees' needs were manipulated between subjects by instructing storytellers that either their addressees were hearing the stories for the first time (expectation high needs), or addressees had heard stories before (expectation low needs). Within one triad, storytellers' expectations of addressees' knowledge were either congruent or incongruent with addressees' actual knowledge: In congruent conditions, the first storyteller expected addressees to have high needs and addressees were indeed hearing the stories for the first time; and the last storyteller expected addressees to have low needs and addressees indeed had heard stories before. In incongruent conditions, the first storyteller expected addressees to have low needs, but in fact addressees were hearing the stories for the first time; and the last storyteller expected addressees to have high needs, but in fact addressees had heard stories before. Mismatching speakers' expectations of addressees' needs with addressees' actual needs allows a direct investigation of how these two sources of information individually contribute to speakers' assessment of addressees' needs.

Across triads (one addressee, two storytellers), this yielded four different experimental groups: (1) List A: both storytellers hold congruent expectations, (2) List B: both storytellers hold congruent expectations; (3) List A: both storytellers hold incongruent expectations; (4) List B: both storytellers hold incongruent expectations.

Hypotheses

In terms of the frequency with which addressees give feedback, the *Feedback for Motivation Hypothesis* would predict that addressees give more feedback when they hear the stories for the first time and increasingly less feedback the more often they hear the stories. In contrast, the *Feedback for Grounding Hypothesis* would predict that addressees give more feedback when stories are known, indicating that the stories have been understood sufficiently.

To further understand the mechanisms by which addressees may express their informational needs, their feedback was analyzed not only on quantitative, but also qualitative dimensions. The pattern of speakers' adjustments in Experiment 1 indicated that addressee feedback enables speakers to adjust their utterances more specifically to the informational needs of their addressees. Experiment 2 makes the clear prediction that atypical instruments will be more relevant to addressees than typical instruments. Therefore, addressees are expected to be more likely to give feedback in reaction to atypical than typical instruments, to respond more promptly (directly after the instrument is first mentioned), and with a more specific type of reaction (e.g., by showing signs of amusement or surprise). However, as addressees become familiar with stories, they are expected to be equally likely to give feedback to atypical and typical instruments, and their feedback is expected to be less well coordinated with speakers' utterances (responding before or much later after the instrument is first mentioned) and less specific (e.g., nodding).

If storytellers take addressees' familiarity with instruments into account, the evidence from Lockridge and Brennan's study would predict that speakers will

emphasize atypical instruments only if addressees have not heard stories before. However, if speakers fail to accommodate addressees' specific needs they should always emphasize atypical over typical instruments, independent of addressees' familiarity with the stories, thereby replicating findings of Brown and Dell.

If storytellers base their assessment of addressees' needs on cues available in addressees' feedback (bottom-up cues), storytellers are expected to emphasize atypical instruments less when addressees' feedback implies that they are familiar with stories (*Feedback Only Hypothesis)*. This effect should hold even for those conditions in which storytellers had expected addressees to be unfamiliar with stories. On the other hand, if storytellers base their assessment primarily on their expectations of addressees' needs (top-down cues), speakers should de-emphasize atypical instruments only when they were told that addressees were familiar with stories (*Expectations Only Hypothesis*). Note that if storytellers' expectations dominate storytellers should de-emphasize atypical instruments even if addressees are entirely unfamiliar with stories.

However, if storytellers integrate their prior expectations with cues available in addressees' feedback, neither of these two cues alone should be able to account for storytellers' behavior. For those conditions in which storytellers hold incongruent expectations about addressees' needs two patterns may emerge: According to the *Feedback Modifies Expectations Hypothesis*, initial stories should reflect storytellers' prior expectations of addresses' needs, while stories told later in the interaction should reflect addressees' feedback. According to the *Expectation Modifies Feedback Hypothesis* storytellers' expectations of addressees' needs may shape addressees'

feedback behavior. In this case, addressees' feedback would reflect the expectations with which storytellers enter the interaction.

In addition to seeking a better understanding of how speakers assess and adapt to their addressees' informational needs, Experiment 2 provides a commentary on the practice of using confederates as addressees in experimental studies of dialogue. The assumption is that as addressees become increasingly familiar with the stories their behavior should resemble more and more the behavior of confederates. If confederates can be thought of as addressees with very low informational needs, storytellers should be increasingly less likely to emphasize atypical instruments as addressees become more and more familiar with stories. Note that this is not the pattern of behavior Brown and Dell (1987) found: In their study, storytellers more often emphasized atypical instruments, although addressees were confederates who presumably had very low informational needs. If confederate addressees' feedback is responsible for the pattern of behavior Brown and Dell found, addressees who have heard stories multiple times before should behave differently than regular addressees with low informational needs. Possibly, in an attempt to pretend to have informational needs, these addressees may behave like generic addressees, thus prompting storytellers to adapt to a generic addressee's needs.

5.1 Experiment 2a: Speakers adapt to addressees' needs

Method

Participants

One hundred-twenty volunteers (79 females, 41 males) participated in a total of 40 sessions with three participants each. Participants signed up for either the role as addressee or storyteller, yielding 40 addressees and 80 storytellers. The gender

distribution within triads resulted coincidentally based on the participants' availability. Table 3 lists the number of male and female storytellers and addressees for each condition. All participants were native speakers of English[10]. Seven sessions were excluded and replaced for the following reasons: One storyteller was not a native speaker of English, two storytellers didn't follow the experimental instructions, and three addressees revealed the experimental manipulation at the beginning of the session (telling the storyteller that they had heard stories before), and one storyteller participated in the experiment twice (his second participation was excluded). All participants received either research credit or $9 per hour for their participation.

Materials

Story sets. Twelve stories were selected from the corpus of stories used by Lockridge and Brennan (2002). Each story has two versions: In one, the main character of the story performs a certain action using a typical instrument (e.g., a computer to write a term paper); in the other, the character uses an atypical instrument (e.g., a typewriter). There were four additional distractor stories, interspersed between target stories. These distractor stories were based on historic events or scientific facts (e.g. the demise of Pompeii, the chemical element carbon) and served the purpose of supporting the cover story given to addressees that the purpose of the experiment was to test their memory.

Stories were printed out on separate sheets of paper, held together on a clipboard. Each story was followed by a second sheet of paper showing an illustration of the story (taken from Lockridge & Brennan, 2002) that served storytellers as a memory cue when recalling the story. On the same sheet of paper was a rating scale asking storytellers to

[10] All participants indicated at the time they volunteered that they were fluent, native speakers of English. Some reported that they were bilingual.

assess, after they had recalled the story, how well they believed their partners had understood the story. This rating was intended to bring addressees' behavior into storytellers' awareness. See Appendix 2 for a complete list of stories, in order of narration, target instruments and target actions underlined.

Exit questionnaires. Written exit questionnaires assessed participants' subjective experiences in the experiment. Specifically, storytellers' questionnaires assessed their beliefs about whether their addressees had heard the stories before or not, and whether their addressees were confederates.

Procedure

For each triad, participants who had signed up to be addressees arrived 15 minutes prior to the first storyteller. The second storyteller was scheduled to arrive 45 minutes after the first storyteller, at a point when the first storyteller had already finished the experiment and left. Upon arrival addressees were informed that this experiment investigated how memory improves upon repeated exposure to the study material[11]. They were told that first one storyteller would tell them a set of stories, then a second storyteller would tell them the same set of stories for a second time, and then there would be a memory test. Addressees were encouraged to interact freely with storytellers.

Storytellers were instructed separately from addresses. Upon arrival, first storytellers were informed that the experiment investigated storytelling. Depending on the experimental condition, storytellers were told that either (a) their addressee had arrived early and was waiting for them in the next room (*Expectation: Stories unknown*),

[11] This cover story gave addressees a reason for the otherwise unusual experience of listening to the same stories multiple times. Although this may have encouraged addressees to listen attentively even when familiar with stories, such behavior would have worked against my hypothesis.

or (b), because the subject pool was small this semester, addressees were "re-used" and their addressee had already participated in the study just previously (*Expectation: Stories known*).

Showing an example story, the experimenter instructed storytellers to study the story silently until they felt ready to retell it, then to flip the page, and retell the story from memory. Participants were instructed not to repeat verbatim what they had read, but instead to retell the story in their own words. After telling the story, they were to rate how well they thought their addressees had understood, judging from their reactions to the story.

In the experimental room, storytellers and addressees became acquainted with each other by playing one round of the word guessing game Taboo. This step was taken to make participants feel comfortable interacting with each other, and acclimated them to the experimental room and the cameras. With the experimenter present, storytellers practiced one narration with the example story they had received instructions on. The experimenter then left the room and participants went through the rest of the stories on their own.

After narrating all stories, the storyteller was led into a different room where they answered an exit questionnaire probing their impression of the experimental situation. They were then fully debriefed on the nature of the experiment. After the first storyteller left, the second storyteller arrived. Second storytellers received identical instructions to first storytellers, except that their expectations of addressees' knowledge were opposite to those of the first storyteller. In congruent conditions, the first storyteller was led to believe stories were unknown to addressees, and the second storyteller was led to believe

stories were known (in other words, they were accurately informed about the addressees' knowledge). In incongruent conditions the first storyteller was led to believe stories were known, and the second was led to believe they were unknown (that is, they were misinformed about the addressee's knowledge). The rest of the experimental procedure during the second round of narrations was identical to the first round. After the second storyteller finished telling the stories, both participants received their exit questionnaires and were then fully debriefed about the nature of the experiment. Experimental sessions were recorded on two video cameras, one focusing on the addressee, and the other focusing on the storyteller.

Analyses

Addressee feedback. Three different coders transcribed in detail addressees' verbal and nonverbal feedback responses. As in Experiment 1, addressee feedback was defined as verbal and/or nonverbal responses on the part of addressees that indicated they were attending, following, appreciating, or reacting to what speakers were saying. Unlike Experiment 1 (where the task required looking at a map), addressees tended to look at storytellers the majority of the time. Therefore, gazes to the storyteller did not count as feedback response. Addressee feedback was set in relation to the number of words used by the speaker (number of addressee feedback responses per 100 speaker words).

Reliability between all three coders in identifying addressee feedback was established for a subset of the data (10 storytellers narrating the full set of stories). Of the maximum number of 452 detected feedback responses, 89% were detected by all three raters. Disagreements were solved through discussion and led to a further specification of the coding criteria.

In addition to recording the frequency with which addressees gave feedback, the coding scheme targeted in more detail feedback responses that co-occurred with storytellers' first reference to target instruments. Specifically, coders recorded whether addressees gave feedback within the same clause the instrument was mentioned, and if they did, whether the response was initiated (a) before the target instrument was mentioned, (b) right after the instrument was mentioned, (c) at the end of the clause. Since an addressee's response is in part determined by the syntactic structure of the storytellers' utterance (e.g., Duncan & Fiske, 1977; Ford & Thompson, 1996; Sacks, Schegloff, & Jefferson, 1974), feedback responses were analyzed with respect to their timing only if they were given in response to utterances introducing the target instrument in the same verb phrase and following the target verb (e.g., "he stabbed him with the knife"). This procedure yielded 92 observations.

As a third variable, coders recorded what type of feedback response addressees produced. Five different coding categories covered the majority of all observed responses to target instruments: laughter, smiles, changes of facial expression indicating surprise (including "funny faces", see Swerts & Krahmer, 2005), nods, and vocalizations (this category included sentences, words and vocalizations such as "uh-hm"). Idiosyncratic responses (such as shaking the head) were coded as "other". For a more targeted analysis, laughter, smiles, and facial expressions were combined to one category of responses that were specific reactions to what storytellers were saying. These responses were considered to express addressees' amusement or surprise about the instrument, and were contrasted with the most common generic response, head nods.

Storytellers' narrations. Narrations were transcribed and the number of words speakers used for narrating the story was recorded. One coder (the author), blind to the experimental conditions, coded all narrations with respect to the mentioning of target instruments. To establish reliability, a second coder (a research assistant, also blind to the experimental conditions) coded a subset of the dataset (22 interactions of 12 stories each). The two coders agreed in 81% of all 264 coding decisions on the exact coding category.

Coding was based on the coding scheme reported in Lockridge and Brennan (2002). In a first step, coders noted whether storytellers mentioned the target instrument. If storytellers mentioned the target instrument, the first reference was identified and coded as to whether it was introduced with a definite or indefinite article.

In a next step, the target instrument was coded in its syntactic relation to the target action (see Table 4). There were four main positions in which target instruments could be mentioned: (a) within the same clause as the target action (categories 1- 3), (b) in a separate clause from the target action (categories 4- 5), (c) implicitly (category 6), or (d) in none of the above categories[12] (category 7). Unless instruments were mentioned implicitly coding further specified whether instruments were mentioned after the verb, before the verb, or, if mentioned within the same clause, incorporated into the verb. In total this yielded seven possible coding categories. Note that all original text stories mentioned instruments implicitly. In order to package instruments in a more prominent position storytellers had to modify the original wording of the stories.

[12] In this category storytellers either didn't mention the target action, or mentioned a different instrument (e.g., atypical instruments were often modified to typical instruments).

For a more focused analysis, the encoding of the target instruments was contrasted for two main positions in the narration: first whether storytellers mentioned target instrument explicitly (categories 1-5) or implicitly (category 6); and second, whether storytellers mentioned the target instrument within the same clause as the target action (categories 1-3), or as an afterthought or repair (category 4).

Statistical analyses

Data were analyzed using 2x2x2 ANOVAs with typicality as a within-subjects and within-items factor. For variables pertaining to addressees' behavior (total amount of feedback, target feedback, and timing of target feedback), addressee knowledge was varied within-subjects and storytellers' expectation as between-subjects variable. For variables pertaining to storytellers (definite reference and syntactic choices), addressees' knowledge and storytellers' expectation were varied between-subjects. For each ANOVA, paired t-tests were used for planned comparisons between typical and atypical stories in the four experimental conditions. Two types of analyses were conducted for each measure: *F1* or *t1* is the analysis by-subjects, which averages values across all twelve stories for each speaker or addressee (depending on the measure) and *F2* or *t2* is the analysis by-items, which averages values across all participants for each of the stories.

Results

Addressee feedback

Quantity of feedback. Addressees tended to give more feedback when they heard stories for the second time, $F1\ (1, 38) = 3.16$, $p = .08$; $F2\ (1, 11) = 26.77$, $p < .01$. Overall, the frequency with which addressees gave feedback was independent of

storytellers' expectations about addressees' knowledge, $F1$ (1, 38) = .67, p = .42; $F2$ (1, 11) = 10.68, p =.01[13]. But as the interaction illustrated in Figures 5a and 5b show, addressees only gave more feedback upon hearing the stories the second time when storytellers also expected stories to be known to addressees, $t1$ (19) = -2.69, p = .01; $t2$ (11) = -5.11, p < .01. Addressees did not increase their feedback when storytellers expected stories to be unknown to addressees, t (19) = -.24, p = .81; $t2$ (11) = -.03, p = .97.

Quality of feedback. Although addressees gave a comparable amount of feedback when hearing atypical- and typical- instrument stories, $F1$ (1, 38) = .48, p = .50; $F2$ (1, 11) = .08, p = .78, they were more likely to give feedback in response to atypical instruments than typical, $F1$ (1, 38) = 7.19, p = .01; $F2$ (1, 11) = 2.29, p = .16, see Figures 6a and 6b. This was true even when considering only those stories in which target instruments were mentioned within the same syntactic environment (within the same verb phrase, after the target action), $F1$ (1, 16) = 11.08, p < .01, $F2$ (1, 4) = 4.24, p = .11[14].

Overall, addressees' knowledge of the stories did not affect whether they gave feedback to instruments, $F1$ (1, 38) = 2.11, p = .15; $F2$ (1, 11) = 4.14, p = .07, nor did speakers' expectations, F (1, 38) = .07, p = .80; $F2$ (1, 11) = 11.19, p < .01. But even though an interaction between typicality of instrument and addressees' knowledge was not significant, $F1$ (1, 38) = 1.52, p = .23; $F2$ (1, 11) = 4.09, p = .07, planned contrast analyses suggest that addressees were marginally more likely to give feedback to atypical

[13] Not all effects that reached statistical significance across participants (by subject) also reached significance across stories (by item); and vice versa. For the current purpose, this section will interpret those effects that are significant by-subjects. Possible reasons why many effects did not yield significance across stories will be brought up in the discussion.
[14] Note the change in degrees of freedom due to a decreased number of observations.

than to typical instruments when they were unfamiliar with stories, in particular when speakers expected them to know stories, $t1$ (19) = -2.27, p = .04; $t2$ (11) = -1.27, p = .23, and less so when they expected them to not know stories, $t1$ (19) = -1.58, p = .13; $t2$ (11) = -1.16, p = .27. But when addressees were familiar with stories they were equally likely to give feedback to atypical and typical instruments, both when speakers expected them to know stories, $t1$ (19) = .25, p = .81; $t2$ (11) = .32, p = .76, and when they expected them to not know stories, $t1$ (19) = -1.16, p = .26; $t2$ (11) = -.70, p = .50.

The data about the timing of feedback responses to instruments are intriguing (see Table 5), although due to the small number of observations distributed unequally across experimental conditions, inferential statistics are not possible. Addressees were overall most likely to give feedback directly after the target instrument was mentioned. As predicted, the timing of their responses was more varied when they were familiar with stories, responding either early, before instruments were actually mentioned, or late, after the clause had ended. Addressees were most likely to respond promptly after the instrument was mentioned when they were unfamiliar with stories and a typical instrument was mentioned.

Addressees were generally much more likely to respond generically by nodding than specifically by showing surprise about instruments (see Table 6). This ratio did not change reliably with addressees' knowledge of the instruments, $F1$ (1, 19) = .25, p = .62; $F2$ (1, 10) = 11.05, p = .01, the typicality of the instrument, $F1$ (1, 19) = .01, p = .97; $F2$ (1, 10) = .06, p = .82, nor speakers' expectations about addressees' knowledge, $F1$ (1, 19) = .28, p = .60; $F2$ (1,10) = .28, p = .61. Judging by mean values only, addressees seemed more likely to show surprise in response to instruments when stories were unknown to

them, independent of the instruments' typicality. Addressees seemed least likely to show surprise when stories were known, storytellers had expected them to be known, and the instrument was typical. Planned t-tests indicated that only addressees who knew stories and were expected to know stories were more likely to show surprise in response to atypical instruments, $t1$ (1, 16) = -2.28, p = .04; $t2$ (1, 11) = -.47, p = .65.

Storytellers' narrations

Table 7 summarizes the frequencies with which storytellers mentioned atypical and typical instruments in each experimental condition, as well as the average number of words used by speakers to tell one story[15].

Explicit instrument mention. Consistent with Brown and Dell (1987), and Lockridge and Brennan (2002), storytellers were more likely to mention atypical instruments explicitly, $F1$ (1, 76) = 47.64, p = .00; $F2$ (1, 11) = 10.70, $p < .01$. However, when storytellers expected stories to be known to addressees, atypical instruments were less often mentioned explicitly, yielding in an interaction between instrument typicality and storytellers' expectations, $F1$ (1, 76) = 4.86, p = .03; $F2$ (1, 11) = .07, p = .80. Planned comparisons suggest that this interaction is driven by those storytellers who expected addressees to know stories and who then encountered addressees who indeed knew the stories (see Figure 7a & 7b). Only in this condition were atypical instruments equally as likely to be mentioned explicitly as typical instruments, $t1$ (19) = -1.34, p = .20; $t2$ (11) = 1.78, p = .10.

[15] One retelling was on average 45.50 (SD = 15.40) words long. Length of narrations did not differ for typical or atypical stories, $F1$ (1, 76) = .48, p = .49; $F2$ (1, 11) = .40, p =.54, and was not reflective of addressees' knowledge, $F1$ (1, 76) = .72, p =.40; $F2$ (1, 11) = 26.99, p =. 00, or storytellers' expectations of addressees' knowledge, $F1$ (1, 76) = .01, p = .91, $F2$ (1, 11) = .37, p = .55.

Definiteness of article. Also consistent with previous studies, atypical instruments were more frequently introduced with an indefinite article, marking the status of these instruments explicitly as not yet mutually known, $F1$ (1, 76) = 18.75, $p < .01$; $F2$ (1, 11) = 3.27, $p = .10$. However, when instruments were known to addressees they were marginally more likely to be introduced with definite than indefinite articles, $F1$ (1, 76) = 3.21, $p = .08$; $F2$ (1, 11) = .60, $p = .46$, suggesting that addressees' knowledge had a direct effect on how storytellers introduced new entities. In parallel with storytellers' decisions to mention atypical instruments explicitly, a planned comparison shows that only when storytellers correctly expected stories to be known to addressees were atypical instruments introduced equally as often with the definite article as typical instruments, $t1$ (19) = .61, $p = .55$; $t2$ (11) = -.57, $p = .58$, see Figure 8a and 8b.

Within clause mention. In contrast to previous findings, storytellers were more likely to mention atypical than typical instruments in a separate clause after the target action, $F1$ (1, 67) = 6.4, $p < .01$; $F2$ (1, 5) = 4.61, $p = .09$[16]. Addressees' knowledge of the stories did not have a direct effect on whether instruments were mentioned within the same clause as the action, $F1$ (1, 67) = .36, $p = .55$; $F2$ (1, 5) = 1.81, $p = .24$, nor did storytellers' expectation about addressees' knowledge, $F1$ (1, 67) = .31, $p = .58$; $F2$ (1, 5) = .02, $p = .89$. But planned comparisons suggest that only those storytellers who expected stories to be unknown when stories were indeed unknown actually distinguished between typical and atypical instruments by mentioning atypical instruments after the main target action, $t1$ (17) = -2.20, $p = .04$; $t2$ (9) = 2.37, $p = .04$.

[16] Note the change of degrees of freedom due to the fact that in some conditions none of the stories mentioned instruments within, or after the target action clause.

Changes in behavior over time

The *Feedback Modifies Expectation Hypothesis* predicts that those storytellers who held wrong expectations about addressees' knowledge will revise their initial expectations over the course of the experiment based on their addressees' feedback behavior. The first atypical story in the list should therefore be shaped by speakers' expectations, while the last atypical story should be shaped by addressees' feedback. Although none of the main coding categories showed significant changes in storytellers' behavior over time, storytellers' decision to mention atypical instruments explicitly was closest to the hypothesized pattern (see Figure 9, 10, 11): Those who had expected addressees to know stories were less likely to explicitly mention atypical instruments initially, but more likely to explicitly mention atypical instruments later. And vice-versa, storytellers who had expected addressees to not know stories were more likely to explicitly mention atypical instruments explicitly initially, but mentioned them implicitly later.

Exit questionnaires

Table 8 summarizes storytellers' answers to two central questions in the exit questionnaire: Whether storytellers thought addressees had heard stories before, and whether they thought addressees were confederates. Consistent with the experimental manipulation of speakers' expectations, the 65% of storytellers who had been led to believe addressees were hearing stories for the first time also reported at the end of the interaction that they thought their addressees had not heard stories before; and correspondingly, 75% of storytellers who had been led to believe addressees had heard stories before reported addressees had known stories. When storytellers' expectations

were contrary to addressees' actual knowledge storytellers were slightly more likely to assess addressees' knowledge based on addressees' actual knowledge. But overall, storytellers' answers to this question seemed to be influenced very little by the actual knowledge of the addressee. This suggests that, when explicitly asked to assess their addressees' knowledge, storytellers referred to information that was given to them in the experiment's instructions.

Over all experimental conditions, 40% of all storytellers' thought that their addressee could have been an experimental confederate, 41% did not think their addressees were experimental confederates, and 19% gave ambiguous answers to this question. Storytellers seemed more likely to believe their addressees were confederates when the experimental instructions had led them to believe addressees already knew the stories (45%) as opposed to when they had led them to believe addressees had not heard stories before (35%).

Discussion

Addressees' feedback was indicative of their informational needs: When hearing stories that involved the use of atypical instruments, they acknowledged this by reacting specifically when these instruments were first mentioned. In parallel, storytellers emphasized atypical instruments by introducing them with the indefinite article and mentioning them explicitly, altering original stories such that these instruments were positioned more prominently. Addressees' responses to atypical instruments did not result from storytellers' syntactic choices: Even when storytellers mentioned typical instruments in the same prominent position as atypical instruments they were less likely to receive a response from their addressees. Addressees' feedback in response to target

instruments can therefore be understood as independent assessment of how noteworthy they found the instruments to be. These findings corroborate previous studies by Brown and Dell (1987), and Lockridge and Brennan (2002) showing that atypical instruments are perceived as noteworthy by speakers. Moreover, the analysis of addressees' feedback indicates that this information is of concern not only to speakers, but also to addressees.

As addressees became familiar with stories, their feedback behavior changed. Upon hearing stories the second time, addressees seemed less likely to distinguish between typical and atypical instruments. Instead, their feedback became less specific: Addressees increased how much feedback they gave overall, but they seemed less likely to show surprise in response to hearing the instrument, and this feedback seemed less closely timed with storytellers' references to target instruments. This feedback behavior is best explained by the *Feedback for Grounding Hypothesis*, which predicts that addressee feedback serves the purpose of indicating that what speakers said was understood sufficiently. Hence, addressees give more feedback when stories were already known. By responding specifically to atypical instruments when this information is new, addressees signal their specific informational needs.

However, addressees' feedback was determined not only by their informational needs; it was also shaped by their partners' expectations. Addressees gave more feedback overall only when storytellers had also expected them to know the stories. Along the same lines, addressees were more likely to show surprise in response to atypical instruments only when they had heard stories before *and* storytellers knew that they had

heard stories before[17]. Such a pattern of behavior indicates that at least some aspects of addressee feedback are not determined from addressees' needs alone, but instead are established jointly in the interaction, providing support for the *Expectations modifies Feedback Hypothesis.*

Storytellers differed in how they introduced typical and atypical instruments in a way that appeared to be shaped by addressees' feedback. When receiving quantitatively more feedback overall, which was however qualitatively less specific, storytellers also stopped distinguishing atypical from typical instruments, introducing them only implicitly, and with the definite article. Note that although storytellers failed to adapt to addressees' knowledge of the stories when they had not expected them to know stories, their behavior aligned with addressees' overall feedback behavior. This suggests that the frequency with which addressees give feedback guides storytellers' decision to emphasize atypical instruments. These findings replicate findings by Lockridge and Brennan, confirming that storytellers do take addressees' specific informational needs into account. Moreover, the current study established that addressee feedback is one possible mechanism by which speakers assess those needs.

Although storytellers reacted to their addressees' overall feedback responses, storytellers appeared indifferent to whether addressees reacted specifically in response to instrument mention. Addressees' lack of feedback to storytellers' mention of atypical

[17] The finding that addressees reacted *more* surprised about atypical instruments, even though they knew about it in advance and storytellers also knew that they knew about it, supports those theories emphasizing the social and communicative nature of emotional displays (e.g., Chovil, 1991; Fridlund, 1991). According to these theories displays of emotions (e.g., surprise) are not only shaped by the individual's state of mind (e.g., whether information is new or known), but instead is also to a large extent shaped by the social context (e.g., whether information is mutually shared).

instruments could have cued storytellers that they were familiar with stories, but the storytellers failed to adapt to this cue when they had not expected addressees to be familiar with stories. Possibly, storytellers interpreted addressees' behavior in the light of their own expectations (ignoring this cue when it did not conform to expectations). Along similar lines, Kuhlen and Brennan (in press) found that speakers adapted to addressees' feedback only when they had an attribution for addressees' behavior. This suggests that neither addressees' feedback alone, nor storytellers' expectations alone predict storytellers' behavior, indicating that storytellers integrate their prior expectations with their addressees' informational needs.

In line with the *Expectations Modifies Feedback Hypothesis*, storytellers may be shaping addressees' feedback behavior such that it corresponds to their expectations. Concerning influence in the other direction, there was only weak support for the *Feedback Modifies Expectations Hypothesis*, which predicts that speakers adapt their expectations over the course of the interaction to addressee feedback. Different experimental paradigms might be more sensitive to measuring such an incremental adaptation. In this experiment, stories were always narrated in the same order. The assessment of how storytellers integrate their prior expectations with incoming feedback over time was probably confounded by differences between stories (e.g., some instruments may have been more atypical than others, or some storylines may have been more conducive to mentioning instruments early on or later).

Such differences between stories likely contributed to the fact that many adaptations observed reliably across storytellers were not reliable across stories. Note that if storytellers revise their initial expectations on the basis of addressee feedback, stories

at the end of the list presumably are influenced more by addressee feedback than stories at the beginning of the list, which are influenced more by storytellers' expectations, resulting in less power to detect effects that hold across all stories. In line with findings from Experiment 1, stories may not constitute independent events (and hence independent observations). Instead, previous experiences in narrating one story may be shaping subsequent stories. Together with possible variations in storylines, this made it difficult to find effects that hold across all stories.

In contrast to Brown and Dell (1987), and Lockridge and Brennan (2002), storytellers in this study were not more likely to mention atypical instruments within the same clause as the target action. Instead they were more likely to mention typical instruments after the target action. This may reflect a different implementation of the coding scheme. While Lockridge and Brennan assigned on average 2.53% of all instrument mentioned to the category "mentioned in separate clause after the target verb", the current coding assigned this coding category about two and a half times as often, 6.39%. The current coders' criterion to code instruments as separate but belonging to the target action may have been less conservative than previous coding criteria.

In summary, addressees indeed seem to be signaling their informational needs through their feedback, confirming Lockridge and Brennan's speculation. In contrast to Brown and Dell, speakers do not judge instruments' noteworthiness based on the needs of a generic addressee (or else they would always distinguish atypical from typical instruments), but instead they seem to be taking addressees' specific needs into account. When assessing addressees' needs, speakers thereby appear to be integrating incoming feedback from addressees with their own expectations of addressees' knowledge, to the

extent that speakers' expectations may even be instrumental in shaping addressees' feedback behavior.

However, the pattern of data from Experiment 2a does not support the conclusion that Brown and Dell's divergent results were due to the fact that addressees were confederates, who revealed their lack of informational needs through their feedback. If that were the case, addressees who heard stories before should have elicited similar (generic) adaptations as Brown and Dell found, prompting storytellers to consistently differentiate between atypical and typical instruments. But instead, addressees who appeared to know stories prompted storytellers to stop differentiating between atypical and typical instruments. Possibly, addressees who hear a story (only) twice are not comparable to confederates who hear a story many times, as they did in Brown and Dell's experiment. Experiment 2b investigates whether storytellers may shift their adaptations when confronted with addressees who not only have low informational needs, but who in fact seem to know the stories better than the storytellers themselves.

5.2 Experiment 2b: Speakers don't adapt to addressees who have no apparent needs

Method

Design

The experimental design of Experiment 2b was identical to that of Experiment 2a, with the exception that addressees' informational needs were manipulated by exposing them to the same set of stories four times. The first time, stories were told by a live storyteller, the second and third time addressees watched video recordings of two storytellers telling the stories, and the fourth time stories were told by a second live

storyteller. Addressees' and storytellers' behavior was analyzed when addressees heard stories for the first time (high needs), and then for the fourth time (very low needs). As in Experiment 2a, storytellers were led to either hold congruent or incongruent expectations about their addressees' knowledge of the stories.

Participants

One hundred-fourteen volunteers (68 females, 46 males) who had not participated in Experiment 1 or 2a participated in a total of 38 triads (38 addressees, 76 storytellers). Table 3 lists the number of male and female storytellers and addressees for each condition. All participants were native speakers of English[18]. Ten sessions were excluded and replaced for the following reasons: One speaker appeared not to be a native speaker of English, one addressee fell asleep as he heard the narrations for the fourth time, and eight addressees mentioned explicitly to storytellers that they had heard stories before, thereby contradicting speakers' prior expectations about addressees' knowledge[19]. Two additional sessions had to be excluded after the data collection was completed due to procedural errors. These last two sessions were both in the condition in which storytellers held incongruent expectations about addressees' needs, resulting in a lower number of observations in this cell of the design. All participants received either research credit or $9 per hour for their participation.

---

[18] All participants indicated that they were fluent, native speakers of English. Some reported that they were bilingual.

[19] These storytellers had been led to believe that addressees were hearing the stories for the first time. Since the goal of the experiment was to investigate how storytellers implicitly consolidate their expectations with addressees' feedback, these sessions were excluded. The fact that in Experiment 2b such a relatively large number of addressees felt the necessity to tell speakers that they had heard stories before (only three addressees did so in Experiment 2a) reinforces how important it is for conversational partners to establish common ground on addressees' prior knowledge.

Materials

Materials were identical to materials used in Experiment 2a, with the addition of two new video-recorded versions of the experimental stories that served the purpose of exposing addressees to the stories a second and third time. These stories were told by two research assistants (one male, one female); since the goal was to make addressees more knowledgeable about the stories than participant storytellers themselves, the research assistant storytellers told the stories in the same wording as the original stories (the text was prompted out of view of the camera). Recorded narrations were compiled in one digital video file, which was played to participants on a Dell computer.

Procedure

Procedures were identical to procedures established for Experiment 2a. Under the pretense of investigating memory, participant addressees were informed that they would be hearing the same set of stories four times: The first time told by a live speaker from the subject pool, the second and third time by two participants who had been video recorded telling the stories in previous experimental sessions, and then a fourth time by a second live speaker. As in Experiment 2a, storytellers were led to believe addressees had either just arrived early (Expectation: Stories unknown), or had participated in experiment previously (Expectation: Stories known). Note that the manipulation of speakers' expectations only indicated that addressees had heard stories before, but it was not informative about how many times they had heard them.

After the first set of narrations, the first storyteller left the experiment room and the addressee was seated in front of a computer to watch the two sets of recorded narrations. Addressees were reminded to listen carefully since they would later be

quizzed on details of the stories. After listening to the narrations, they were asked to return to their seat and wait for the second live storyteller to arrive.

Data analyses

Data was analyzed in identical fashion to Experiment 2b. In addition to the regular statistical analyses, addressees' behavior in Experiment 2b was directly compared to addressees' behavior in Experiment 2a by entering data jointly into one ANOVA with Experiments 2a and 2b as additional between subject factor. For a more targeted analyses, independent t-Tests compared addressees' behavior in Experiment 2a with Experiment 2b for selected contrasts.

Results

Addressee feedback

Quantity of feedback. Unlike in Experiment 2a, addressees in Experiment 2b did not increase the overall frequency of their feedback once they became familiar with stories, $F1\ (1, 36) = .09$, $p = .77$; $F2\ (1, 11) = .10$, $p = .75$. Specifically, in the condition that yielded the clearest effect of addressees' knowledge in Experiment 2a, namely when storytellers expected addressees to know stories, addressees in Experiment 2b did not give more feedback once they were familiarized with stories, $t1\ (36) = 1.19$, $p = .24$; $t2\ (11) = -3.96$, $p < .01$. If anything, the tendency was the opposite, see Figure 12a and 12b.

A direct comparison of addressees' behavior in Experiment 2a with Experiment 2b suggests that, overall, addressees in Experiment 2a did not differ from addressees in Experiment 2b in their frequency of giving feedback, $F1\ (1, 74) = .22$, $p = .64$; $F2 = (1, 22) = 1.53$, $p = .23$, and an interaction between experiments and addressees' knowledge did not reach significance across participants, $F1\ (1, 74) = .84$, $p = .36$; $F2\ (1, 22) =$

16.09, $p < .01$. But crucially, when storytellers expected addressees to know stories, addressees who heard stories four times behaved differently from addressees who heard stories only twice, failing to indicate their familiarity with stories by increasing their feedback, $t1\ (38) = 1.93$, $p = .06$; $t2\ (22) = 4.42$, $p < .01$.

Quality of feedback. As in Experiment 2a, addressees acknowledged atypical instruments more often than typical ones by giving feedback specifically when instruments were first introduced, $F1\ (1, 35) = 6.14$, $p = .02$; $F2\ (1,11) = 6.03$, $p = .03$. Once familiar with the stories, addressees were less likely to differentiate between typical and atypical instruments than when stories were new, yielding an interaction between typicality of instrument and addressee knowledge, $F1\ (1, 35) = 7.37$, $p < .01$; $F2\ (1, 11) = 3.92$, $p = .07$, see Figure 13a and 13b.

Although addressees' response to instruments in Experiment 2b marginally differed from addressees' response in Experiment 2a, $F1\ (1, 73) = 2.52$, $p = .12$; $F2\ (1, 22) = 3.90$, $p = .06$, they did not seem to differ with respect to how they expressed their familiarity with stories or their reaction to atypical instruments, as suggested by the insignificant interaction between the experiments and addressees' knowledge, $F1\ (1, 73) = .43$, $p = .54$; $F2\ (1, 22) = 1.78$, $p = .20$, and between the experiments and typicality of instrument, $F1\ (1, 73) = .05$, $p = .82$; $F2\ (1, 22) = .24$, $p = .63$.

However, addressees' informational needs did not influence other qualitative dimensions of their feedback responses as clearly. Unlike patterns found in Experiment 2a, addressees did not seem to indicate their familiarity with stories through the type of responses, see Table 6. Accordingly, addressees in Experiment 2b were equally as likely to display amusement upon hearing the stories for the first time and for fourth time, $F1$

(1, 14) = .02, p = .90; $F2$ (1, 8) = 3.19, p = .11. Opposite to Experiment 2a, addressees who knew stories before and were expected to know stories were not more likely to display amusement towards atypical than typical instruments, $t1$ (13) = .68, p = .51; $t2$ (10) = .57, p = .59. In fact, they seemed *least* likely to show amusement towards atypical instruments in this condition.

A direct comparison between addressees in Experiment 2a and Experiment 2b confirmed that indeed, addressees in Experiment 2a gave a different type of feedback than addressees in Experiment 2b, $F1$ (1, 33) = 3.77, p = .06; $F2$ (1, 18) = 2.7, p = .12. And more specifically, when addressees knew stories and speakers expected them to know stories, addressees in Experiment 2b were more likely to display surprise towards atypical instruments than addressees in Experiment 2a, $t1$ (32) = 2.53, p = .02; $t2$ (21) = 1.70, p = .10.

No clear pattern emerged as to when addressees gave feedback to target instruments, see Table 5. While addressees in Experiment 2a seemed to respond more promptly when stories were unknown to them, addressees in Experiment 2b did not confirm this pattern.

Storytellers' narrations

Table 9 summarizes the frequencies with which storytellers mentioned atypical and typical instruments in each experimental condition, as well as the average number of words used by speakers to narrate one story[20].

[20] On average, storytellers used 46.10 (SD = 13.26) words for each narrations. There was a slight tendency to use more words when stories were atypical, $F1$ (1, 75) = 3.39, p = .07; $F2$ (1, 11) = 3.67, p = .08.

Instrument mention. When instruments were atypical, storytellers across all experimental conditions introduced them more often explicitly (instead of implicitly later in the story) than when instruments were typical, $F1\ (1, 72) = 45.25$, $p < .01$; $F2\ (1, 11) = 8.95$, $p < .01$ (see Figures 14a & 14b). Atypical instruments were also more often introduced with an indefinite article than typical instruments, $F1\ (1, 72) = 12.11$, $p < .01$; $F2\ (1, 11) = 2.28$, $p = .16$ (see Figure 15a & 15b). Unlike in Experiment 2a, storytellers' expectations, addressees' knowledge, or a combination of these factors had no effect on how instruments were introduced. Neither instruments' typicality, storytellers' expectations, nor addressees' knowledge had an influence on whether storytellers mentioned instruments within the same clause or later as an afterthought.

Changes in behavior over time. As Table 10 indicates, there is no clear indication that storytellers revised their initial expectations in response to feedback incongruent with their expectations. More specifically, in Experiment 2a storytellers who expected stories to be unknown seemed initially more likely to mention atypical instruments explicitly than later in the interaction, when addressees' feedback indicated that stories were known. In Experiment 2b, storytellers always seemed to be less likely to mention explicitly atypical instruments towards the end of the interaction, independent of whether addressees' feedback could have revised their expectations.

Exit questionnaires

Table 8 reports storytellers' answers in the exit questionnaire. On average, 54% of storytellers who were led to believe addressees were hearing stories for the first time reported at the end of the interaction that they thought addressees had not heard stories before. Likewise, 53% of storytellers who were led to believe addressees had heard

stories before reported later that addressees had heard stories before. Compared to Experiment 2a, a lot more storytellers who had been misinformed about addressees' knowledge of the stories judged addressees on their actual knowledge.

In comparison to Experiment 2a, only a few more storytellers in Experiment 2b thought that their addressee was a confederate. However, storytellers in Experiment 2b were considerably more likely to think addressees were confederates when they had heard stories before: When interacting with addressees who knew stories, storytellers reported about four times as often that their addressees were confederates.

Discussion

Experiment 2b confirms that addressee feedback contains cues about addressees' informational needs. Addressees are more likely to give feedback about atypical than typical instruments, supporting the assumption that these instruments are particularly noteworthy (and appropriate to acknowledge by grounding). Similar to Experiment 2a, storytellers as well acknowledged the special status of atypical instruments by introducing them early on in the narrative and marking them as indefinite.

Once addressees were well familiar with stories, they stopped responding preferentially to atypical instruments. Similarly to Experiment 2a, addressees' lack of reaction to atypical instruments could have indicated to speakers that addressees judged these instruments no longer noteworthy. However, also in parallel with Experiment 2a, storytellers did not appear sensitive to this dimension of addressees' feedback behavior and failed to adapt to addressees' specific needs.

Instead of aligning with addressees' response to instruments, storytellers' behavior in Experiment 2a had aligned with addressees' overall frequency of giving

feedback. Crucially, on this dimension, addressees in Experiment 2b failed to indicate their informational needs: In contrast to addressees who heard stories for the second time (Experiment 2a), addressees did not increase the overall frequency of their feedback upon hearing stories for the fourth time (Experiment 2b). Corresponding to this feedback cue, storytellers in Experiment 2b continued to emphasize atypical over typical instruments, independent of addressees' actual informational needs, and independent of their own expectations about addressees' informational needs.

Storytellers' behavior when interacting with addressees who were hearing stories for the fourth time thereby resembled storytellers' behavior when interacting with confederate addressees, as reported in Brown and Dell (1987): Storytellers did not seem sensitive to addressees' specific informational needs, emphasizing atypical instruments more when these instruments were unknown to addressees, but instead seemed to be adapting to a generic addressee's needs, *always* emphasizing atypical over typical instruments. But under the assumption that addressees' overall feedback frequency is the decisive cue in shaping storytellers' behavior, storytellers in Experiment 2b should be showing exactly the behavior they are showing: Since addressees don't indicate having any specific needs (by giving more feedback overall), storytellers don't take their specific needs into account.

Addressees' feedback behavior may therefore have contributed to speakers' lack of specific adaptations. While addressees who heard stories twice may have appeared knowledgeable, addressees who heard stories four times may have appeared sated. In this way, upon hearing stories for the fourth time, addressees started behaving somewhat generically, themselves, giving feedback with stereotypical frequency. Possibly,

qualitative cues in addressees' behavior (e.g., failing to display surprise upon mentioning atypical instruments) may have signaled storytellers that addressees were only pretending to listen. Participants may have established a culture of pretense, behaving "as if" they were listening and "as if" they were telling a story. Instead of behaving in accordance with their actual informational needs, these participants may be adapting to the demands of the conversational task.

In summary, addressees who heard stories for the second time indicated their low informational needs through the frequency with which they gave feedback. Subsequently, storytellers adapted to their informational needs by de- emphasizing atypical information. Addressees who heard stories for the fourth time did not signal their lack of informational needs (at least not by how often they give feedback). Subsequently, storytellers did not adapt. Under the assumption that addressees who heard stories four times resembled more closely confederate addressees such as those in Brown and Dell (1987) who heard stories many more than four times, the contrasts between Experiment 2a's and Experiment 2b's results suggest that confederates may be changing the nature of the conversational "game" by behaving like generic addressees, who in turn elicit more generic behavior from speakers.

6 General Discussion

The goal of this dissertation was to reach a better understanding of the processes underlying speakers' adaptations to addressees' informational needs. Two different experimental settings were used to observe how speakers accommodate to addressees' informational needs while either giving directions or narrating short stories. Addressees' informational needs were manipulated either by providing them with the route speakers were describing, or by familiarizing them with the stories speakers were telling. In addition, speakers' expectations about addressees' needs were manipulated through prior experiences interacting with the same addressee, or through giving speakers explicit information about addressees' knowledge. The two experimental approaches yielded converging evidence for the questions at stake, as follows.

Addressee feedback shapes speakers' utterances

Addressee feedback appears to shape speakers' utterances on different levels of speech production. A particularly relevant cue is the frequency with which addressees give feedback. Depending on how much feedback they received, speakers shifted how clearly and deliberately they articulated central referring expressions, whether they introduced salient information as new or old discourse entities, where they positioned this information, and how much additional detail they provided. Speakers were not always sensitive to qualitative changes in addressees' feedback behavior, such as whether addressees gave feedback specifically in response to certain information. This may suggest that, at least in the conversational settings presently studied, evidence about the

addressees' informational needs accumulates over the course of the conversation through the overall frequency with which addressees give feedback.

Speakers adapted in response to their addressees on various levels of speech production. The present two experiments investigated how addressees shape processes of articulation and syntactic encoding. Together with previous studies showing that speakers adapt to their addressees in articulation (Galati & Brennan, 2010), syntactic encoding (Lockridge & Brennan, 2002), semantic encoding (Kuhlen & Brennan, in press), and speech accompanying hand gestures (Kuhlen, Galati, & Brennan, under review), results support the assumption that information about the conversational partner can in principle influence a speakers' utterance on multiple levels. There is no strong evidence that certain processes of utterance production are encapsulated from influences by the conversational partner.

The parallel pattern in addressees' behavior on the one hand and speakers' behavior on the other hand observed in both experimental settings provides strong support for the idea that addressees' feedback is one of the mechanisms responsible for shaping speakers' utterances. Nevertheless, in both experimental settings, addressees' feedback was manipulated only indirectly, by controlling addressees' informational needs. In order to exclude the possibility that both addressees' and speakers' behavior is influenced by a third factor, more controlled follow-up studies will need to manipulate addressee feedback directly. The advantage of the current approach was in gaining insight into the question of whether and how addressees, in relatively natural and unrestrained situations, use their feedback to indicate their informational needs.

Addressee feedback reflects their informational need, and speakers' expectations

To a certain extent, addressees' feedback reflected their informational needs. When receiving directions on how to follow a route on a map addressees initially gave more feedback when they had to draw the route into the map (and hence needed the information speakers were providing), and they gave less feedback when they already had the route marked into their map (hence did not need the information speakers were providing). Likewise, addressees gave less feedback when hearing stories they were unfamiliar with, and they gave more when they had heard stories before (at least when speakers also expected addressees to be familiar with stories, as will be discussed next). Addressees' feedback also indicated which particular information (e.g., an atypical instrument) was noteworthy by giving feedback specifically in response to this information.

Whether addressees with high informational needs gave more or less feedback varied with the conversational task. This aligns with the idea that feedback is instrumental in grounding speakers' utterances. More collaborative conversational tasks, such as giving and receiving directions, require incremental coordination between speakers and addressees, and addressees take considerable initiative in grounding. Hence, when addressees truly collaborated with speakers they gave more feedback. Conversational tasks such as narrating require one partner to take most of the initiative, and therefore do not afford such close coordination. Hence, addressees generally give less feedback. The overall pattern of how feedback was shaped by addressees' informational needs supports the idea that feedback enables speakers to tailor their utterances more specifically to the needs of their addressees. In both experimental settings, when

receiving more feedback, speakers were able to better accommodate addressees' needs by leaving out superfluous information and attenuating information that is already known to addressees.

However, addressees' feedback was shaped by more than addressees' informational needs. In both experimental settings, addressees often continued to display behavior they had established in previous interactions, although their actual informational needs had changed. Experiment 1 suggests that addressees' initial informational needs set the standard for later interactions. But this does not mean that conversational partner establish their informational needs in a "one-shot" fashion and then fail to accommodate to a change in their needs. As Experiment 2a shows, addressees do change their feedback behavior when speakers' expectations justified such a change of behavior: When storytellers expected stories to be known, addressees gave more feedback upon hearing them for a second time. But when storytellers expected stories to be unknown addressees were less likely to adapt their feedback upon hearing stories for the second time. Likewise, matchers in Experiment 1 failed to adapt their feedback to their needs in those situations when directors couldn't anticipate a change. Thus, in both experimental settings, addressees' feedback failed to reflect their changed informational needs whenever speakers' expectations about addressees' needs were incongruent with addressees' actual needs.

Possibly, speakers' expectations enable addressees' expression of their informational needs. Such an interpretation is in line with studies in social psychology that propose people seek to confirm their own expectations of the other, thereby actively contributing to shaping the other's behavior according to their expectations. The current

studies support such a self-fulfilling prophecy and may have captured these processes on the level of conversational dynamics. Exactly *how* speakers may shape addressees' feedback is however not clear. My first pilot analyses following up on Experiment 1 were not able to establish that speakers were prompting addressees to give feedback by seeking out their gaze. Experiment 2 excluded the possibility that speakers' syntactic choices were eliciting addressees' feedback. Possibly, a combination of factors promotes addressee feedback, such as speakers' gaze in combination with the utterance intonation or pauses in speech.

Despite the clear parallels between addressees' informational needs and their feedback behavior, addressees' feedback was altogether a somewhat weak cue. Changes in addressees' behavior in response to their informational needs were often only marginally significant. In particular, qualitative dimensions of addressee feedback, such as the type of response or the timing of the response, did not yield clear patterns of results. This may be due to large inter-individual differences in giving feedback. And it may be an attribute of studying conversation in rich, relatively unrestrained, and diverse conversational situations.

Addressees may have also felt compelled to pretend to be listening, in order not to be impolite towards the speaker. In conversation people often spontaneously take on roles that involve a certain amount of pretense (e.g., Clark & Gerrig, 1990; Goffman, 1959). This brings up the point that conversation not only serves the purpose of transferring information. Even in task-related experimental settings, conversation also serves other purposes such as interpersonal affiliation, and in fact may be instrumental to interactively

constructing an understanding of one's self and the conversational situation (e.g., Davies & Harré, 1990; Goffman, 1959; Shotter & Gergen, 1989).

Conversational partners thereby do not always avoid talking about things that are already known. Revisiting the scenario outlined in the introduction, we may recall situations in which we patiently (and maybe even gleefully) listen to our friends' stories again and again. In this case, the topic of conversation is not new to either of us. And the function of the conversation is not so much the information itself, but instead the sharing of mutual experiences. In a similar way conversationalists in the present experimental settings may have jointly constructed a definition of the conversational situation, or what conversational "game" they are playing. This may entail how much accuracy the task affords, how to tell stories, what information is noteworthy (recall that in Experiment 2a addressees were more likely to express their surprise about atypical instruments when it was mutually understood that stories were well familiar), and how to play along and listen, even when the information is not really relevant.

Speakers integrate top-down cues with bottom-up cues

When assessing addressees' informational needs, speakers in both experimental settings seemed primarily influenced by addressees' feedback. Their expectations about addressees' informational needs seemed to influence this assessment only indirectly by guiding how they interpreted addressees' feedback or what opportunities they gave addressees to give feedback. Some previous studies have shown a stronger and more direct influence of speakers' expectations of their conversational partner. As discussed in the introduction, Russell and Schober (1999) showed that speakers use their expectations about their addressees' needs to guide utterance planning, irrespective of the addressee's

actual behavior. And other studies find that at least initially speakers' expectations influence speakers' utterances (Brennan, 1991; Norlund et al., 2009). Speakers' expectations may have a more pronounced influence if the manipulation of their expectation is stronger, or the assessment of such an influence is more fine-grained.

There was some indication that speakers may update their initial expectations based on addressees' feedback. This would be in line with studies by Brennan (1991) and Norlund et al. (2009) showing that speakers' utterances are initially shaped by speakers' expectations about their addressee, but later are based more heavily on addressees' actual behavior. Temporally more fine-grained measures such as eye tracking or EEG may yield more sensitive investigations of such an integration process (see e.g., Tanenhaus & Brown- Schmidt, 2007; van Berkum, in press). Also, a randomization of stimuli material could eliminate any confounding effects that may be due to variation in the stimuli.

Confederates may be establishing a different type of conversational game

One goal of this project was to investigate possible consequences of using confederates as conversational partners. The present experiments approximated the confederate role by revealing to addressees the map routes speakers were describing, or by repeatedly exposing them to the same stories. Thus, although none of the present experiments actually employed confederates, addressees with low informational needs may have an important characteristic in common with the typical confederate: They do not really need the information they are presently receiving. In fact, they can probably even anticipate what speakers will say next. Thus, instead of having a genuine interest, these addressees play along and pretend to be listening.

The present work suggests that this can have two effects: Addressees may express their lack of genuine interest in their feedback. This may happen involuntarily: Although addressees in the present experiments were encouraged not to reveal to their lack of informational needs their feedback still reflected it. In return, speakers may adapt to this feedback by attenuating crucial information. Observed behavior may therefore illustrate how speakers behave when interacting with addressees who already know everything, but may not be representative of speakers' behavior in general. This concern seems particularly relevant in situations in which speakers have reason to assume their addressees have no real informational needs, for example, if speakers are aware they are interacting with confederates. Typically, studies that use confederates try to prevent participants from knowing that they are interacting with confederates. But not the fact that conversational partners are confederates, but, more accurately, the knowledge associated with being a confederate is what may trigger a behavioral adaptation in speakers.

Aside from this concern, there appears to be a second, more severe risk to using confederates. Confederates may be changing the nature of the conversational game. Experiment 2b suggests that addressees who have very low informational needs, and in fact are more familiar with the experimental task than speakers themselves, may disengage from their role as genuine addressee and instead display generic feedback behavior. Speakers respond to this by behaving generically, themselves. Especially when studying partner-specific adaptations, using confederates may lead to wrong conclusions: Speakers appear to be ignoring their addressees' informational needs, when in fact, they

are responding to a different conversational game: How to behave as a generic speaker interacting with a generic addressee.

In summary, the current data suggest that addressees' informational needs shape their feedback, addressees' feedback shapes speakers' utterances, and speakers in return shape addressees' feedback. Such a complex structure of interrelated behavior emphasizes that conversation is a dynamic system of collaborative actions. Unilateral actions, such as "addressees shape speakers", fall short of capturing the complexity of the processes underlying interaction. Even less justice is done by investigating such processes with paradigms that replace one conversational partner with an experimental confederate.

7 Conclusions

In order to understand a complex phenomenon like dialogue, researchers often strive to reduce it to more basic features or sub-processes. Taken to the extreme, such an approach might suggest that complex behavior such as a conversation is nothing more than the sum of its parts-- for example, the contributions of each individual conversational partner. The complex pattern of coordination between speakers and addressees as evident in the currently presented data, however, suggests otherwise. For example, neither addressees' informational needs alone, nor speakers' expectations of addressees' needs alone, but instead the combination of these two factors was able to explain speakers' and addressees' behavior in the interaction. When speakers and addressees act, they therefore do not act autonomously, but instead mutually shape each other. Dialogue needs to be understood (and investigated) as a collaborative and dynamically evolving process. This has not only theoretical implications for our understanding of dialogue; it also has implications for the cognitive architecture in the individual mind that is needed to support such processes. And it brings about methodological challenges for how to go about studying dialogue.

In addition to conversational partners coordinating with each other, language processes also appear to be coordinated within the individual mind. In order for speakers to incorporate perceptual (bottom-up) cues available in addressees' feedback speakers must be planning and articulating their utterances in parallel to monitoring addressees' reactions to what speakers are saying, and integrating these reactions with their prior (top down) expectations about their addressees' needs. Addressees, on the other hand, must be able to simultaneously listen to speakers' utterances, interpret them, and decide how to

react to them, perhaps based on cues available in speakers' behavior on what reaction is expected of them. Such processes require a cognitive architecture that allows various subprocesses of speech production and speech comprehension to run in parallel and influence each other in a rather fine-grained way.

The close coordination between speakers and addressees in dialogue situations brings methodological challenges. For one, conversational partners jointly construe their understanding of the conversational situation, incorporating their prior expectations and experiences. An experimental manipulation at "Time 1" therefore is likely to have a different effect as the identical manipulation at "Time 2" (see Experiment 1). In other words, in within-subject designs, experimental manipulations may carry over and inform participants' later behavior. Researchers must therefore consider whether a within-subject manipulation is appropriate and, if so, carefully balanced the order of the manipulation.

Another, perhaps more severe, methodological challenge is that the appropriate unit of study for dialogue might not be the individual, but instead the conversational dyad. If dialogue is understood as a truly collaborative activity, both conversational partners play an important role on a moment-by-moment basis. By replacing conversational partners with confederates, researchers therefore may be reducing dialogue beyond its basic features and transforming it to something else entirely.

I have argued that confederate addressees may be particularly prone to shape an interaction in undesirable ways, because they typically have more knowledge than their conversational role affords. Should researchers therefore refrain from using confederates as addressees? While clever experimental designs may allow an experimentally controlled investigation of dialogue processes without the use of confederates (see e.g.,

Brown-Schmidt & Tanenhaus, 2008; Brown-Schmidt, Gunlogson, Tanenhaus, 2008; Hanna & Brennan, 2007; Kraljic & Brennan, 2005) confederates may remain unavoidable for other investigations. Nevertheless, it seems to be a reasonable goal for future research to supplement experiments that use confederates with studies that use naïve conversational partners. This way the chance of studying behavior that only arises as an artifact of the experimental setting can be reduced.

Using naïve participants instead of confederates as conversational partners might not always solve the problem. Naïve participants, as well, can be bored by uninteresting experimental procedures (see Experiment 2b, in which one participant even fell asleep). It could be argued that a well-trained and motivated confederate may do a better job than an unmotivated, naïve participant. The current set of experiments suggest that, rather than omitting confederates entirely from the role of being an addressee, another solution may be to develop experimental tasks that engage addressees, confederate or naïve, by giving them real informational needs. If addressees do not only have to "sit and listen", but the conversational tasks requires them to respond contingently on their partners' behavior, their informational needs are likely to develop and be expressed quite naturally. Therefore, confederates who have informational needs and a genuine interest in incorporating the information they are receiving are likely to elicit behavior in their conversational partners that reflect more closely natural conversation.

More remains to be learned about how conversational partners coordinate. For example, the present experiments were limited in their scope of addressing temporally more fine-grained processes. Within these limitations, the data suggest that the temporal coordination between conversational partners is an important dimension when

considering how addressees express their informational needs. What is more, an accurate understanding of how quickly speakers integrate incoming feedback cues bears important distinctions for the type of models proposed to account for partner-specific information influencing speech production. Building upon my present work, I am interested in understanding further how an addressee's feedback response is integrated, moment-by-moment, with a speaker's expectation of the addressee's behavior using temporally fine-grained measures such as EEG. Such an investigation would not only allow further insight in how flexibly speakers can adjust their assessment of addressees' needs on a moment-by-moment basis, but would also allow a deeper understanding of the neural underpinnings of coordinative processes between conversational partners.

Of further interest would also be a closer investigation of how affordances of the conversational task ("language game") shape addressees' feedback and the coordination between conversational partners. The current project suggests that the grain of coordination required by the task shapes how frequently addressees give feedback. Qualitative differences in addressee feedback may also be of relevance here. For example, addressees' responses may be more precisely timed in situations where speakers and addressees need to coordinate more closely. Establishing reliable differences in qualitative dimensions of addressee feedback however is challenging. As done in the current projects, large individual differences in giving feedback make a within-subject investigation of addressee behavior preferable. Moreover, due to co-dependencies between speakers' and addressees' behavior, it is necessary to control for the syntactic environment and possibly other factors such as speakers' gaze and intonation that may be prompting feedback at a particular point in time. As the present data show, such

restrictions drastically decrease the number of naturally occurring observations, and hence the statistical power.

The currently developed understanding of dialogue as a process in which conversational partners shape each other in a reciprocal fashion also has implications for our understanding of how social dynamics unfold. A different line of future work could therefore investigate the social consequences suggested by current findings. This project indicates that speakers' expectations about their addressees may be instrumental in shaping addressees' behavior, to the extent that addressees confirm speakers' expectations. Such a vicious cycle of self-fulfilling prophecies may also apply in situations in which individuals hold certain expectations about their partners based on existing stereotypes. Similar experimental designs as presently used could be transferred into a setting that manipulates speakers' expectations by evoking cultural stereotypes about the addressee. If such investigations can replicate, on the level of dialogue processes, that speakers shape addressees' behavior to match their expectations, this would be relevant to understanding the mechanisms behind the self-perpetuating nature of stereotypes. A further investigation of *how* speakers shape addressees will be important to verify these processes.

The field of cognitive psychology in general, and psycholinguistics more specifically, tend to study individual minds and language use in isolation. However, much of what we do and what constitutes us as human beings involves social activities. A deeper understanding of how people coordinate and shape each other is therefore of vital interest, not only to psycholinguistics, but social psychology, and the study of joint action in general.

Table 1

Main Landmarks and their Phonological Characteristics Used for Measuring Word Duration and Rating Word Clarity in Order of Appearance in Target Map 1 and 2.

Articulatory measure	Phonological characteristic	Target Map 1	Target Map 2
Word duration	begins and ends with stop consonant	camera shop	bake shop
Word duration/ word clarity rating	begins and ends with stop consonant, t/d reduction	cobbled street	crumpled spaceship
Word clarity rating	t/d reduction, place assimilation	white mountain	white cottage
Word clarity rating	t/d reduction	lost steps	mountain stream
Word duration/ word clarity rating	begins and ends with stop consonant, place assimilation	gurgling brook	broken gate
Word duration	begins and ends with stop consonant	beech forest	pine forest
Word clarity rating	t/d reduction	round rocks	sandstone cliffs
Word clarity rating	t/d reduction, place assimilation	ruined monastery	eight monkeys
Word duration/ word clarity rating	begins and ends with stop consonant, t/d reduction	great lake	great rock
Word duration	begins and ends with stop consonant	potting shed	tool shed

Table 2

Mean Values and Standard Errors for General Measures of Directors' Instructions for Target Map 1 and Target Map 2 Relative to Matchers' Informational Needs.

	Target Map 1		Target Map 2	
	High needs	Low needs	Low needs	High Needs
Extra landmarks mentioned	2.63 (.53)	3.75 (.41)	2.63 (.57)	4.13 (.30)
Words used by director	354.25 (95.56)	279.38 (53.25)	302.38 (65.02)	303.88 (51.93)
Duration of interaction	167.63 (45.53)	106.12 (15.72)	136.87 (30.78)	123.00 (20.27)

Table 3

Gender Composition Within Conversational Dyads for Every Experimental Condition in Experiment 2a and Experiment 2b.

	First storyteller				Second storyteller			
	Expects stories unknown		Expects stories known		Expects stories unknown		Expects stories known	
Addressees	female	male	female	male	female	male	female	male
Experiment 2a								
females	11	2	7	4	8	3	8	5
males	4	3	6	3	4	2	5	2
Experiment 2b								
females	5	4	8	4	5	7	4	5
males	7	4	2	4	6	0	10	1

Table 4.
Coding Schemes and Examples (Based on Lockridge & Brennan, 2002).

Coding Categories		Example
Within clause	(1) After the verb	Adolph *stabbed* the man with a *knife*.
		Jill *lay* down on her *cot*.
	(2) Before the verb	Adolph used a *knife* to *stab* the man.
		She used *chopsticks* to *eat* rice.
	(3) Incorporated into the verb	Adolph *knifed* the man.
		Sam was working on his room, *hammering* some nails in.
Separate clause	(4) After the verb	Adolph *stabbed* the man. He used an icepick.
		Sam *pounded* the nails. He used a *mallet*.
	(5) Before the verb	Adolph had a *knife.* He *stabbed* the man.
		He put water in the *kettle*, and let it *boil.*
(6) Implicit mention	Mentioned only in conjunction with an action subsequent to the target action, or at the end of the story.	He wiped he blood off the *knife.*
		The police investigators found his fingerprints all over the *knife.*
(7) Other	Not in any of the previous categories	
(8) No mention	Target instrument was not mentioned.	

Table 5.
Timing of Feedback Response to Typical and Atypical Instruments Mentioned Within the Same Clause, After Target Action (in Percentages) for Each Condition in Experiment 2a and 2b.

	Speaker expects stories to be unknown				Speaker expects stories to be known			
	Stories unknown		Stories known		Stories unknown		Stories known	
Timing of Feedback	Typical	Atypical	Typical	Atypical	Typical	Atypical	Typical	Atypical
Experiment 2a								
Directly before target instrument	0.00	11.80	16.7	26.70	0.00	12.50	7.70	10.50
Directly after target instrument	100.00	70.60	50	60.00	100.00	62.50	76.90	78.90
After target instrument clause	0.00	17.60	33.3	13.30	0.00	25.00	15.40	10.50
Total	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
N	4	4	6	15	7	8	13	19
Experiment 2b								
Directly before target instrument	9.09	9.09	14.29	35.00	0.00	4.76	0.00	0.00
Directly after target instrument	63.64	68.18	71.43	60.00	71.43	85.71	100.00	87.50
After target instrument clause	27.27	22.73	14.29	5.00	28.57	11.11	0.00	12.50
Total	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
N	11	22	7	20	7	21	6	8

Table 6.

Type of Feedback Response to Typical and Atypical Instruments for Each Condition (in percentages) in Experiment 2a and 2b.

	Speaker expects stories to be unknown				Speaker expects stories to be known			
	Stories unknown		Stories known		Stories unknown		Stories known	
Type of Feedback	Typical	Atypical	Typical	Atypical	Typical	Atypical	Typical	Atypical
Experiment 2a								
Laughter	10.30	7.30	14.3	4.90	9.10	4.70	2.00	12.50
Smile	0.00	4.90	5.7	4.90	6.10	14.00	7.80	8.30
Facial expression	20.70	14.60	0	7.30	12.10	11.60	3.90	2.10
Total, specific response	31.00	26.80	20.00	17.10	27.30	30.30	13.70	22.90
Nods (generic response)	69.00	73.20	68.6	80.50	60.60	69.80	82.40	72.90
Vocalisation	0.00	0.00	8.6	2.40	12.10	0.00	3.90	4.20
Other	0.00	0.00	2.9	0.00	0.00	0.00	0.00	0.00
Total	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Experiment 2b								
Laughter	2.60	6.00	11.60	8.20	9.70	17.00	7.00	5.40
Smile	0.00	2.00	0.00	4.10	3.20	3.80	4.70	0.00
Facial expression	5.30	4.00	4.70	6.10	0.00	7.50	0.00	0.00
Total, specific response	7.90	12.00	16.30	18.40	12.90	28.30	11.70	5.40
Nods (generic response)	84.20	76.00	69.80	71.40	61.30	54.70	81.40	81.10
Vocalisation	7.90	12.00	14.00	10.20	25.80	17.00	7.00	13.50
Other	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Total	100.00	100.00	100.00	100.00	100.00	100.00	100.10	100.00

Table 7

Percentages of Explicit Mention for Typical and Atypical Instruments for Each Condition in Experiment 2a.

	Speaker expects stories to be unknown				Speaker expects stories to be known			
	Addressee hears stories first time		Addressee hears stories second time		Addressee hears stories first time		Addressee hears stories second time	
Category	Typical	Atypical	Typical	Atypical	Typical	Atypical	Typical	Atypical
Explicit mention								
Within clause								
After the verb	24.17	32.50	21.67	37.50	25.83	33.33	28.33	31.67
Before the verb	1.67	10.00	1.67	5.83	1.67	8.33	0.83	1.67
Incorporated into verb	1.67	1.67	1.67	1.67	1.67	0.00	3.33	1.67
Total, within clause	27.50	44.17	25.00	45.00	29.17	41.67	32.50	35.00
Separate clause								
After the verb	1.67	6.67	4.17	6.67	3.33	9.17	2.50	4.17
Before the verb	4.17	7.50	5.83	7.50	7.50	7.50	5.00	9.17
Total, explicit mention	33.33	58.33	35.00	59.17	40.00	58.33	40.00	48.33
Implicit mention	51.67	26.67	52.50	29.17	47.50	28.33	51.67	40.00
No mention	9.17	6.67	10.83	8.33	12.50	6.67	8.33	2.50
Other	5.83	8.33	1.67	4.17	0.00	6.67	0.00	9.17
Total	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Average number words per story (standard deviation)	43.58 (11.70)	44.25 (12.42)	46.71 (14.06)	46.84 (14.61)	43.72 (13.37)	45.85 (12.32)	47.19 (21.24)	45.53 (20.74)

Table 8
Storytellers' Answers to Exit Questionnaires (in Percentages) for Experiment 2a and 2b.

		Speaker expects stories to be unknown		Speaker expects stories to be known	
Question		Stories unknown	Stories known	Stories unknown	Stories known
Experiment 2a					
Do you think your partner has heard the stories before?	Yes	10.00	20.00	75.00	75.00
	No	65.00	65.00	15.00	5.00
	Ambiguous	25.00	15.00	10.00	20.00
Do you think your partner might have been part of our research group (an experimental confederate)?	Yes	35.00	35.00	45.00	45.00
	No	50.00	45.00	45.00	25.00
	Ambiguous	15.00	20.00	10.00	30.00
Experiment 2b					
Do you think your partner has heard the stories before?	Yes	20.00	44.44	52.63	52.63
	No	70.00	38.89	42.11	26.32
	Ambiguous	10.00	16.67	5.26	21.05
Do you think your partner might have been part of our research group (an experimental confederate)?	Yes	35.00	57.89	22.22	68.42
	No	50.00	15.79	50.00	15.79
	Ambiguous	15.00	26.32	27.78	15.79

Table 9
Percentages of Explicit Mention for Typical and Atypical Instruments for Each Condition in Experiment 2b.

	Speaker expects stories to be unknown				Speaker expects stories to be known			
	Addressee hears stories first time		Addressee hears stories fourth time		Addressee hears stories first time		Addressee hears stories fourth time	
Category	Typical	Atypical	Typical	Atypical	Typical	Atypical	Typical	Atypical
Explicit mention							120.00	119.00
Within clause								
After the verb	30.25	33.33	25.93	36.11	20.37	31.48	20.83	31.09
Before the verb	2.52	5.83	1.85	5.56	0.00	4.63	2.50	5.04
Incorporated into verb	1.68	2.50	3.70	1.85	0.93	0.93	2.50	0.84
Total	34.45	41.67	31.48	43.52	21.30	37.04	25.83	36.97
Separate clause								
After the verb	3.36	5.00	0.93	1.85	0.93	4.63	1.67	3.36
Before the verb	4.20	8.33	2.78	9.26	11.11	13.89	5.00	8.40
Total, explicit mention	42.02	55.00	35.19	54.63	33.33	55.56	32.50	48.74
Implicit mention	47.90	35.00	53.70	32.41	45.37	25.93	53.33	31.09
No mention	7.56	5.83	9.26	7.41	12.96	10.19	11.67	10.92
Other	2.52	4.17	1.85	5.56	8.33	8.33	2.50	9.24
Total	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Average number words used per story (standard deviation)	44.69	45.63	44.99	47.17	48.85	51.21	44.07	43.00

Table 10
Changes over time in mention of first and last atypical instruments in Experiment 2b.

	Speaker expects stories to be unknown				Speaker expects stories to be known			
	Addressee hears stories first time		Addressee hears stories fourth time		Addressee hears stories first time		Addressee hears stories fourth time	
Type of Feedback	First atypical story	Last atypical story	First atypical story	Last atypical story	First atypical story	Last atypical story	First atypical story	Last atypical story
Percentage explicit mention	60.00	50.00	61.54	44.44	80.00	50.00	71.43	42.11
Percentage definite article	71.43	60.00	69.23	77.78	60.00	56.25	50.00	68.42
Percentage within clause mention	100.00	100.00	100.00	87.50	100.00	66.67	88.89	100.00

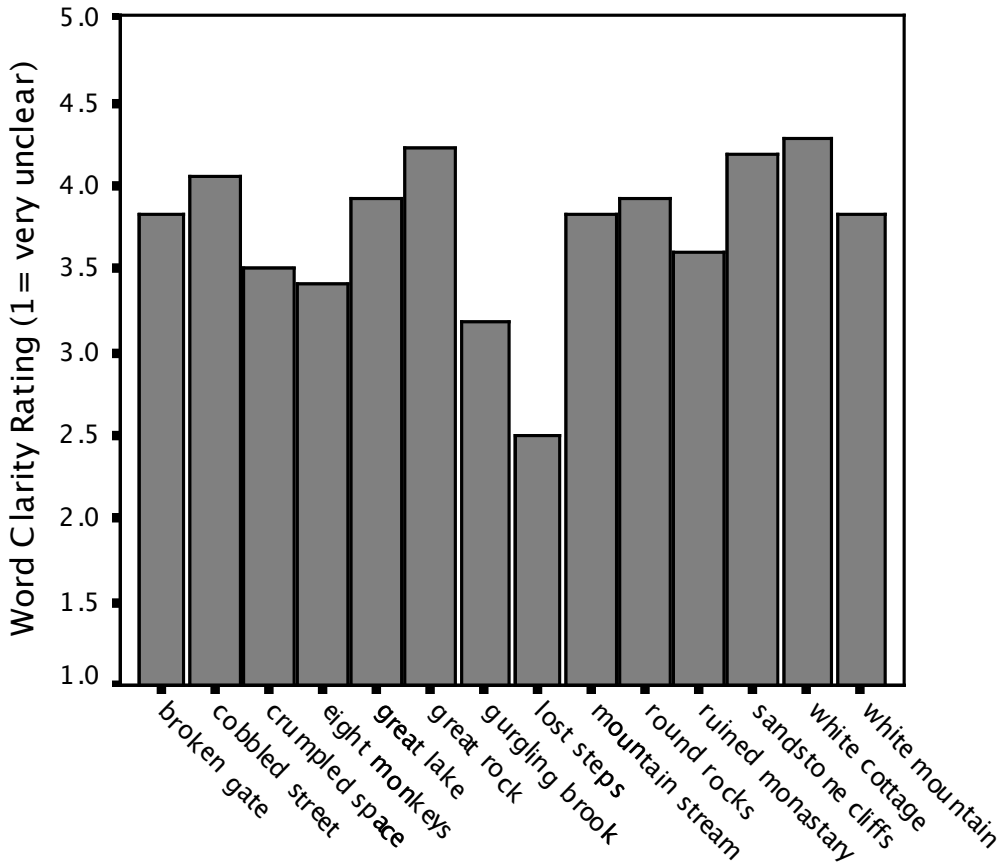


Figure 1. Mean ratings of clarity for each landmark individually, Experiment 1.

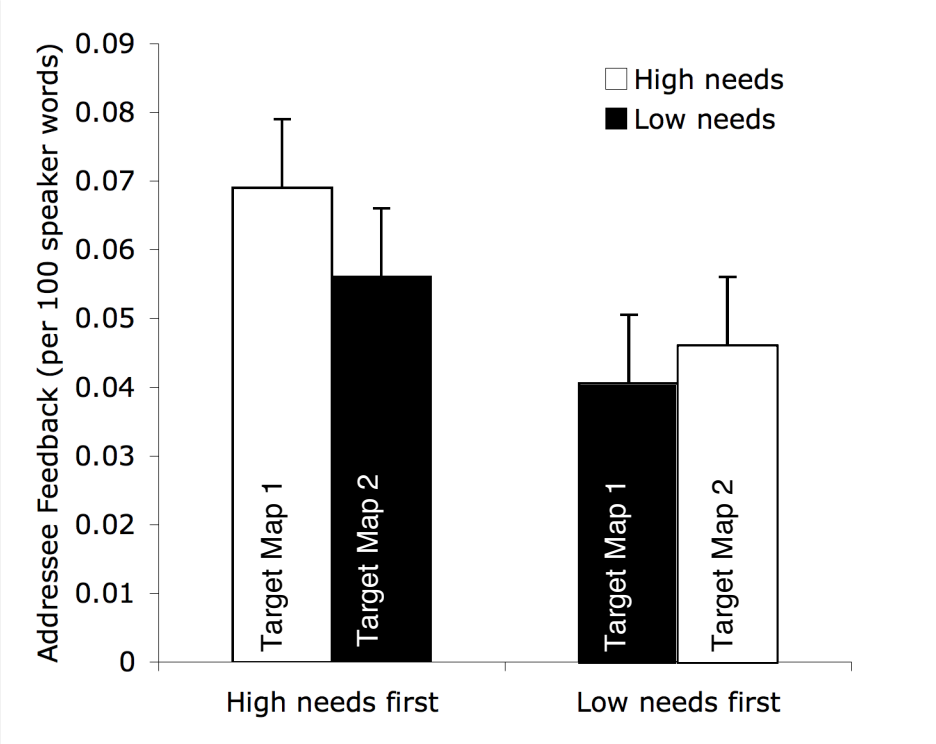


Figure 2. Mean values and standard errors for number of matcher feedback per 100 words by director relative to matchers' informational needs and target map in Experiment 1.

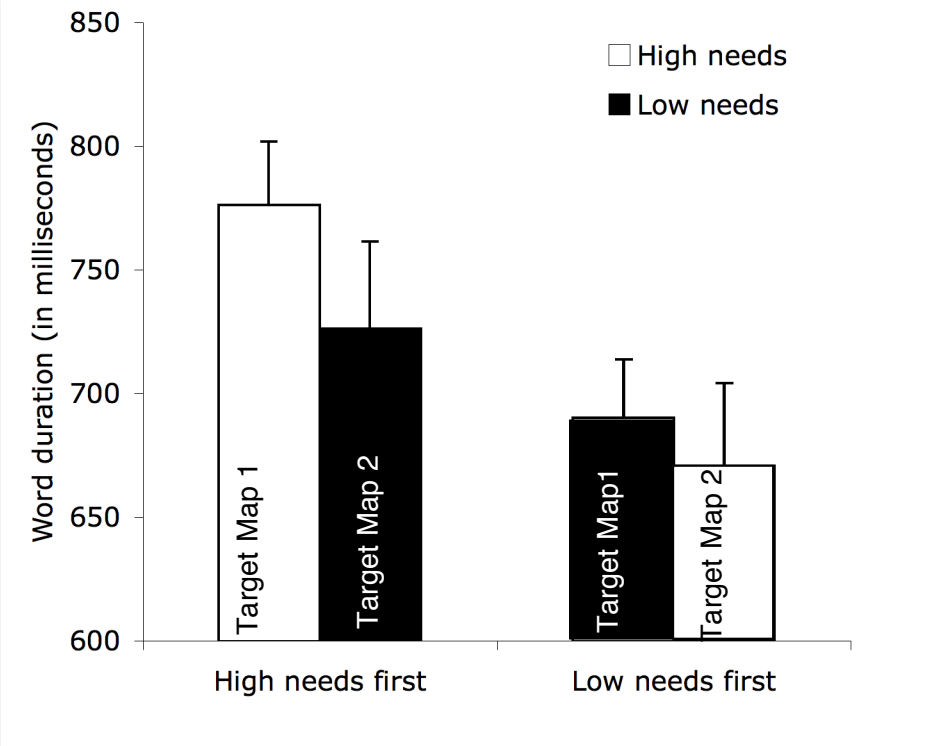


Figure 3. Mean values and standard errors for duration of target landmarks in milliseconds relative to matchers' informational needs and target map in Experiment 1.

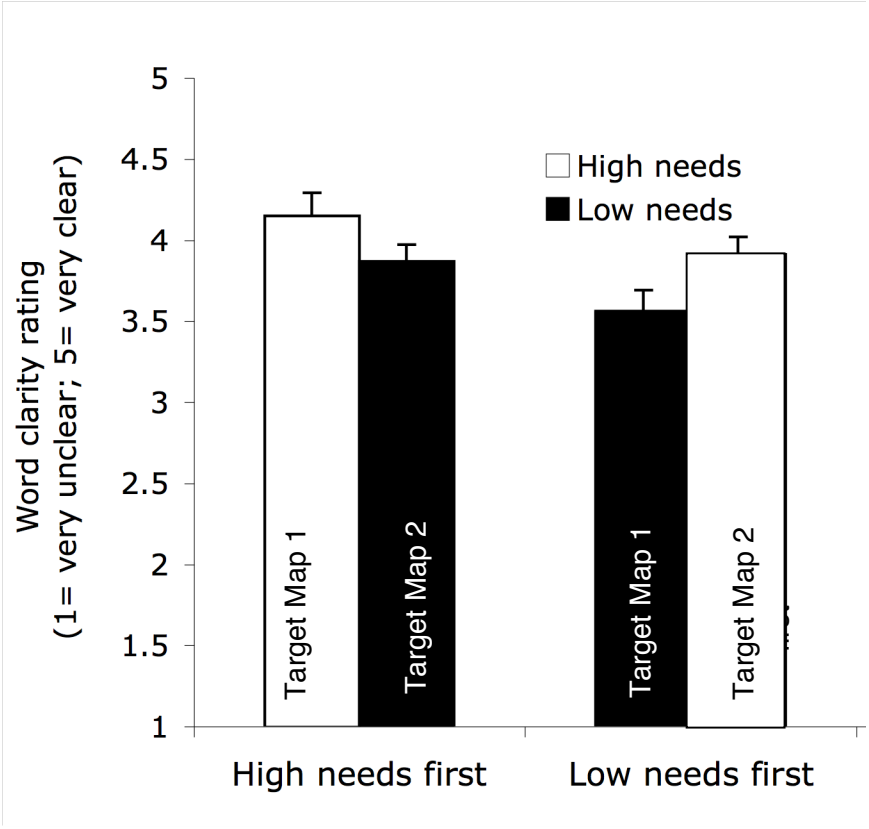
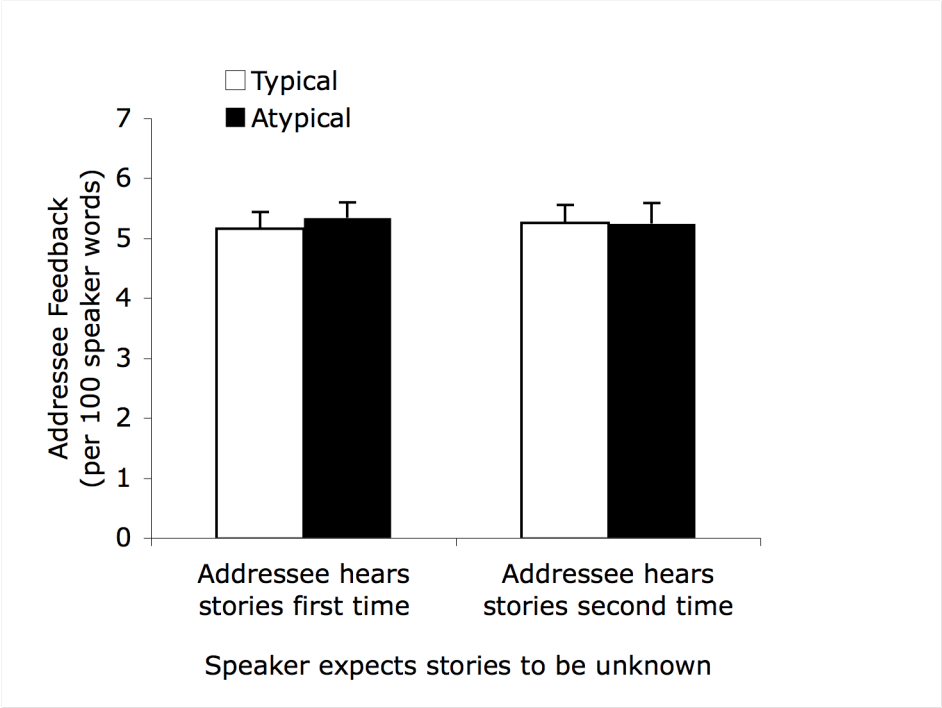


Figure 4. Mean values and standard errors for word clarity rating relative to matchers' informational needs and target map. in Experiment 1



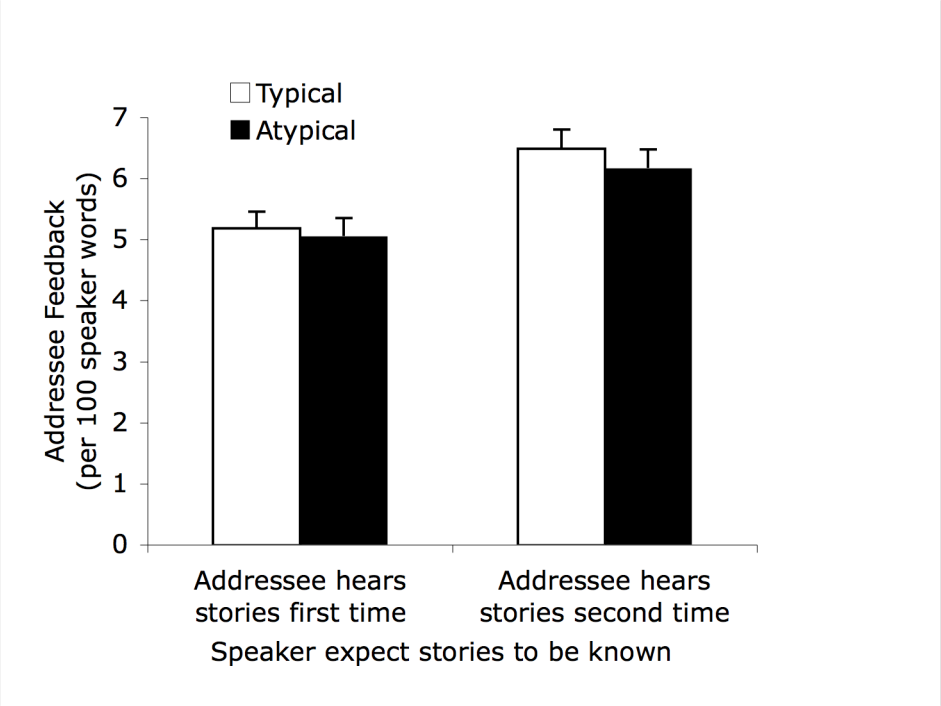
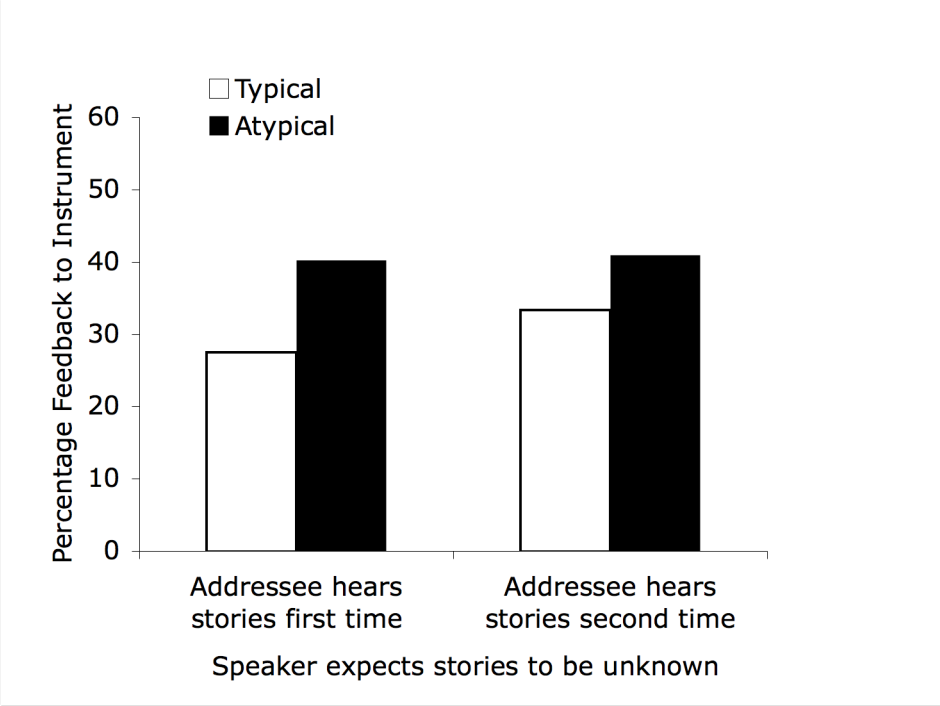


Figure 5a & 5b. Mean values and standard errors for number of matcher feedback (per 100 words by storyteller) relative to typicality of instrument, addressees' knowledge and storytellers' expectations in Experiment 1.



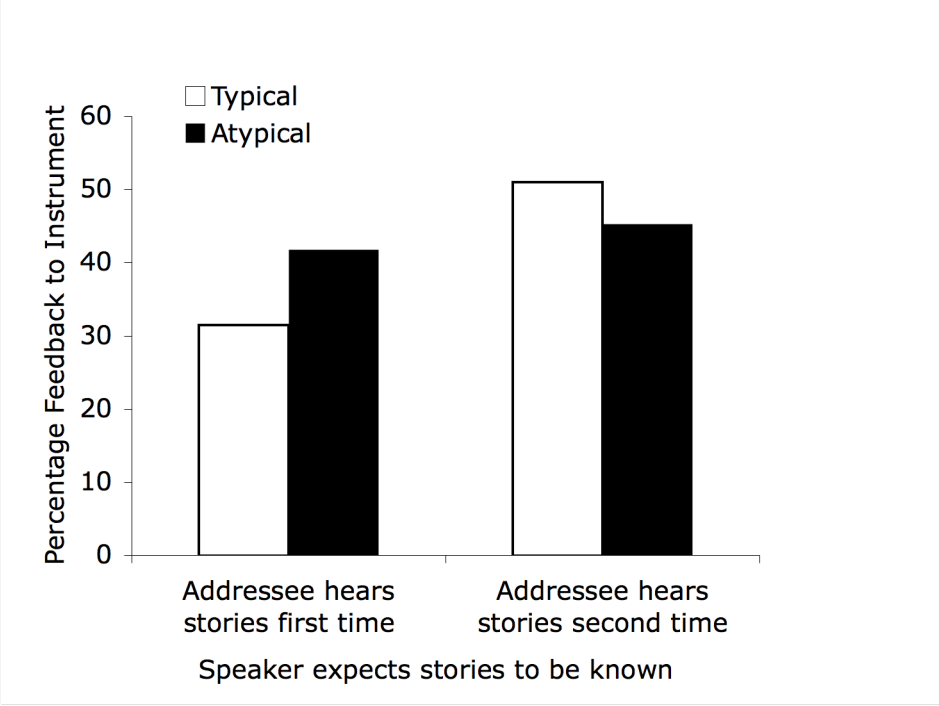
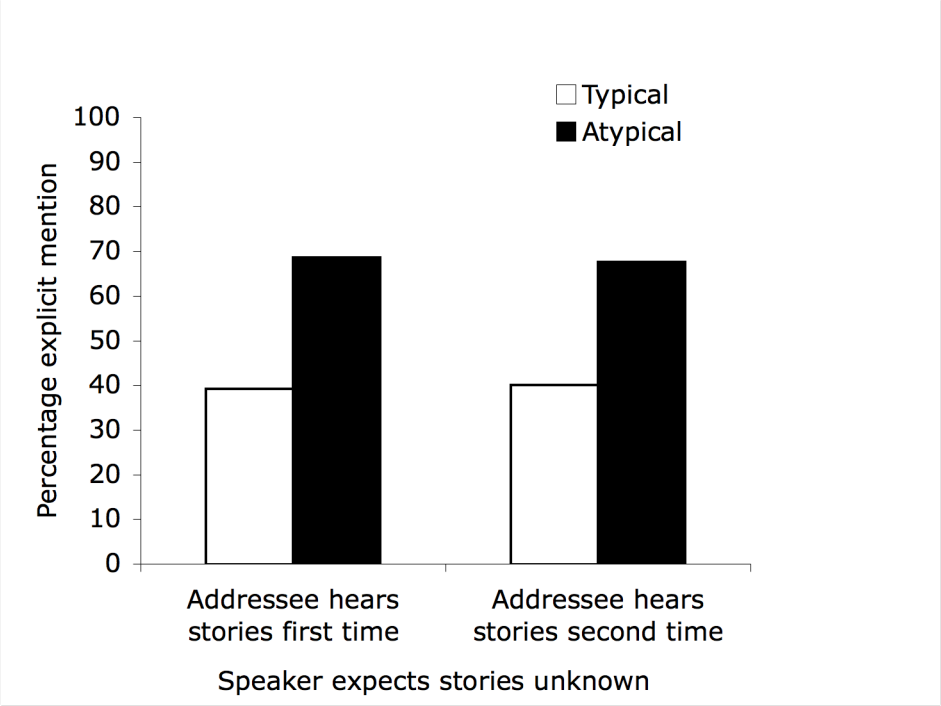


Figure 6a & 6b. Percentage of mentioned instruments accompanied by feedback relative to typicality of instrument, addressees' knowledge and storytellers' expectations in Experiment 2a.



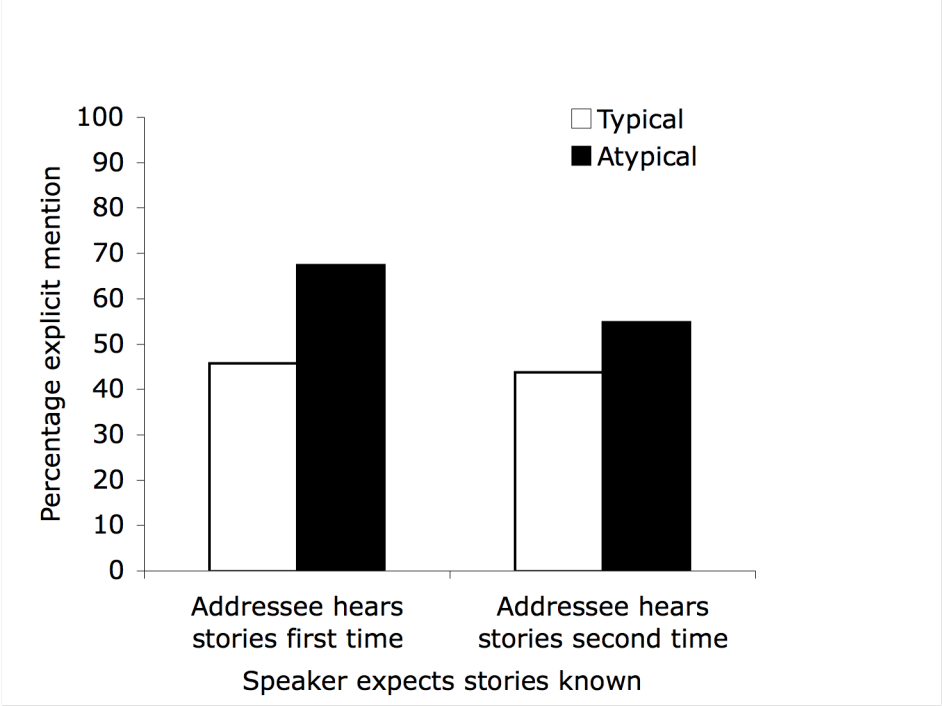
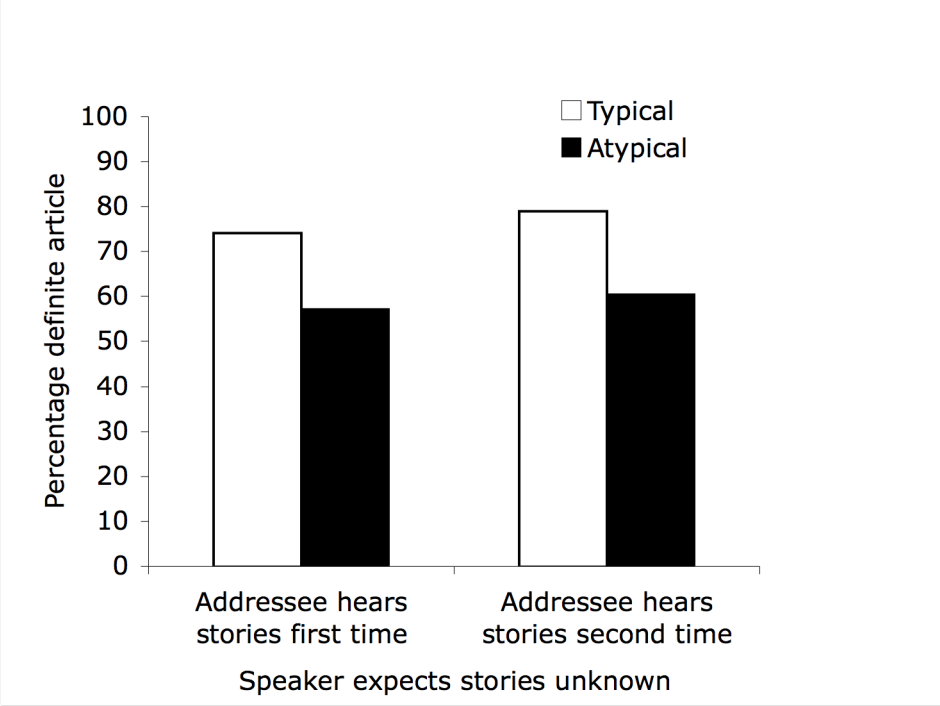


Figure 7a & 7b. Percentage of target instruments mentioned explicitly relative to typicality of instrument, addressees' knowledge and storytellers' expectations.



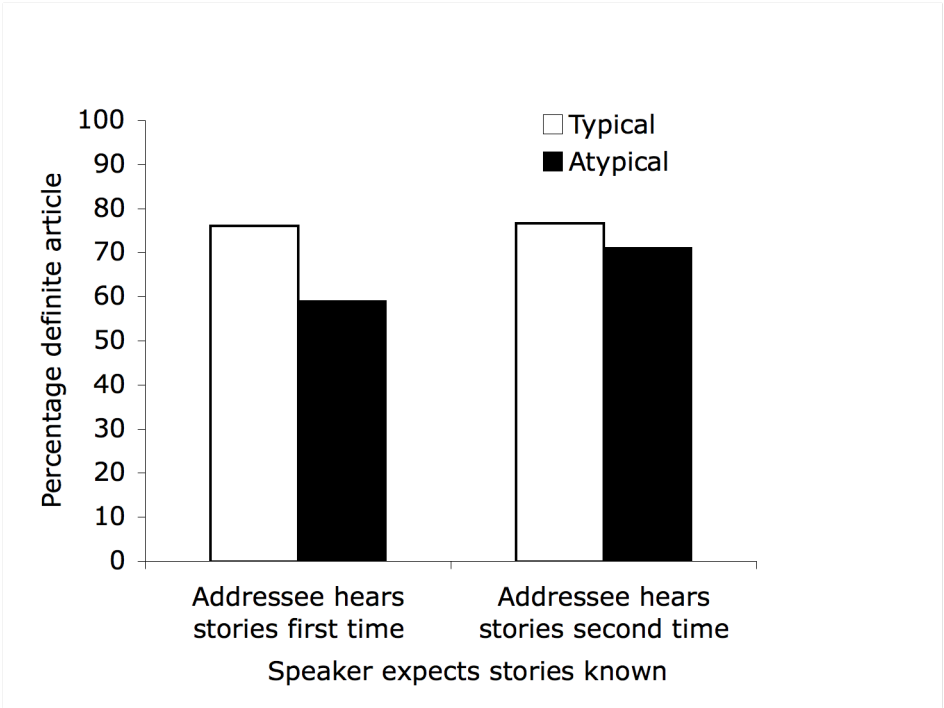


Figure 8a & 8b. Percentage of target instrument introduced with definite article relative to typicality of instrument, addressees' knowledge and storytellers' expectations in Experiment 2a

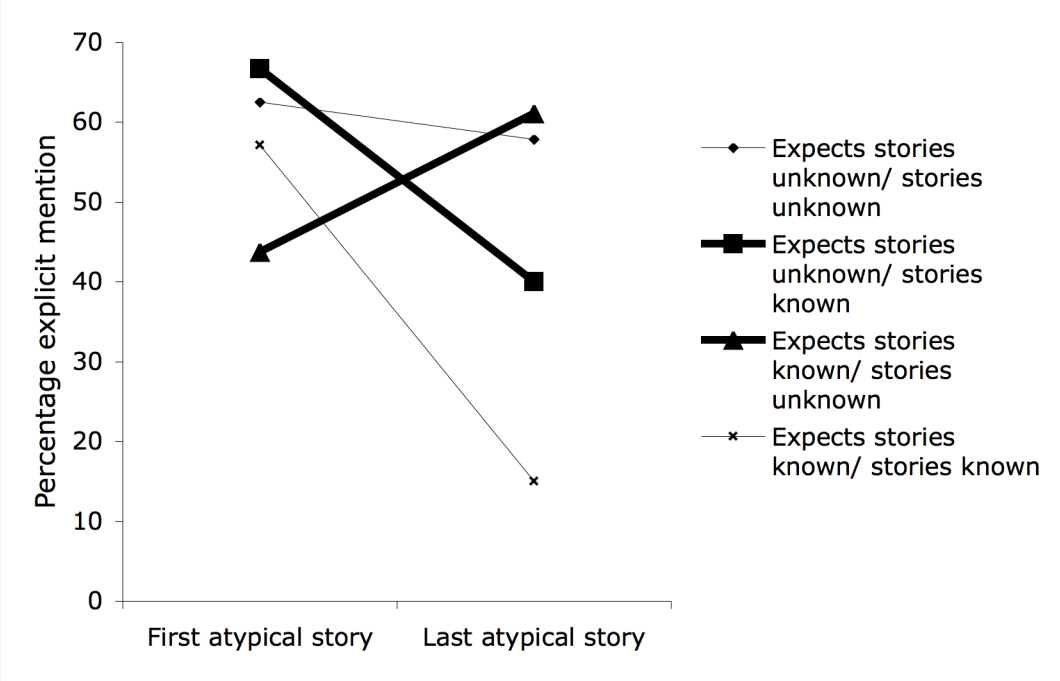


Figure 9. Percentage of first and last atypical stories mentioning instruments explicitly in Experiment 2a.

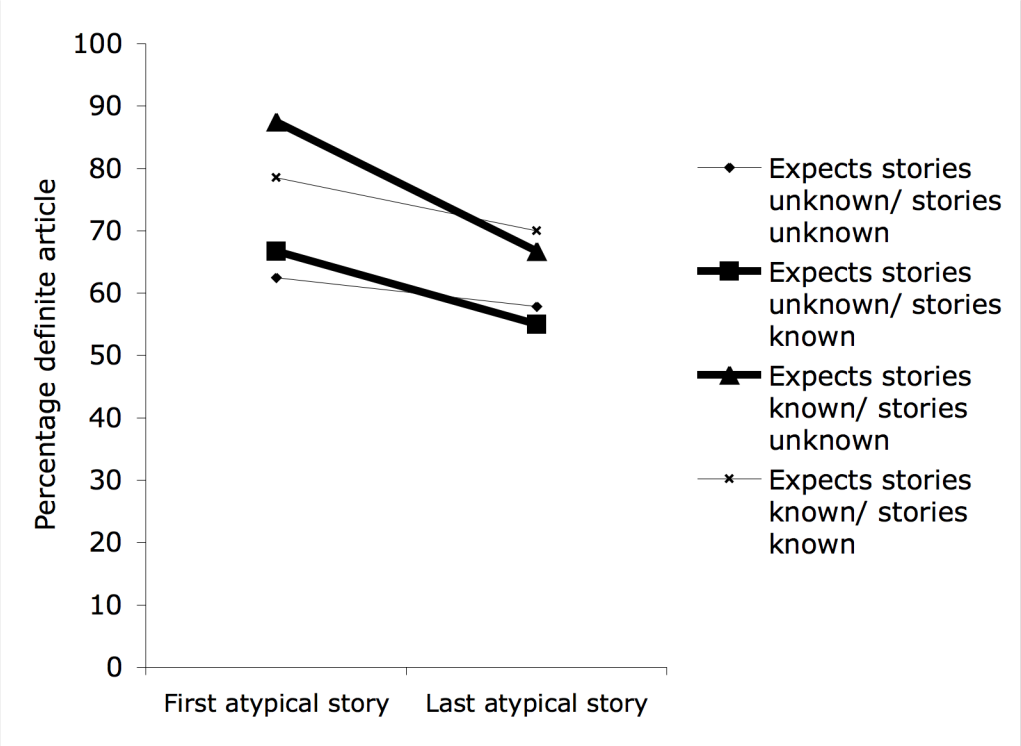


Figure 10. Percentage of first and last atypical stories mentioning instruments with definite article in Experiment 2a.

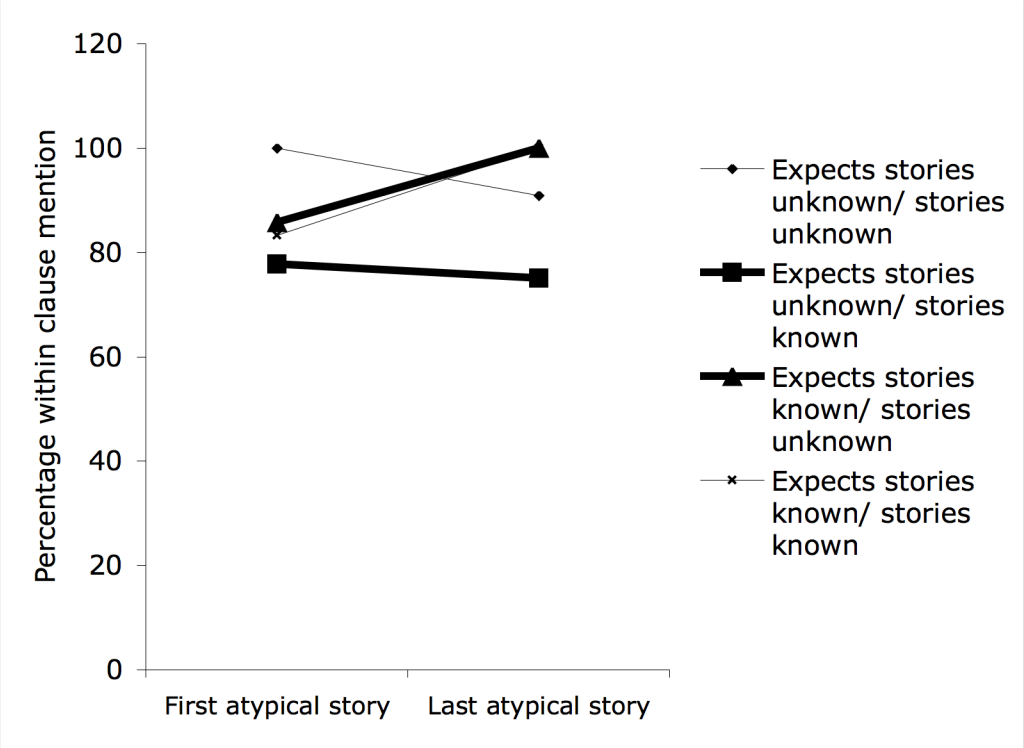
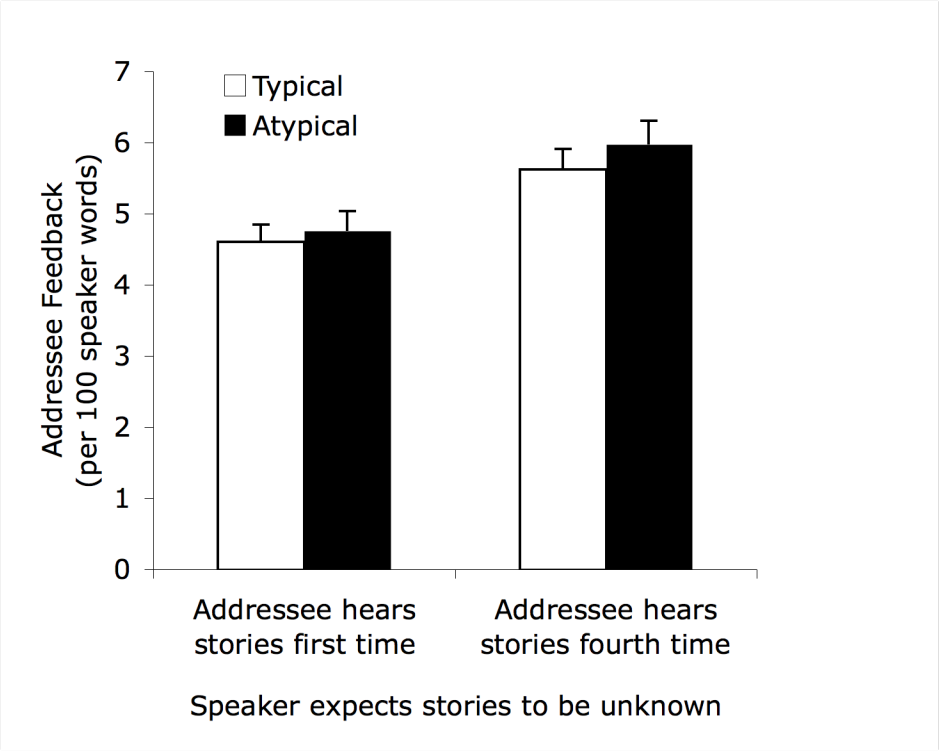


Figure 11. Percentage of first and last atypical stories mentioning instruments within the same clause as target action in Experiment 2a.



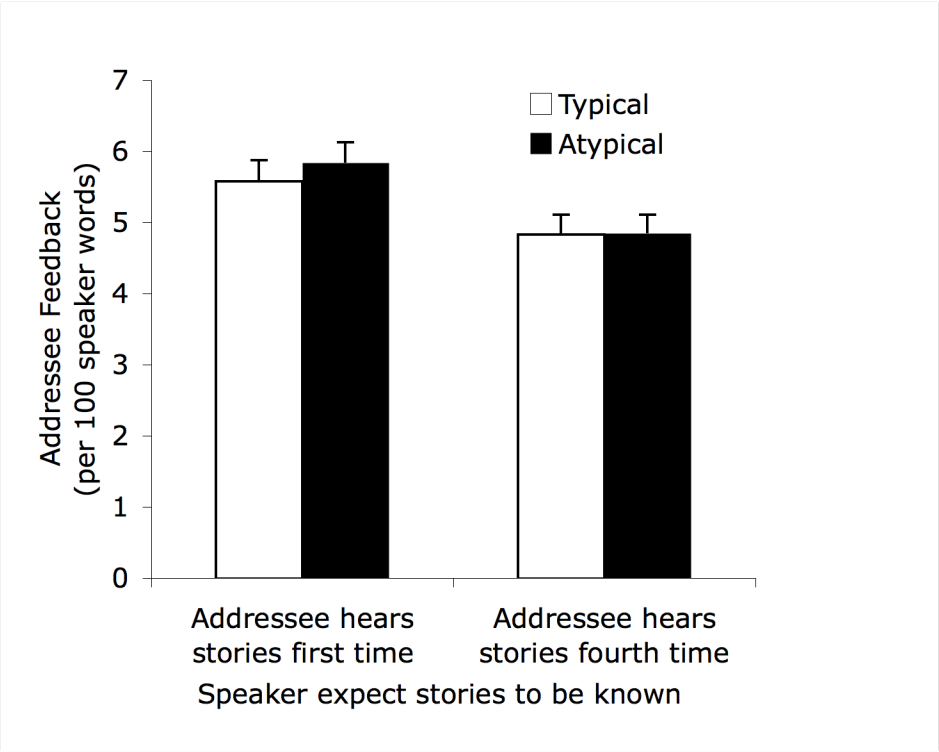
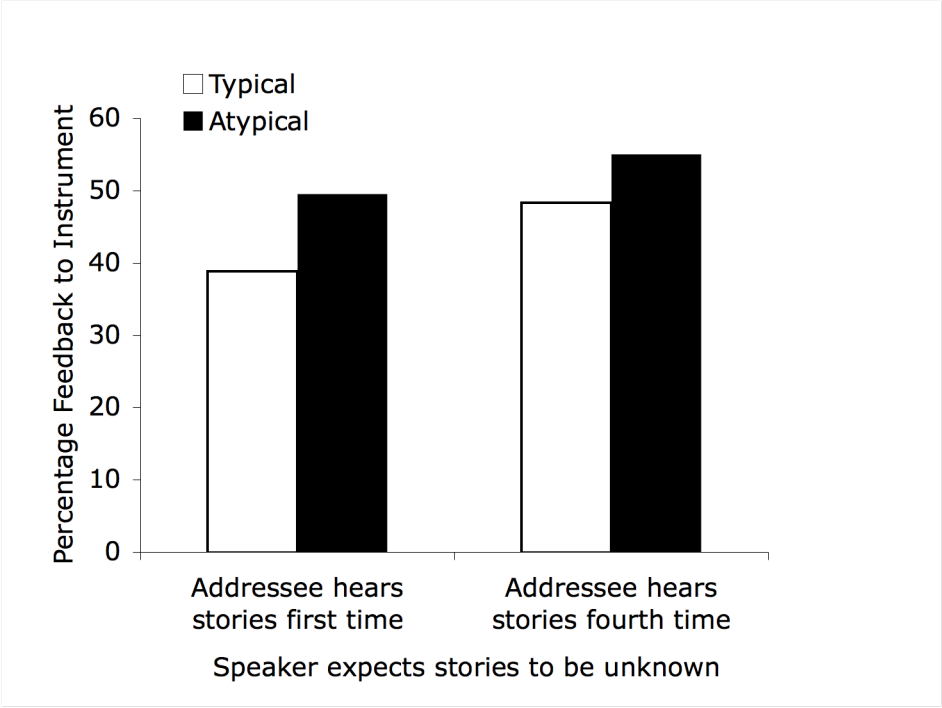


Figure 12a & 12b. Mean values and standard errors for number of matcher feedback (per 100 words by storyteller) relative to typicality of instrument, addressees' knowledge and storytellers' expectations in Experiment 2b



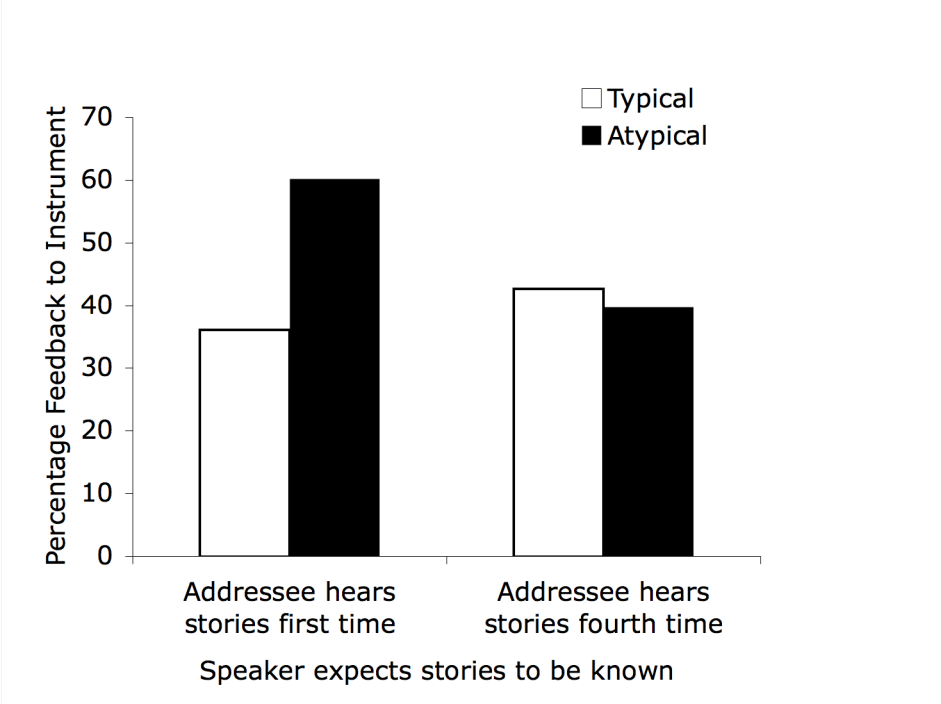
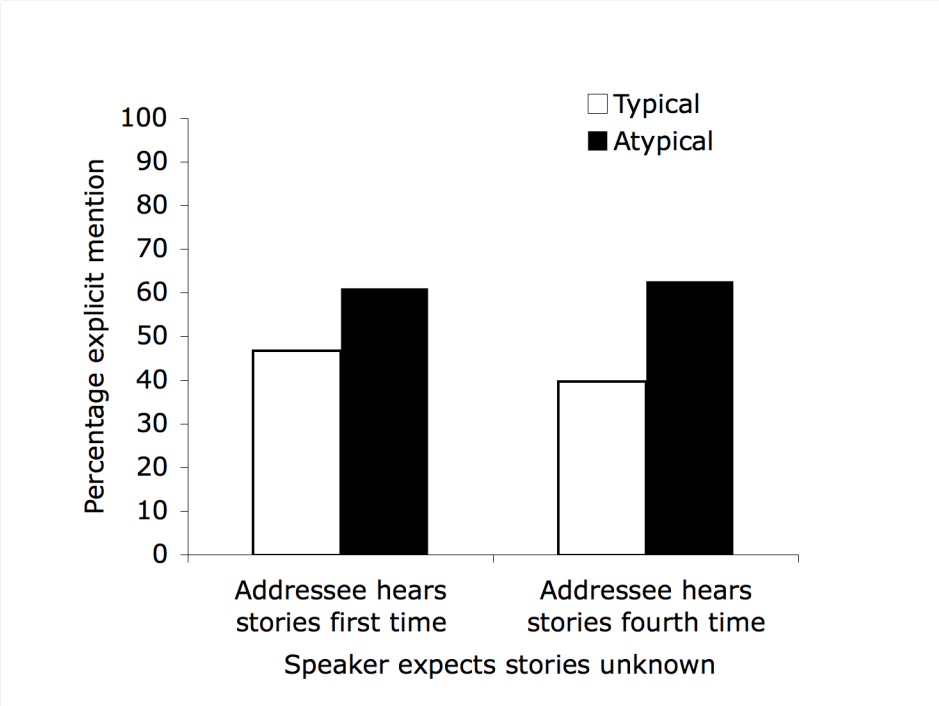
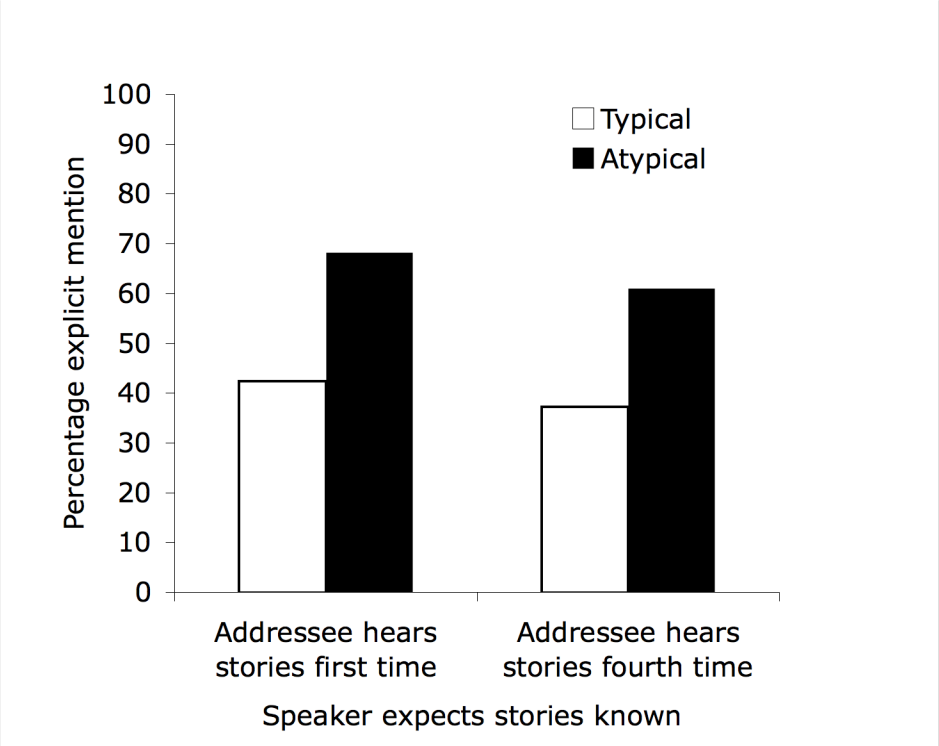
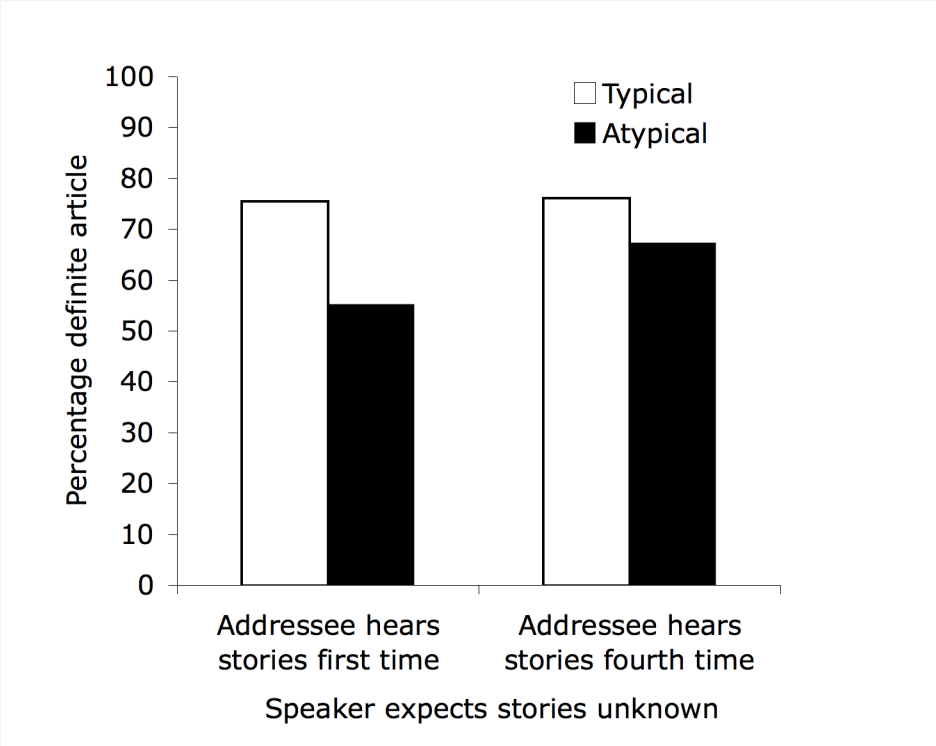


Figure 13a & 13b. Percentage of mentioned instrument accompanied by target feedback relative to typicality of instrument, addressees' knowledge and storytellers' expectations in Experiment 2b.





Figures 14 a & 14b. Percentage of target instruments mentioned explicitly relative to typicality of instrument, addressees' knowledge and storytellers' expectations in Experiment 2b.



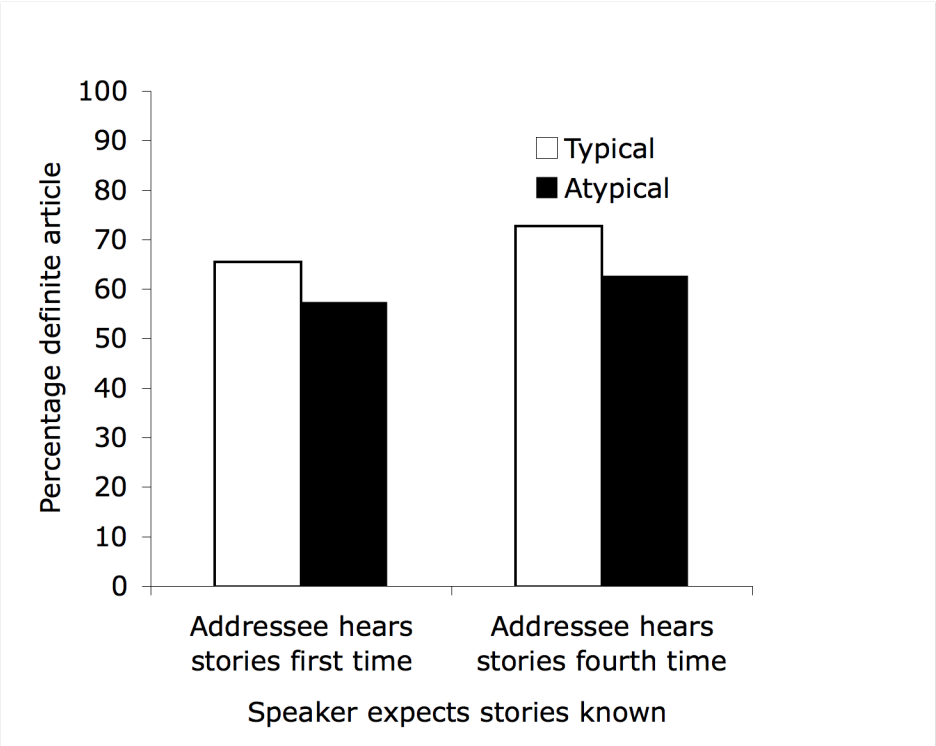


Figure 15a & 15b. Percentage of target instrument introduced with definite article relative to typicality of instrument, addressees' knowledge and storytellers' expectations in Experiment 2a.

References

Anderson, A. H., Bader, M., Boyle, E., Bard, E. G., Doherty, G., Garrod, S., Isard, S. D., Kowtko, J., MacAllister, J., Miller, J., Sotillo, C., Thompson, H. S., & Weinert, R. (1991). The H.C.R.C. Map Task Corpus. *Language and Speech*, 34, 351- 366.

Anderson, A. H., Bard, E. G., Sotillo, C., Newlands, A., & Doherty-Sneddon, G. (1997). Limited visual control of the intelligibility of speech in face-to-face dialogue. *Perception and Psychophysics,* 59, 580- 592.

Anderson, A., Clark, A., & Mullin, J. (1991). Introducing information in dialogues: How young speakers refer and how young listeners respond. *Journal of Child Language, 18,* 663- 687.

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

Bard, E. G., Aylett, M. (2001). Accessibility, duration, and modeling the listener in spoken dialogue. *Gtelog 2000*, *Fourth workshop on the semantics and pragmatics of dialogue.* Gtalog, Sweden: Goethenberg University.

Barr, D. J., & Keysar, B. (2002). Anchoring comprehension in linguistic precedents. *Journal of Memory and Language*, 46, 391-418.

Barr, D. J., & Keysar, B. (2005). Mindreading in an exotic case: The normal adult human. In Malle, B. F., & Hodges, S. D. (Eds.). *Other minds: How humans bridge the divide between self and other.* New York: Guilford Press. pp. 271-283.

Bavelas, J. B., Coates, L., Johnson, T. (2000). Listeners as Co-Narrators. *Journal of Personality and Social Psychology,* 6, 941- 952.

Bavelas, J. B., Coates, L., Johnson, T. (2002). Listener Responses as Collaborative Process: The Role of Gaze. *Journal of Communication,* 52, 566- 580.

Bavelas, J. B., Gerwing, J., Sutton, C., & Prevost, D. (2008). Gesturing on the telephone: Independent effects of dialogue and visibility. *Journal of Memory and Language, 58*, 495-520.

Bell, A. (1984) Language Style as Audience Design. In: Coupland, N. & A. Jaworski (Eds.) *Sociolinguistics: A Reader and Coursebook*, pp. 240-50. New York: St Mattin's Press Inc.

Bernolet, S., Hartsuiker, R. J., & Pickering, M. J. (2007). Shared syntactic representations in bilinguals: Evidence for the role of word-order repetition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 931- 949.

Bock, J. K. (1996). Language production: Methods and methodologies. *Psychonomic Bulletin and Review*, 3, 395-421.

Bock, J. K., & Levelt, W. J. M. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (Ed). *Language processing* (pp. 945-984). Elsevier/ Academic Press.

Boersma, P., & Weenink, D. (2006). Praat: Doing phonetics by computer (Version 4.6.09) [Computer program]. Retrieved date, from http://www.praat.org/

Bortfeld, H., & Brennan, S. E. (1997). Use and acquisition of idiomatic expressions in referring by native and non-native speakers. *Discourse Processes, 23,* 119- 147.

Branigan, H. P., Pickering, M. J., McLean, J. F., & Cleland, A. A. (2007). Syntactic alignment and participant role in dialogue. *Cognition*, 104, 163-197.

Brennan, S. E. (1991). Conversation with and through computers. *User Modeling and User-Adapted Interaction, 1*, 67-86.

Brennan, S. E. (2005). How conversation is shaped by visual and spoken evidence. In J. Trueswell & M. Tanenhaus (Eds.), *Approaches to studying world-situated language use: Bridging the language-as-product and language-action traditions* (pp. 95-129). Cambridge, MA: MIT Press.

Brennan, S. E., & Hanna, J. E. (2009). Partner-specific adaptation in dialogue. *Topics in Cognitive Science*, 1, 274-291.

Brennan, S. E., & William, M. (1995). The feeling of another's knowing: Prosody and filled pauses as cues to listeners about the metacognitive states of speakers. *Journal of Memory and Language*, 34, 383-398.

Brennan, S. E., Chen, X., Dickinson, C., Neider, M., & Zelinsky, G. (2008). Coordinating cognition: The costs and benefits of shared gaze during collaborative search. *Cognition*, 106, 1465-1477.

Brennan, S. E., Galati, A., & Kuhlen, A. (2010). Two minds, one dialog: Coordinating speaking and understanding. In B. Ross (Ed.), Psychology of Learning and Motivation, Vol. 53 (pp. 301- 344). Academic Press/Elsevier.

Brown, P. M., Dell, G. S. (1987). Adapting production to comprehension: The explicit mention of instruments. *Cognitive Psycholog*, 19, 441- 472.

Brown, G., Anderson, A., H., Yule, G., & Shillcock, R. (1984). *Teaching talk.* Cambridge, UK: Cambridge University Press.

Brown-Schmidt, S., & Tanenhaus, M. K. (2008). Real-time investigation of referential domains in unscripted conversation: A targeted language game approach. *Cognitive Science*, 643 – 684.

Brown-Schmidt, S., Gunlogson, C., & Tanenhaus, M. K. (2008). Addressees distinguish shared from private information when interpreting questions during interactive conversation. *Cognition*, 107, 1122-1134.

Bruehl, D., U. (1970). Systematic variation in the clarity of demand characteristics in an experiment employing a confederate. *Psychological Reports*, 27, 55-60.

Chovil, N. (1991). Social determinants of facial displays. *Journal of Nonverbal Behavior*, 15, 141- 154.

Clark, H. H. (1996). *Using Language*. Cambridge. MA: Cambridge University Press.

Clark, H. H., & Brennan, S. A. (1991). Grounding in communication. In L.B. Resnick, J.M. Levine, & S.D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 127-149). Washington: APA Books.

Clark, H. H., & Carlson, T. B. (1982). Hearers and speech acts. *Language*, 332- 373.

Clark, H. H., & Gerrig, R. J. (1990). Quotations as demonstrations. *Language,* 66, 764-805.

Clark, H. H., & Haviland, S. E. (1977). Comprehension and the given-new contract. In R. O. Freedle (ed.), *Discourse production and comprehension* (pp. 1- 40). Norwood, NJ: Ablex Publishing.

Clark, H. H. & Krych, M. A. (2004). Speaking while monitoring addressees for understanding. *Journal of Memory and Language*, 50(1), 62-81.

Clark, H. H., & Murphy, G. L. (1983). Audience design in meaning and reference. In J. F. LeNy and W. Kintsch (Eds.), *Language and comprehension* (pp. 287-299). Amsterdam: North-Holland Publishing Co.

Clark, H. H., & Schaefer, E. F. (1989). Contributing to discourse. *Cognitive Science, 13,* 259- 294.

Clark, H. H., & Wilkes- Gibbs, D. (1986). Referring as a collaborative process. *Cognition, 22,* 1- 39.

Davies, B., & Harré, R. (1990). Positioning: The discursive production of selves. *Journal for the Theory of Social Behaviour*, 20, 43-63

Duncan, S. (1973). Toward a grammar for dyadic conversation. *Semiotica*, 9, 29–47.

Duncan, S. & Fiske, D. W. (1977). *Face-to-face interaction: Research, methods, and theory*. Hillsdale, NJ: Lawrence Erlbaum.

Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. Developmental Psychology, 20, 104-113.

Ford, C. E., Thompson, S. A. (1996). Interactional units in conversation: Syntactic, intentional, and pragmatic resources for the management of turns. In E. Ochs, E. A. Schegloff, & S. A. Thompson (Eds.), *Interaction and grammar* (pp. 134- 184). Cambridge, UK: Cambridge University Press.

Fridlund, A. J. (1991). Sociality of solitary smiling: Potentiation by an implicit audience. *Journal of Personality and Social Psychology*, 20, 229- 240.

Fussell, S. R., Krauss, R. M. (1992). Coordination of knowledge in communication: Effects of speakers' assumptions about others' knowledge. *Journal of Personality and Social Psychology*, 62, 378-391.

Galati, A. (2009). Assessing common ground in conversation: The effect of linguistic and physical co-presence on early planning. *Unpublished doctoral dissertation*. State University of New York: Stony Brook, NY.

Galati, A., Brennan, S. E. (2010). Attenuating information in spoken communication: for the speaker or for the addressee? *Journal of Memory and Language,* 62, 35- 51.

Garrett, M. (1980). Levels of processing in speech production. In B. Butterworth (Ed). *Language production* (vol 1). Academic Press.

Goffman, E. (1959). The Presentation of Self in Everyday Life. New York: Doubleday Anchor Books.

Goodwin, C. (1981). *Conversational Organization. Interaction between speakers and hearers*. New York: Academic Press.

Goodwin, C. (1986). Between and within: Alternative sequential treatments of continuers and assessments. *Human Studies*, 9, 205- 217.

Hanna, J. E., & Brennan, S. E. (2007). Speakers' eye gaze disambiguates referring expressions early during face-to-face conversation. *Journal of Memory and Language, 57,* 596- 615.

Hanna, J. E., Tanenhaus, M. K., & Trueswell, J. C. (2003). The effects of common ground and perspective on domains of referential interpretation. *Journal of Memory and Language, 49,* 43- 61.

Hanna, J. E., & Tanenhaus, M. K. (2004). Pragmatic effects on reference resolution in a collaborative task: Evidence from eye movements. *Cognitive Science, 28,* 105- 115.

Haviland, S., Clark, H. H. (1974). What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior, 13,* 512- 521.

Heider, F. (1946). Social perception and phenomenal causality. *Psychological Review, 51,* 358- 374.

Heritage, J. (1984). A change-of-state token and aspects of its sequential placement. In Atkinson & Heritage (eds.), *Structure of Social Actions: Studies in Conversation Analysis* (pp. 299- 345). Cambridge: Cambridge University Press.

Howarth, B., Anderson, A. H. (2007). Introducing objects in spoken dialogue: The influence of conversational setting and cognitive load on the articulation and use of referring expressions. *Language and Cognitive Processes,* 22, 272- 296.

Isaacs, E.A., & Clark, H.H. (1987). References in conversation between experts and novices. *Journal of Experimental Psychology: General*, 116(1), 26-37.

Kaplan, R., & Bresnan, J. (1982). Lexical-functional grammar: A formal system for grammatical representation. In J. Bresnan (Ed.), *The mental representation of grammatical relations.* MIT press.

Kendon, A. (1967). Some functions of gaze-direction in social interaction. *Acta Psychologica,* 26, 22- 63.

Keysar, B., Barr, D. J., Balin, J. A., & Paek, T. S. (1998). Definite reference and mutual knowledge: Process models of common ground in comprehension. *Journal of Memory and Language, 39*, 1-20.

Keysar, B., Barr, D. J., Balin, J. A., & Brauner, J. S. (2000). Taking perspective in conversation: The role of mutual knowledge in comprehension. *Psychological Science, 11*, 32-38.

Kraljic, T., & Brennan, S. E. (2005). Prosodic disambiguation of syntactic structure: For the speaker or for the addressee? *Cognitive Psychology*, 50, 194-231.

Krauss, R. M., & Weinheimer, S. (1966). Concurrent feedback, confirmation, and the encoding of referents in verbal communications. *Journal of Personality and Social Psychology*, 4, 343–346.

Krauss, R. M., & Bricker, P. D. (1967). Effects of transmission delay and access delay on the efficiency of communication. *Journal of the Acoustical Society of America*, 41, 286-292.

Krauss, R. M., Garlock, C. M., Bricker, P. D., & McMahon, L. E. (1977). The role of audible and visible back-channel responses in interpersonal communication. *Journal of Personality and Social Psychology*, 35, 523–529.

Kraut, R. E., Lewis, S. H., & Swezey, L. W. (1982). Listener responsiveness and the coordination of conversation. *Journal of Personality and Social Psychology*, 43, 718-731.

Kronmüller, E., & Barr, D. J. (2007). Perspective-free pragmatics: Broken precedents and the recovery-from-preemption hypothesis. *Journal of Memory and Language*, 56, 436-455.

Kuhlen, A. K. (2004). Der Einfluss von Geschlechterstereotypen auf die Nutzung von transaktiven Wissenssystemen. *Unpublished master thesis.* Humboldt University: Berlin, Germany.

Kuhlen, A. K., Brennan, S. E. (in press). Anticipating distracted addressees: How speakers' expectations and addressees' feedback influence storytelling. *Discourse Processes.*

Kuhlen, A.K., & Brennan, S. E. (2008). Addressees shape speaking: When confederates may be hazardous to your data. Abstracts of the Psychonomic Society, 49th Annual Meeting (p. 6), Chicago, IL.

Kuhlen, A. K., Galati, A., Brennan, S. E. (under review). Gesturing integrates top-down and bottom-up information: Effects of speakers' expectations and addressees' feedback. *Cognitive Science.*

Levelt, W. J. M. (1989). *Speaking: From intention to articulation.* MIT Press.

Lockridge, C. B., Brennan, S. E. (2002). Addressees' needs influence speakers' early syntactic choices. *Psychonomic Bulletin & Review, 9I,* 550- 557.

Martin, J. D. (1970). Suspicion and the experimental confederate: A study of role and credibility. *Sociometry*, 33, 178-192.

Martin, J. D. (1973). Pejorative factors in the suspicion of deception. The case of 'detection' of confederates. *Journal of Social Psychology*, 91, 155-156.

Metzing, C., & Brennan, S. E. (2003). When conceptual pacts are broken: Partner-specific effects on the comprehension of referring expressions. *Journal of Psycholinguistic Research*, 25, 319-344.

Newman- Norlund, S. E., Noordzji, M. L., Newman- Norlund, R. D., Volman, I. A. C., de Ruiter, J. P., Hagoort, P., Toni, I. (2009). Recipient design in tacit communication. *Cognition, 111,* 46- 54.

Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, 17, 776-783.

Pasupathi, M., Stallworth, L. M., & Murdoch, K. (1998). How what we tell becomes what we know: Listener effects on speakers' long-term memory for events. *Discourse Processes*, 26, 1–25.

Pollard, & Sag, I. A. (1994). *Head-driven phrase structure grammar.* University of Chicago Press and CSLI.

Rosenberg, M. J. (1965). When dissonance fails: On eliminating evaluation apprehension from attitude measurement. *Journal of Personality and Social Psychology*, 1, 24-42.

Rosenthal, R., & Jacobson, L. (1968). Pygmalion in the classroom. *The Urban Review*, 3, 16- 20.

Rosenthal, R., & Jacobson, L. (1992). *Pygmalion in the classroom*. Expanded edition. Ew York: Irvington.

Ross, L., D., Greene, O., & House, P. (1977). The false consensus phenomenon: An attributional bias in self-perception and social perception processes. *Journal of Experimental Social Psychology, 13,* 279- 301.

Russell, A. W., Schober, M. F. (1999). How beliefs about a partner's goals affect referring in goal-discrepant conversations. *Discourse Processes*, 27, 1-33.

Sacks, H., Schegloff, E. A., Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language, 59,* 696- 735.

Samuel, A. G., Troicki, M. (1998). Articulation quality is inversely related to redundancy when children or adults have verbal control. *Journal of Memory & Language, 39,* 175-194.

Schegloff, E. A. (1981). Discourse as an interactional achievement: Some uses of uh-huh and other things that come between sentences. In D. Tannen (Ed.), *Analyzing Discourse: Text and Talk.* Georgetown University Roundtable on Language and Linguistics (pp. 71-93). Washington, DC: Georgetown University Press.

Schober, M. F., & Brennan, S. E. (2003). Processes of interactive spoken discourse: The role of the partner. In A. C. Graesser, M. A. Gernsbacher, & S. R. Goldman (Eds), *Handbook of discourse processes*. Hillsdale, NJ: Lawrence Erlbaum.

Schoonbaert, S., Hartsuiker, R., & Pickering, M. J. (2007). The representation of lexical and syntactic information in bilinguals: Evidence from syntactic priming. *Journal of Memory and Language*, 56, 153-171.

Shatz, M., & Gelman, R. (1973). The development of communication skills: Modification in the speech of young children as a function of listener. *Monographs of the Society for Research in Child Development, 38,* 1- 38.

Shotter, J., & Gergen, K. (Eds.) (1989). *Texts of identity*. London: Sage Publications.

Smith, V. L., & Clark, H. H. (1993). On the course of answering questions, *Journal of Memory and Language, 32,* 25- 38.

Snyder, M. (1984). When beliefs create reality. In L. Berkowitz (Ed.), Advances in experimental psychology, (Vol. 18, pp. 248- 305). Orlando, FL: Academic Press.

Snyder, M., & Klein, O. (2005). Construing and constructing others. On the reality and the generality of the behavioral confirmation scenario. Interaction Studies, 6, 53- 67.

Snyder, M., & Stukas, A. A. Jr. (1999). Interpersonal processes: The interplay of cognitive, motivational, and behavioral activities in social interaction. Annual Review of Psychology, 50, 273- 303.

Snyder, M., Tanke, E. D., & Berscheid, E. (1977). Social perception and interpersonal behavior: On the self-fulfilling nature of social stereotypes. *Journal of Personality and Social Psychology, 35*, 656- 666.

Stivers, T. (2008). Stance, alignment, and affiliation during storytelling: When nodding is a token of affiliation. *Research on Language & Social Interaction, 41*, 31- 57.

Stricker, L. J., Messick, S., & Jackson, D. N. (1967). Suspicion of deception: Implications for conformity research. *Journal of Personality and Social Psychology, 5*, 379-389.

Swerts, M., Krahmer, E. (2005). Audiovisual prosody and feeling of knowing. Journal of *Memory and Language*, 53, 81-94.

Tanenhaus, M. K., & Brown-Schmidt, S. (2007). Language processing in the natural world. In Morre, B. C. M., Tyler, L. K. & Marslen-Wilson, W. D. (Eds.) *The perception of speech: From sound to meaning. Philosophical Transactions of the Royal Society B: Biological Science.*

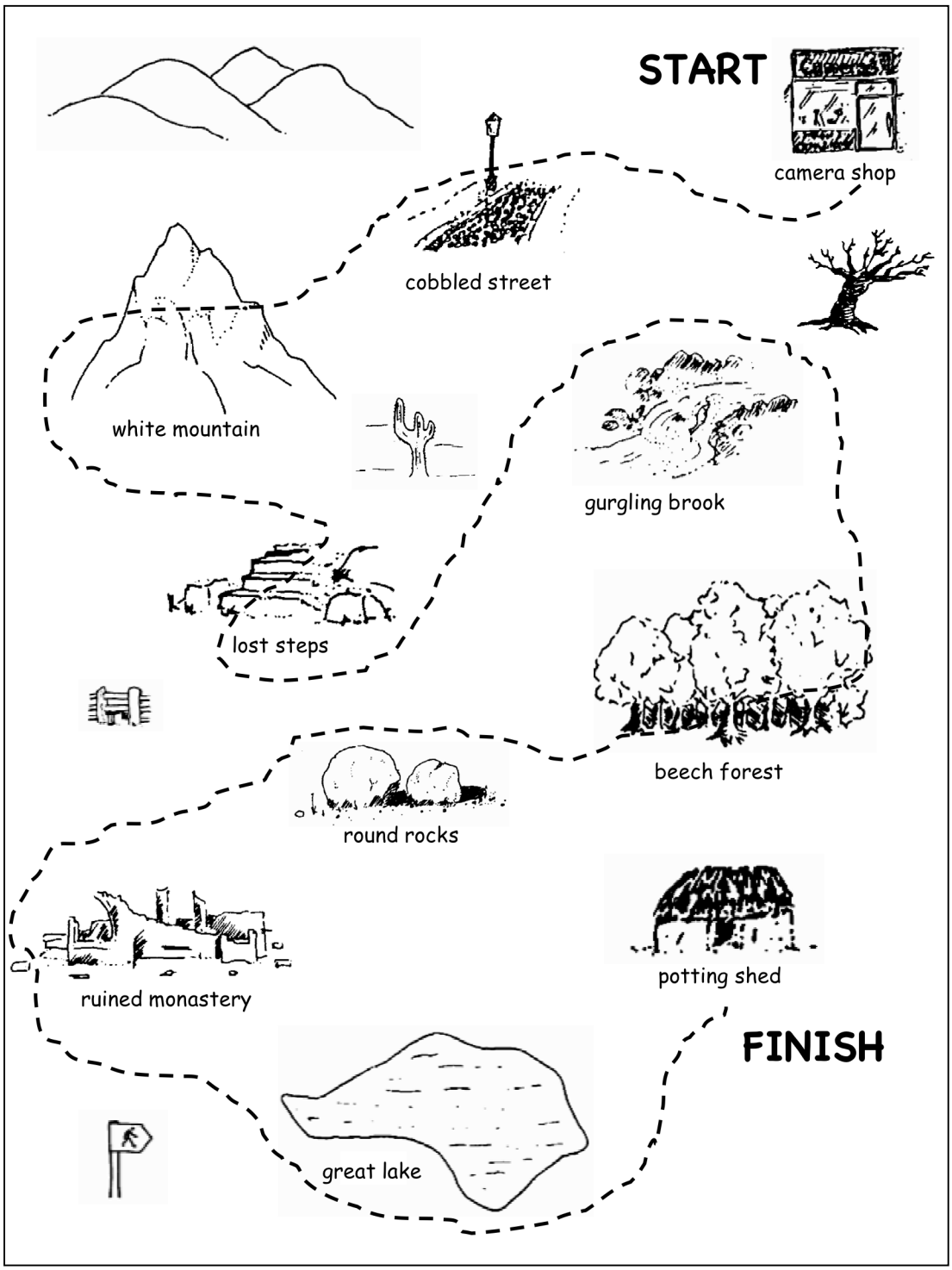
The HCRC Map Task Corpus. http://www.hcrc.ed.ac.uk/maptask/#top

Van Berkum, J. J. A. (in press). The electrophysiology of discourse and conversation. In M. Spivey, M. Joanisse, & K. McRae (Eds), *The Cambridge Handbook of Psycholinguistics.* Cambridge: Cambridge University Press.

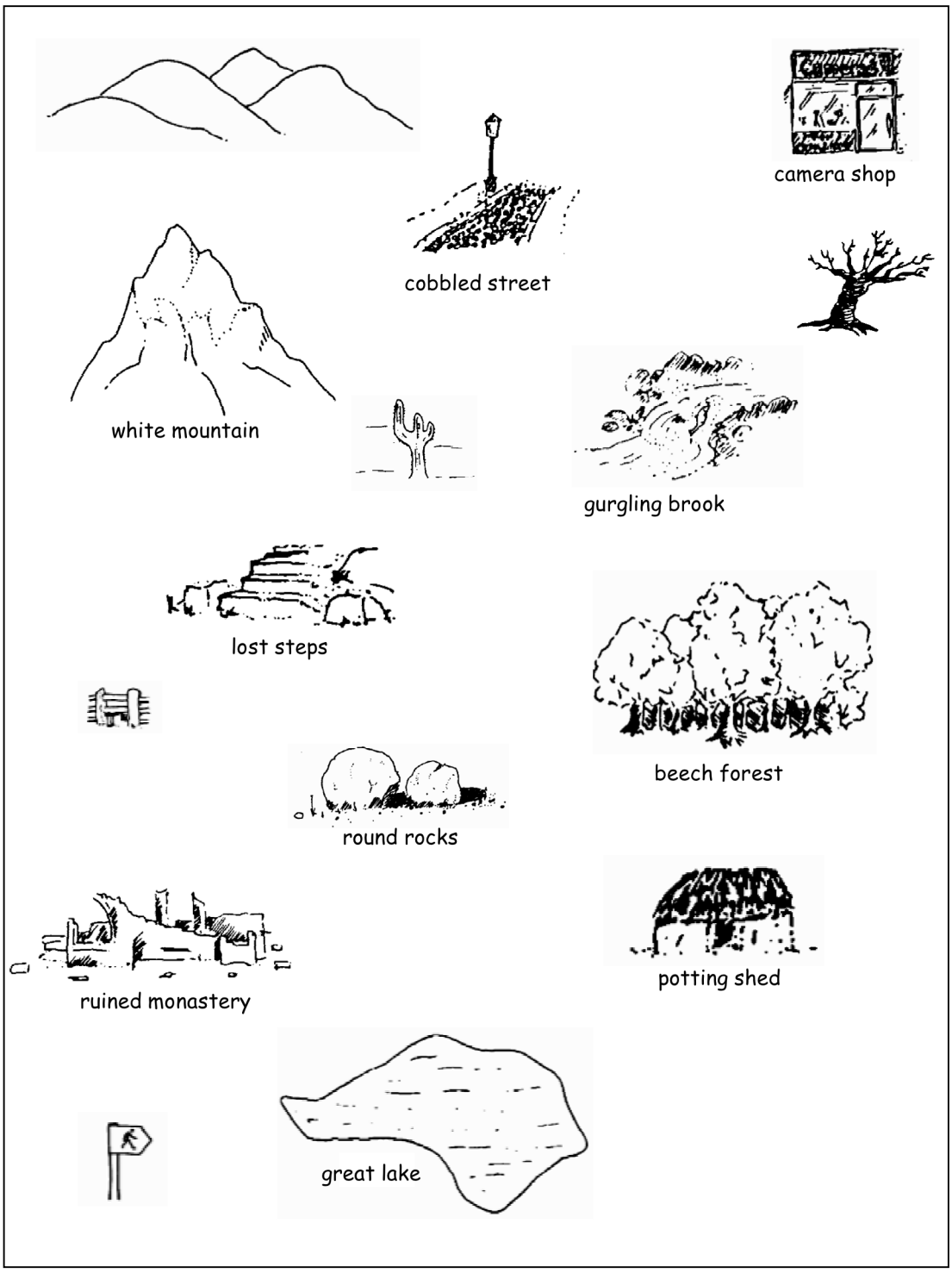
Yngve, V. H. (1970). On getting a word in edgewise. *Papers form the Sixth Regional Meeting of the Chicago Linguistic Society* (pp. 567- 578). Chicago: Chicago Linguistic Society.

Appendix 1
Map Task

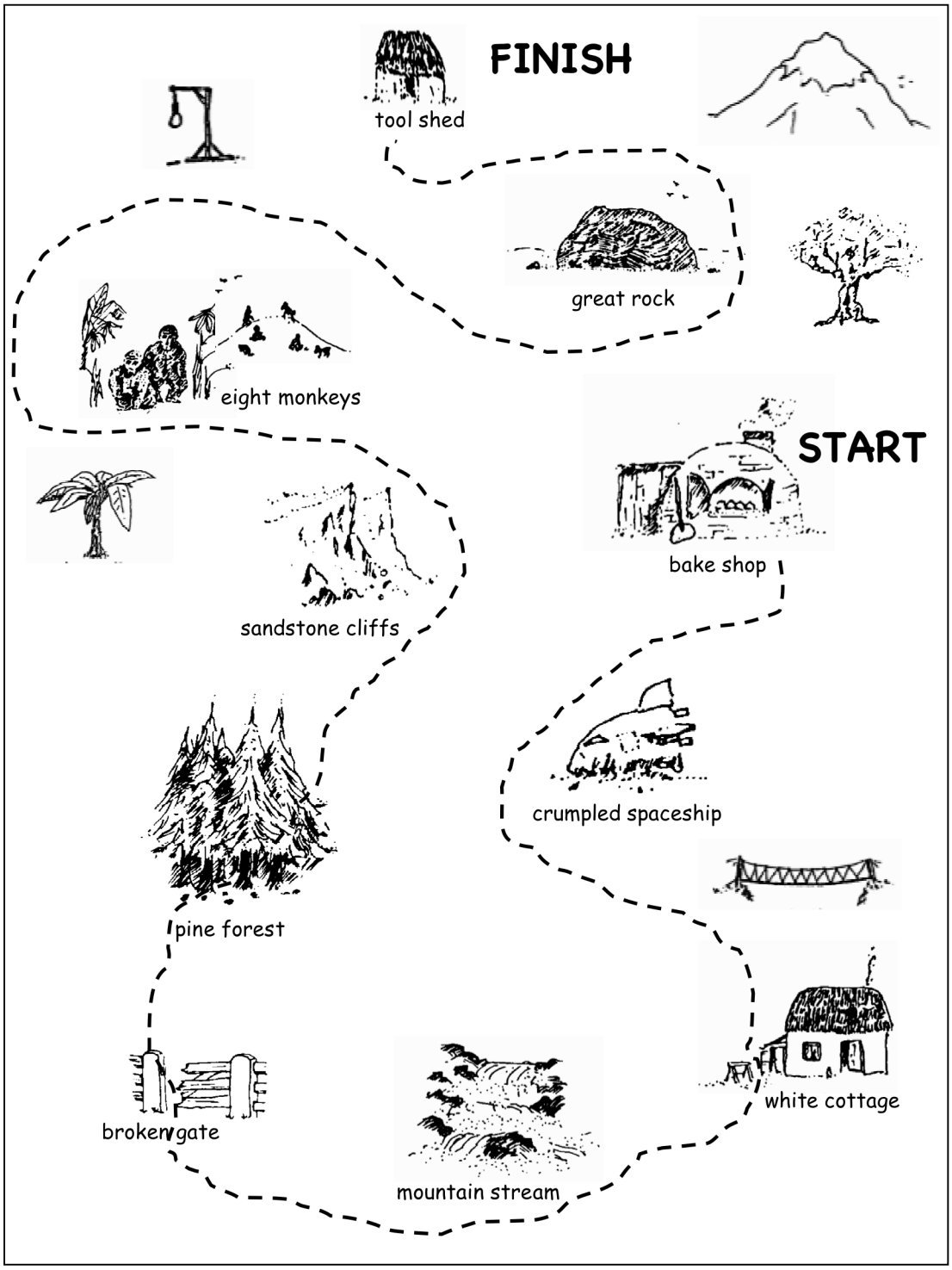
Target Map 1. Directors and matchers with low informational needs



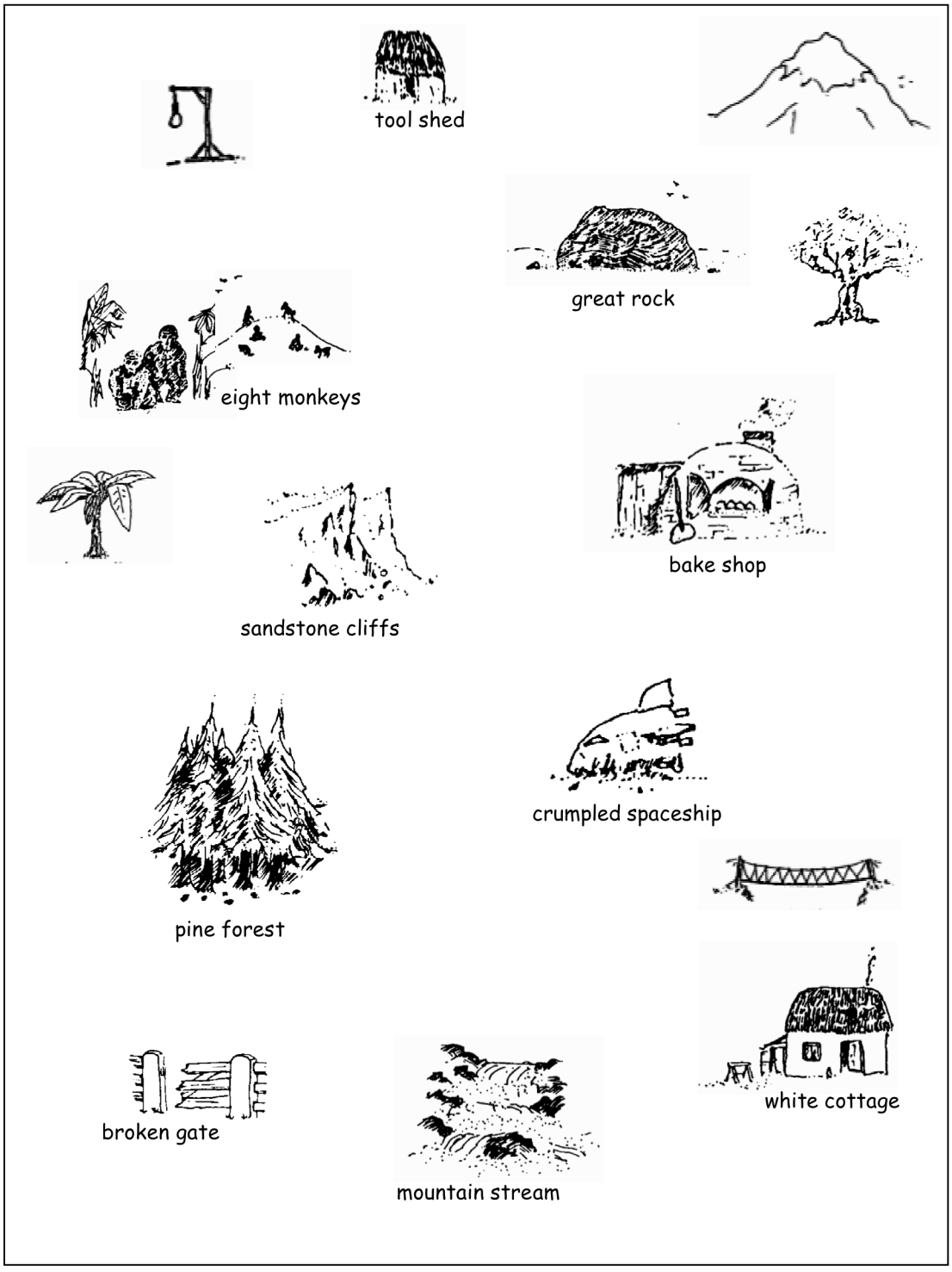
Target Map 1. Matchers with high informational needs



Target Map 2. Director & matcher with low informational needs



Target Map 2. Matcher with high informational needs



Appendix 2
Narrative Task

Typical and atypical stories in order of narration (target action and target instrument underlined).

Typical version	Atypical version
Sally turned off the light and lay down to sleep. Just as she was drifting off she heard a squeaky sound under her bed, which woke her up. She thought, "Tomorrow, I need to get a mouse trap to put underneath the bed."	Sally turned off the light and lay down to sleep. Just as she was drifting off she heard a squeaky sound under her cot, which woke her up. She thought, "Tomorrow, I need to get a mouse trap to put underneath the cot."
Henrietta put some cream in her coffee and stirred it. She put her spoon down and took a big gulp of the coffee. It was so hot that she dropped the cup, knocking the spoon off the table.	Henrietta put some cream in her coffee and stirred it. She put her swizzle stick down and took a big gulp of the coffee. It was so hot that she dropped the cup, knocking the swizzle stick off the table.
Tim was afraid that his wife would find out that he still smoked. He would sit in his favorite chair and enjoy the taste of his favorite brand of cigarettes as he read. However, he always left the window open so the smoke from his cigarettes wouldn't alert his wife.	Tim was afraid that his wife would find out that he still smoked. He would sit in his favorite chair enjoying the taste of his favorite pipe as he read. However, he always left the window open so the smoke from his pipe wouldn't alert his wife.
Adolph hid behind the door and when the man entered the kitchen he stabbed him in the back. He wiped the blood off the knife and rummaged through the house. Later police investigators found his fingerprints all over the knife and had no trouble catching him.	Adolph hid behind the door and when the man entered the kitchen he stabbed him in the back. He wiped the blood off the ice pick and rummaged through the house. Later police investigators found his fingerprints all over the ice pick and had no trouble catching him.
Jane went to the kitchen and poured herself some hot chocolate. She carried the cup into the living room and sat down to collect her thoughts. Her dog ran into the room wanting to play, and knocked the cup over.	Jane went to the kitchen and poured herself some hot chocolate. She carried the glass into the living room and sat down to collect her thoughts. Her dog ran into the room wanting to play, and knocked the glass over.

Michael took careful aim at the rabbit and shot at it. However, his gun misfired, startling the rabbit. He thought, "I've got to get a new gun; there goes my dinner running off into the woods."	Michael took careful aim at the rabbit and shot at it. However, his crossbow misfired, startling the rabbit. He thought, "I've got to get a new crossbow; there goes my dinner running off into the woods."
Shirley was hungry, so she got out an English muffin and heated it. When she removed the muffin from the toaster it was nice and warm and she spread it with jam. In the pleasure of eating it, she didn't notice the smoke in the toaster.	Shirley was hungry, so she got out an English muffin and heated it. When she removed the muffin from the oven it was nice and warm and she spread it with jam. In the pleasure of eating it, she didn't notice the smoke in the oven.
Elizabeth started to eat her rice. When she heard the sound of her door being forced open, she became so scared that she dropped her fork and spat out the rice. "If it's a dangerous intruder," she thought, "my only weapon is my fork."	Elizabeth started to eat her rice. When she heard the sound of her door being forced open, she became so scared that she dropped her chopsticks and spat out the rice. "If it's a dangerous intruder," she thought, "my only weapon is my chopsticks."
After Lisa washed a load of clothing, she decided to dry them. Later when they were dry, she took them out of the dryer and started to fold them. She thought, "Boy, these are wrinkled. I shouldn't have left them in the dryer for so long."	After Lisa washed a load of clothing, she decided to dry them. Later when they were dry, she took them off the line and started to fold them. She thought, "Boy, these are wrinkled. I shouldn't have left them on the line for so long."
Henry peered into the dark cave. He entered it carefully, holding his flashlight high. At that moment there was a loud growling sound and suddenly a strong gust of wind blew through the cave and blew the flashlight out of his hand.	Henry peered into the dark cave. He entered carefully, holding his lantern high. At that moment there was a loud growling sound and suddenly a strong gust of wind blew through the cave and blew the lantern out of his hand.
When Mr. White was getting ready to go to work, he noticed the temperature was dropping outside, so he wanted to wear something on his hands to keep them warm. He went to the dresser by the door and looked for his gloves. When he put them on, he noticed one glove had a hole.	When Mr. White was getting ready to go to work, he noticed the temperature was dropping outside, so he wanted to wear something on his hands to keep them warm. He went to the dresser by the door and looked for his mittens. When he put them on, he noticed one mitten had a hole.

Patricia was writing her term paper until her roommate interrupted her. She turned off her computer when Jill suggested that they go to the movies. After they came back Patricia found that her computer didn't work anymore.	Patricia was writing her term paper until her roommate interrupted her. She turned off her typewriter when Jill suggested that they go to the movies. After they came back Patricia found that her typewriter didn't work anymore.