
Volume 62, Issue 1, January 2010
ISSN 0749-596X
ELSEVIER
Journal of Memory and Language
Editor
Victor Ferreira
Associate Editors
Sally Andrews
Simon Farrell
Robert Greene
James Magnuson
Editorial Board
Available online at www.sciencedirect.com
ScienceDirect

Journal of Memory and Language 62 (2010) 35–51

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

Attenuating information in spoken communication: For the speaker, or for the addressee?

Alexia Galati *, Susan E. Brennan

Department of Psychology, Stony Brook University, Stony Brook, NY 11794-2500, United States

ARTICLE INFO

Article history:
Received 27 May 2009
Revision received 1 September 2009
Available online 13 October 2009

Keywords:
Audience design
Attenuation
Common ground
Spontaneous speech
Addressees
Partner-specific adaptation
Communication

ABSTRACT

Speakers tend to attenuate information that is predictable or repeated. To what extent is this done automatically and egocentrically, because it is easiest for speakers themselves, and to what extent is it driven by the informational needs of addressees? In 20 triads of naive subjects, speakers told the same Road Runner cartoon story twice to one addressee and once to another addressee, counterbalanced for order (Addressee1/Addressee1/Addressee2 or Addressee1/Addressee2/Addressee1). Stories retold to the same (old) addressees were attenuated compared to those retold to new addressees; this was true for events mentioned, number of words, and amount of detail. Moreover, lexically identical expressions by the same speaker were more intelligible to another group of listeners when the expressions had been addressed to new addressees than when they had been addressed to old addressees. We conclude that speakers' attenuating of information in spontaneous discourse is driven at least in part by addressees. Such *audience design* is computationally feasible when it can be guided by a "one-bit" model (*my audience has heard this before*, or not).

Introduction

As a conversation unfolds, speakers adapt what they say. This includes attenuating words, expressions, and information repeated over time. For example, when speakers introduce a referent, they tend to use a full noun phrase, and when they mention the same referent again, they often use a shortened form such as a pronoun (Ariel, 1990; Chafe, 1976; Grosz, Joshi, & Weinstein, 1995; Gundel, Hedberg, & Zacharski, 1993). Articulation is attenuated as well; a word representing new information is pronounced relatively clearly, whereas its subsequent mention is often shorter, de-stressed, and less intelligible (Bard et al., 2000; Fowler & Housum, 1987; Samuel & Troicki, 1998). When no conventional lexicalized form is available, or if it is unclear how best to conceptualize a referent for current purposes, the speaker may introduce it with a lengthy and provisional description that becomes shorter later in the conversation, after it has been ratified or modified by the addressee during an interactive process of conversational grounding (Clark & Brennan, 1991; Clark & Schaefer, 1989; Clark & Wilkes-Gibbs, 1986). During grounding, hedge words (e.g., *sort of*; *reddish*) are dropped (Brennan & Clark, 1996; Horton & Gerrig, 2002), expressions tend to become more compact overall, and interlocutors adopt one another's terms. By converging on the same term, interlocutors can signal that they believe they are talking about the same entity and taking the same perspective (Brennan & Clark, 1996; Isaacs & Clark, 1987). Addressees appear to assume that speakers with whom they have established such precedents do not abandon them without reason (Matthews, Lieven, & Tomasello, 2008; Metzing & Brennan, 2003). Conversational partners converge on attenuated forms, evoking and using the meanings they have created together, such that successful communication is often marked by formal efficiency.

In this paper, we investigate what drives attenuation in spontaneous spoken discourse: to what extent does it emerge as a by-product of speaker-oriented information

* Corresponding author.
E-mail address: alexia.galati@gmail.com (A. Galati).

0749-596X/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.jml.2009.09.002

processing, and to what extent does it reflect an addressee's communicative needs? These questions are relevant to the debate about how and when social context is taken into account during the planning of utterances, and therefore have implications for the architecture of speech planning. One theoretical orientation focuses on utterance planning as the result of largely autonomous cognitive processes that can incorporate socio-contextual information only as an adjustment *after* initial egocentric processing, when necessary. Another theoretical orientation sees utterance planning as largely constraint-based, incorporating all available information (including social information such as what is known about an addressee's idiosyncratic communicative needs) in a highly interactive manner. These two extremes need not be mutually exclusive, as planning may differ depending on the kind of constituent being planned (for a summary of the debate, see Brennan & Hanna, 2009).

To the extent that partners in conversation share a dialog context, convergence on the same attenuated phrasal, lexical, and articulatory forms is consistent with the processing needs of both the speaker and addressee (Pickering & Garrod, 2004). That is, the second time around, speakers should find it easier and faster to tell an anecdote, plan a syntactic phrase, or articulate a referring expression than they did the first time, due to rehearsal and to priming. At the same time, addressees should require weaker cues during dialog to reinstate an entity or episode than to initially establish or retrieve it.

In fact, attenuation itself can provide informative cues for utterance interpretation. This is true at several grains of analysis. For instance, at the fine grain represented by articulatory cues, listeners rapidly interpret de-accented nouns as anaphoric or given (Dahan, Tanenhaus, & Chambers, 2002). At the grain of lexical accessibility, pronouns tend to cue discourse entities that are in the center of attention, and interpretation is actually slowed when a full noun phrase is used to refer to an entity already in the center of attention (Hudson, Tanenhaus, & Dell, 1986; Hudson-D'Zmura, 1988; this has been dubbed *the repeated name penalty* by Gordon, Grosz, & Gilliom, 1993). And at a larger grain, there is pressure to avoid telling people what they already know (as laid out in Grice's *cooperative principle* and *maxim of quantity*; Grice, 1975); in fact, addressees often presuppose information that is not stated explicitly in an utterance, but is needed to make sense of it (Stalnaker, 1974). In these ways, attenuating phrasal, lexical, and articulatory forms would appear to serve communication quite well and may be fundamental to interactive language use. For instance, on the level of repeated references, attenuation depends to a large extent on the ability of partners to interact, occurring substantially less when speakers address a silent listener, an imaginary listener, or a tape recorder (Schober, 1993).

What is not clear, however, is what actually drives attenuation of repeated information: does this represent simply what is easiest for speakers themselves (benefiting addressees fortuitously), or is it an adaptation by speakers to addressees' informational needs (and therefore a form of audience design)? The problem is that what looks like audience design may not actually *be* audience design. Investigations of audience design are confounded when both partners happen to know the same information, hold the same perspective, or experience the same dialog context (for discussion, see Keysar, 1997). It is possible that speakers plan their utterances "egocentrically" and fail to consider addressees' needs unless they have to, as an afterthought or repair (Brown & Dell, 1987; Dell & Brown, 1991; Horton & Keysar, 1996; Keysar, Barr, & Horton, 1998; Pickering & Garrod, 2004). To establish whether speakers attenuate forms in dialog automatically and egocentrically, or whether they can do this flexibly *for* their addressees, an experimental design must tease apart the distinct knowledge states or perspectives of partners in dialog.

In the work reported here, we have speakers retell stories in order to test two hypotheses that make different predictions about the length, completeness, level of detail, and clarity of articulation of repeated information. According to the *for-the-speaker* hypothesis, speakers attenuate repeated material if they have told the story before. This hypothesis predicts that a speaker will attenuate even if the material is new to an addressee, because for the speaker, the story's content, structure, and expressions are simply more accessible. According to the *for-the-addressee* (or audience design) hypothesis, speakers are more flexible and can tailor their stories to the informational needs of their addressees. We begin by surveying some previous tests of the for-the-speaker and for-the-addressee hypotheses, in order to motivate Experiment 1. Of course, these hypotheses need not be mutually exclusive, in that adaptations by a speaker could be driven by *both* egocentric processing and audience design. However, some accounts have simply assumed that any adaptation that aids addressees by definition represents audience design, and others have lacked a comparison that would enable teasing these for-the-speaker and for-the-addressee effects apart, concluding that adaptation is done egocentrically by speakers.

What drives attenuation in spoken dialog?

To test whether speakers attenuate repeated information egocentrically or else influenced by the informational needs of addressees, Bard and colleagues (2000; see also Bard & Aylett, 2000) sought to tease apart speakers' and addressees' perspectives by comparing the intelligibility and duration of a speaker's first production of a referring expression directed toward a first addressee, to that speaker's production of the same referring expression directed toward a second addressee. They reasoned that if speakers take addressees' knowledge into account, information new to an addressee would not be attenuated and would not decrease in intelligibility even when it had been previously mentioned by the speaker (to a different addressee). Their finding was that referring expressions directed to second addressees decreased in intelligibility (both in absolute terms and relative to citation forms produced by the same speaker) despite being new to those addressees, at least for dialogs for which speakers and addressees could not see one another. As for duration, speakers attenuated referring expressions slightly but not reliably when speaking to second addressees compared to first addressees. The conclusion from this study was that speakers are egocentric and do *not* use addressees' knowledge to guide articulation (Bard et al., 2000).

These analyses, however, were missing a key control: Bard et al.'s corpus did not include a condition in which the same speaker directed the same referring expression again to the same addressee (Speaker$_{old}$–Addressee$_{old}$, in which the information would be repeated for both interlocutors). If the duration and intelligibility of referring expressions are guided by *both* speaker-centric and addressee-centric factors, speakers may attenuate more when re-telling something to the same addressee (Speaker$_{old}$–Addressee$_{old}$) than to a new addressee (Speaker$_{old}$–Addressee$_{new}$). An unpublished study by Gregory, Healy, and Jurafsky (presented as a poster at AMLaP, 2001 and described in an unpublished ms in 2002) included this necessary condition in order to test whether the duration of referring expressions was shortened more for an old or knowledgeable addressee than for a new or naive addressee. In their study, speakers studied and then repeated three-sentence discourses containing target expressions either twice to the same addressee or once to a new addressee. For speakers' repeated discourses, they found significantly more shortening when information was known to addressees than when it was new to addressees.

Even though Gregory et al.'s study contained a control condition that Bard et al.'s did not, there were enough other differences between these two studies that make it difficult to judge the implications of their contrasting conclusions for audience design. Gregory, Jurafsky, and Healy's (2001) task, which involved studying and repeating short discourses, did not allow for spontaneous planning. The task used to generate Bard et al.'s (2000) corpus was spontaneous and interactive as well as much noisier, involving planning and production of referring expressions in a direction-giving task (HCRC Map Task Corpus, Anderson et al., 1991).

The speech planning system is characterized by incremental processing and a cascading architecture (e.g., Bock, 1995; Bock & Levelt, 1994; Dell, 1986; Dell & O'Seaghdha, 1992; Levelt, 1989). Utterances can be tailored in a variety of ways—not only in how they are articulated, but also in how they are planned and encoded (including variation in perspective, syntactic packaging, and lexical choices). This suggests that, with respect to potential for audience design, the system may be more nimble on some levels than on others. The audience design question has been addressed at several grains of analysis, including duration of repeated words (for which Bard et al. (2000) found no effect), intelligibility (which they concluded was egocentrically driven), and the definiteness of referring expressions (for which Bard and Aylett (2001), using part of the same corpus, found evidence of audience design). Together, this evidence was taken to support a "dual process model" in which articulation is encapsulated from partner-specific knowledge and defaults automatically to being egocentric, whereas other processes such as the planning of referring expressions can be guided by inferences about a partner's needs (Bard & Aylett, 2001).

However, others have accommodated such evidence as the result of probabilistic constraints on information processing rather than as the result of architectural constraints such as modularity (see, e.g., Brennan & Hanna, 2009; Hanna & Tanenhaus, 2004; Hanna, Tanenhaus, & Trueswell, 2003). On this sort of account, if audience design seems to emerge late in a particular utterance, that does not mean that early processing is necessarily egocentric, but perhaps information about the addressee's needs was not available early enough. To the extent that audience design depends on information about an addressee's needs, it cannot emerge unless that information is activated early enough to be used in utterance planning or interpretation (Horton & Gerrig, 2002; Kraljic & Brennan, 2005). Therefore we would expect to see evidence of audience design in situations where a relevant aspect of an addressee's needs or perspective (as distinct from a speaker's own) can be easily tracked—for example, a single either/or cue, learned or computed in advance or else supported by perceptual evidence.

In the studies we report here, we examined whether speakers' adjustments to addressees' knowledge would extend to a variety of production measures related to attenuation of repeated material, including articulation, amount of detail, completeness of the story, word counts, and the perspectives used across narrations. Our design was inspired by Gregory et al.'s (2001), in that we included situations in which material: (1) was being presented for the first time by speakers and was also new to addressees, (2) was repeated by speakers but new to addressees (Bard et al.'s comparison condition), and (3) was repeated by speakers to addressees who had heard the material from the same speaker before (Gregory et al.'s control). Unlike Gregory et al., we varied addressees' knowledge status during the re-tellings within-speakers instead of between-speakers, and our speakers performed a memory-intensive task that required planning utterances: watching a cartoon and then generating their own stories.

In Experiment 1, speakers told the same story to two addressees a total of three times, such that the status of information (old vs. new) varied across narrations (Speaker$_{new}$–Addressee$_{new}$, Speaker$_{old}$–Addressee$_{old}$, and Speaker$_{old}$–Addressee$_{new}$). Speakers narrated the story of a cartoon that had no dialog or voiceover and therefore produced spontaneously planned narrations. We examined whether speakers adjusted their stories to addressees' knowledge states via any of the following measures: completeness of the story in terms of number of narrative events and number of words, consistency of expression in terms of perspective and amount of detail, and duration of lexically identical expressions. In Experiment 2, we examined the intelligibility of a sample of lexically identical expressions excised from the re-tellings in Experiment 1. If speakers retell stories according to what is easiest for themselves (*for-the-speaker* hypothesis), each re-telling would have fewer narrative elements, fewer words, less detail, and shorter or less intelligible referring expressions. On the other hand, if speakers retell stories by taking their addressees' informational needs into account (*for-the-addressee* hypothesis), they would attenuate their stories and expressions more when re-telling to the same partner than when re-telling to a new partner.

Experiment 1

Design

Adjustments in speaking were assessed across the speaker's first narration to the original addressee (A1) and their subsequent narrations to the same addressee and to a new addressee (A2). The order of re-tellings of the story to the same and new partners was counterbalanced to control for any net effects of hypermnesia (increased recall about an event with repeated attempts; see Payne (1987) for a review), forgetting, or fatigue. Speakers narrated the cartoon story in either the order A1–A1–A2 ($\text{Speaker}_{\text{new}}$–$\text{Addressee}_{\text{new}}$, $\text{Speaker}_{\text{old}}$–$\text{Addressee}_{\text{old}}$, and $\text{Speaker}_{\text{old}}$–$\text{Addressee}_{\text{new}}$) or the order A1–A2–A1 ($\text{Speaker}_{\text{new}}$–$\text{Addressee}_{\text{new}}$, $\text{Speaker}_{\text{old}}$–$\text{Addressee}_{\text{new}}$, and $\text{Speaker}_{\text{old}}$–$\text{Addressee}_{\text{old}}$).

Participants

Sixty-nine students from Stony Brook University participated in triads. Speakers were all native English speakers; addressees were all fluent English speakers. Three triads were excluded without analysis and replaced due to idiosyncrasies of the speaker or addressee, or because of how they did the task.[1] Of the remaining 60 participants, 20 served in the role of the speaker and 40 in the role of the addressee. Forty-five of these participants were female and 15 were male. In none of the 20 triads did participants know each other in advance. Addressees were randomly assigned to the roles of A1 and A2. Participants were given credit toward a requirement in a psychology course.

Materials

A Looney Tunes cartoon without dialog (entitled "Beep Beep") starring Road Runner and Wile E. Coyote was used to elicit narratives.[2] The cartoon was edited for length such that it had four distinct episodes, corresponding to four attempts of Coyote to capture Road Runner. The length of the resulting cartoon was 3 min and 10 s.

Procedure

All participants were told that the study investigated storytelling and memory, and that the addressees, after hearing the stories of cartoons told by the speakers, would be tested for their memory of the stories at the end of the session. Speakers watched the cartoon alone, twice, on a computer. Then they moved to another room to narrate the story. They were informed in advance that they would be narrating each story twice to one of the addressees and once to the other, but they were not told to whom they would be narrating the story twice or in which order they would interact with each addressee. Construing the purpose of the experiment as a memory test for the addressees provided a convenient pragmatic rationale for why speakers would be telling the same story twice to one of the addressees; the speakers were asked to narrate the story in as much detail as possible each time. Addressees were told that they would hear stories of a cartoon once or twice and that they would have to remember as much of the narrations as possible for an upcoming memory test at the end of the storytelling session. They were also told that they could interact with the speaker (e.g., ask questions for clarification or make comments if they wished) during the narration.

Only one addressee was present in the room during each storytelling session, which was videotaped with a digital camcorder. The experimenter was present at the far end of the room during all the storytelling sessions to supervise the recording; speakers and addressees were instructed to face each other and to ignore the experimenter and the camera when doing the storytelling task. When the three storytelling sessions were completed, participants were told that there would not in fact be a memory test for the addressees, and they were all debriefed.

Narrative script

Each speaker's three narrations of the Road Runner cartoon were transcribed in detail by the first author. The transcripts were then filtered such that the following were excluded from word counts and narrative content measures: words interrupted and followed by speech repairs (which typically removed from only one to a few words), fillers such as "uh" or "um", the questions and contributions of the addressee, and metacomments or interactive exchanges that did not add propositional content that advanced the plot of the story. This precaution was intended to avoid any undue inflation of word counts, as the word counts were intended as an approximate measure of narrative content rather than as a measure of metacommentary or any production difficulties (see Arnold, Tanenhaus, Altmann, & Fagnano, 2004). Although measures of disfluencies, interaction, and metacommentary could exhibit sensitivity to the knowledge status of the addressee, they are beyond the scope of our current focus on the realization and articulation of propositional content.

To be able to compare the narrations within and across speakers, we created a script for narrative elements for the Road Runner cartoon by watching the cartoon and compiling a list of all observable events relevant to the plot that seemed likely to be mentioned by the speakers. Here, a narrative element referred to a proposition or set of propositions forming a sub-event that advanced the plot of the story. Then we used the narrations produced by the first five speakers to reduce the original list of 108 narrative elements to 85 by deleting those narrative elements that were not mentioned by any speakers and concatenating

[1] Two triads were excluded because the interaction between speaker and addressee were compromised; in one, one of the addressees had his eyes closed throughout the speaker's narration (trying to visualize the cartoon events, as he later reported); in the other, the speaker had a broken arm and was unable to gesture comfortably while speaking, which could affect storytelling (see Rauscher, Krauss, & Chen, 1996). A third triad was excluded because the speaker ran out of time and did not complete the task.

[2] Additional cartoons were also narrated, one with Tweety and Sylvester and one with Bugs Bunny and Yosemite Sam. The third cartoon was dropped from the task after the first six triads, because narrating a total of nine stories (three for each cartoon) was often too tiring for the speakers. The second cartoon was not analyzed.

those that were typically packaged together in a single description. Some narrative elements that were not in the original script but were consistently mentioned by the first five speakers were added to the script. These were typically inferences that arose from the observable events of the cartoon (e.g., *Coyote's intention is to drop the anvil on Road Runner*).[3]

Appendix A lists the narrative elements of the Road Runner cartoon, and Appendix B illustrates how the script was applied to coding, with an example of the last episode of the three narrations produced by Speaker 7, with narrative elements demarcated.

Content coding

The script was used to segment each narration's three transcripts and to match the segments with the corresponding elements of the script. Then each element could be compared for its presence and form across all three tellings of the narration by the same speaker. All comparisons were done blind as to order, as well as to which condition the addressee was in (old or new). Each pair-wise comparison was coded for: (1) mention of each element, (2) word counts for each element, and (3) similarity of expression of narrative elements in terms of: (i) the relative amount of detail for that element across narrations and (ii) the similarity of perspective for that element across narrations. The first author segmented the transcripts into narrative elements and coded for the mention of narrative elements, word counts, relative amount of detail, and similarity of perspective. Coding was done blind to experimental condition.

Narrative elements

The first narrative content measure, mention of an element, referred to whether a particular narrative element was realized across the three narrations. A narrative element could be present or absent in a given narration.

Word counts

The second measure, word counts, was used as a first approximation of the informativeness of each element across narrations. Although the word counts for a narrative element across the three narrations could generally indicate whether speakers attenuated or elaborated their expressions, they were not a precise indicator of such adjustments because words do not map directly to propositional content. In order to supplement the word count measure, two additional measures examined adjustments of expressions across narrations: the relative amount of detail and the similarity of perspective. By definition, if a speaker realized the same narrative element across narrations, there would be some degree of propositional similarity for those sets of utterances to have been categorized as the same element; however, the speaker's utterances could differ in terms of how much information was encoded and how the event was conceptualized.

Detail

The measure of amount of detail attempted to capture the relative amount of propositional content for a particular element across the three narrations that the word counts may not have reflected. In a particular re-telling, speakers could have encoded more, less, or the same amount of detail as the first telling of the story; these judgments were coded as +1, −1, and 0, respectively. Relative amount of detail was coded with respect to both the amount of propositional content reported and its specificity. Pronouns were not coded as being less detailed than noun phrases across narrations as long as the pronouns were anaphorically resolved by an equivalent noun phrase in a preceding narrative element. For each narrative element, this pair-wise comparison for amount of detail was made for Session 2's story compared to Session 1's, and for Session 3's story compared to Session 1's. A re-telling (either the second or third time a speaker told the story) was coded as having more or less detail when information within the narrative element had been added or omitted or when the lexical choices of the subsequent telling were more or less specific than those of the first telling.

Perspective

The measure of similarity of perspective attempted to capture the similarity of lexical and syntactic choices made by the speakers. In a re-telling, speakers could have used the same lexical and syntactic perspective as the first telling or they could have switched perspective; these judgments were coded as 0 and 1, respectively. Due to minor variability in the speakers' re-tellings, it was not required for the realizations of elements to be the same across narrations in order for expressions to be coded as having the same perspective. Synonymous expressions were coded as instantiating the same perspective as long as they were judged as conceptually equivalent. A difference in perspective could involve a difference in the assignment of thematic roles in the utterance (usually reflected by the syntactic structure or the choice of main verb); such a difference often reflected a switch in narrative voice or in how an action, object, or scene was conceptualized across narrations. Table 1 presents examples of perspective similarity coding judgments.

Reliability

To assess reliability for the content coding, we had a second coder, an undergraduate research assistant (blind to the order of the conditions and the identity of the addressee in the re-tellings as well as to the experimental hypotheses) redundantly code approximately 20% of the corpus for the relative amount of detail and the similarity of perspectives. This amounted to randomly choosing four speakers and making the two pairs (Sessions 1–2 and 1–3)

[3] An additional set of 11 "structural elements" (remarks that commented on episode boundaries, questions regarding the addressee's familiarity with the Road Runner and Coyote characters, or other metanarrative commentary regarding the cartoon genre) was also coded initially, but power was relatively low, so we do not include these in 'Results'. However, speakers were marginally likely to realize more structural elements in a re-telling to a new addressee than to an old addressee. In the transcript excerpts included in Appendix B, these structural elements are marked as "metacomments". Whenever the topic of familiarity came up in a metacomment, partners indicated that they were familiar with these cartoon characters.

Table 1
Experiment 1, examples of coding similarity of perspective for utterances encoding narrative elements across pairs of narrations.

Same perspective, despite not being verbatim	
1a.	And they give them like the Latin names underneath, it's like Acceleratis, you know, Maximus, and Carnivorus Vulgaris is the Coyote
1b.	And they give them the Latin names underneath? And he's got like Acceleratis Maximus and then the Coyote's got Carnivorus Vulgaris or something
Different perspective due to difference in main verb and thematic roles	
2a.	They're running after each other
2b.	Coyote is chasing after the Road Runner
Different perspective due to difference in thematic roles	
3a.	And then it slingshots him into the air
3b.	And then like he gets slingshot right up into the sky
Different perspective due to switch in narrative voice	
4a.	And he's like, alright, I'm gonna punch the Road Runner as he runs by
4b.	And he's gonna punch the Road Runner as he runs by
Different perspective due to difference in conceptualization of object	
5a.	And like around the rock you have like a harness of metal
5b.	Like with the metal apparatus all connected around it

of detail and perspective judgments, resulting in a total of 316 judgments comparing the relative detail and perspective of speaker's first mention of a narrative element and a subsequent mention. For reliability in detail coding, we calculated the proportion of cases for which the coders' judgments agreed about the relative amount of detail (less, same, more detail) when comparing a narrative element from the first telling to that in a re-telling. The coders agreed 83% of the time (Kappa = .71). For similarity of perspective in comparing a narrative element from the first telling to a re-telling (a binary decision between same vs. different perspective), coders agreed 93% of the time (Kappa = .85).

Duration measures

We also examined whether speakers would attenuate their articulation according to the knowledge status of their addressees. We focused on the introductory mentions of referring expressions produced by the speakers in each of the three re-tellings of the Road Runner cartoon. If the duration of key words depends only on whether the information is newly expressed or repeated for speakers, then we would expect tokens to become shorter over time (or at least for the second and third tokens to be shorter than the first). If, however, the duration of a speaker's referring expressions depends on the information status of the referent for their addressee, then we would expect tokens directed to knowledgeable addressees to be shorter than those directed to naive addressees. We included this measure because even though Bard et al. (2000) did not find any significant adjustments in duration across referential communication sessions, the (unpublished) study of read and retold short discourses by Gregory, Healy, and Jurafsky (2002) did.

Selection of tokens

Twenty salient objects (listed in Appendix C) were chosen because they were instrumental to the story and appeared with high probability and relative consistency across narrations by all or most of the speakers. These were either objects that Coyote used in his schemes to capture Road Runner (e.g., *boxing glove, anvil*), or else key objects in the environment in which the characters were situated (e.g., *cliffs, old cactus mine*). These objects were coded for whether speakers referred to them identically across the three re-tellings. Objects denoted by noun phrases were selected rather than verbs because speakers demonstrated more variability across a triplet of re-tellings in encoding actions (with respect to lexical choice, as well as tense and aspect) and this would have limited the sample of lexically identical tokens available for coding. The tokens coded for duration had to be realized in lexically identical form in all three narrations; each constituted the first mention of the referent in a narration, within the same syntactic expression (either the head of the noun phrase in all three narrations or the modifier of the noun in all three narrations; e.g., "the tightrope thing" could not be compared to "tightrope" because any differences in articulation could be syntactic in origin) and the same position in the utterance (utterance-initial, utterance-medial, and utterance-final, because speakers tend to elongate words at the end of an utterance; Cooper & Paccia-Cooper, 1980).

To level the playing field for comparing duration (or in Experiment 2, intelligibility), none of the tokens chosen could contain disfluencies or overlapping speech from an addressee. Furthermore, if the speaker paused immediately before or after a particular token, the pause had to occur in the same position relative to the comparable token in all three narrations. There was a concern that, to the extent that pausing before a word reveals difficulties with planning or retrieval, this criterion could have disqualified triplets that exhibited systematic differences in pausing across first vs. later re-tellings, which may in turn be associated with differences in word duration. This concern was dispelled because only 3.48% (*SD* = 18.35%) of any eligible tokens that appeared in triplets were directly preceded by a pause, and there were no systematic differences in pausing before first vs. later re-tellings. Triplets were *not* excluded if a silent or filled pause preceded a determiner associated with one of its tokens (e.g., "he picked up uh.. an anvil"). These criteria yielded a set of 89 triplets, 47 from speakers narrating in order A1–A1–A2 and 42 from order A1–A2–A1. The words were excerpted from digitally recorded materials and digitized at a sampling of 16 kHz (12 bit analog-to-digital conversion) using Goldwave sound processing software.

Segmentation heuristics

A linguist trained in phonetics measured the duration of these sets of lexically identical expressions using Praat software (Boersma & Weenink, 2006). The following heuristics were used: for both word/phrase-initial and word/phrase-final fricatives, nasals, and liquids if the frication, nasal murmur or liquid signal was too weak to be distinguished from the background noise, then the initial boundary was placed at the onset of the following vowel [fricative + vowel] word/phrase-initially, (i.e. at the start of visible periodicity in the waveform), or at the offset of the preceding vowel [vowel + fricative] word/phrase-finally (i.e. at the cessation of visible periodicity in the waveform). In the case of initial plosives, if the stop closure was saturated with background noise and not easily discernible in the waveform and spectrogram, then the boundary was placed at the release (burst) of the plosive if visible, or at the onset of the following vowel. For final plosives if the stop closure was not discernible or if there was a pause, the boundary was placed at the offset of the preceding vowel. These heuristics were applied in the same way to each token of a particular referring expression.

Analyses

Analyses of duration of expression, mention of element, word counts, and similarity of expression were conducted as 2 × 3 ANOVAs with knowledge status ($Speaker_{new}$–$Addressee_{new}$, $Speaker_{old}$–$Addressee_{old}$, or $Speaker_{old}$–$Addressee_{new}$) as the within-subjects factor and addressee order (A1–A1–A2 vs. A1–A2–A1) as the between-subjects factor. Except where indicated otherwise, two planned contrasts were examined: the first contrast looked for speaker-driven effects by comparing the speaker's first telling to A1 and subsequent re-telling to A2 ($Speaker_{new}$–$Addressee_{new}$ vs. $Speaker_{old}$–$Addressee_{new}$); in this comparison, information is held constant (new) for addressees. The second contrast looked for addressee-driven effects by comparing in triads of subjects the speaker's second telling to A1 and first telling to A2 ($Speaker_{old}$–$Addressee_{old}$ vs. $Speaker_{old}$–$Addressee_{new}$); in this comparison, information status is held constant for the speaker. For each result, we report three analyses: $F1$ or $t1$ is the analysis by-subjects (for which means are computed for triads of participants), $F2$ or $t2$ is the analysis by-items (for which means are computed for script elements or excised expressions, as relevant), and min F' is a more conservative test computed from $F1$ and $F2$ (see Clark, 1973). Confidence intervals are provided around the significant contrasts of interest.

Results

Mention of narrative elements

If speakers tailor their stories to what their addressees know, they should mention fewer narrative elements when re-telling the story to the same (old) addressee than to a new addressee. This prediction was supported; Fig. 1 illustrates the mean proportion of narrative elements mentioned

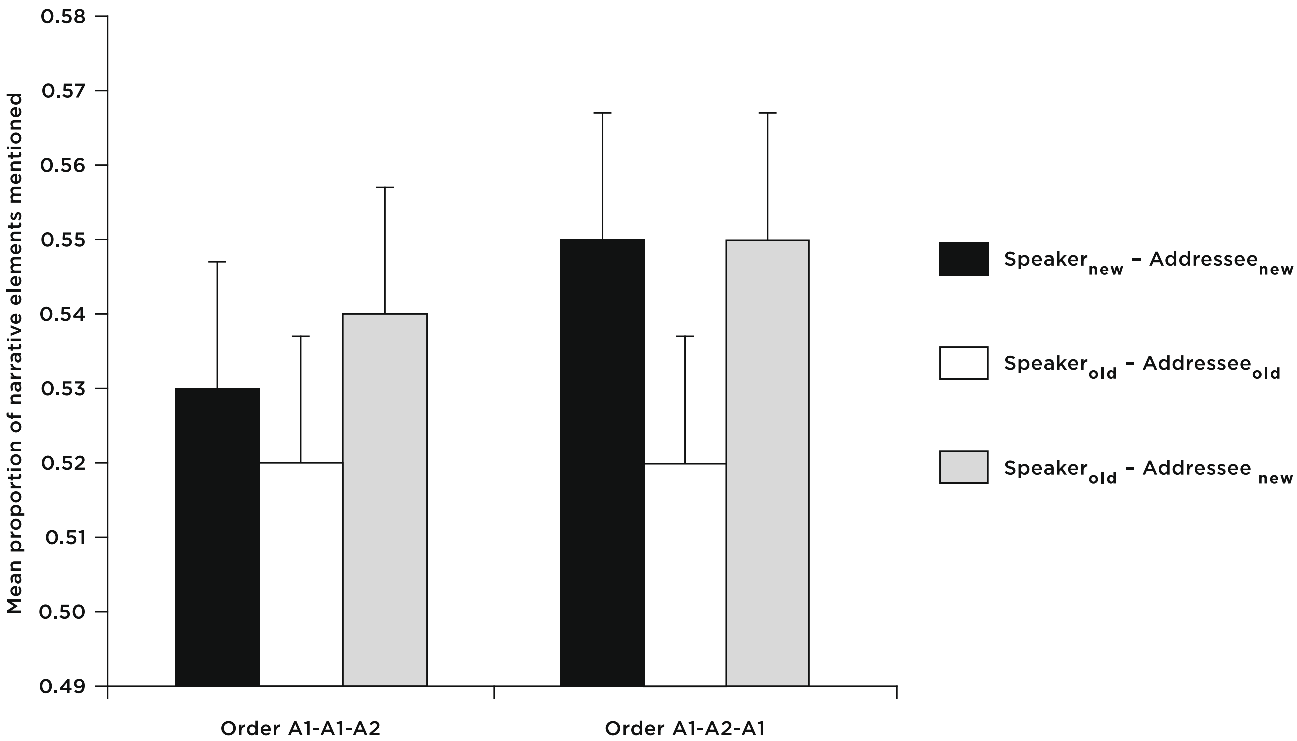

Fig. 1. Means for proportion of narrative elements mentioned according to old–new status in the two narration orders. Error bars represent standard error of the mean.

Table 2
Experiment 1, Partner-specific contrasts for percentage of narrative elements realized in stories, number of words per narrative element, and duration of lexically identical expressions. Significant effects are highlighted.

Contrasts	Narrative elements	Words per narrative element	Duration
For-the-speaker effect:			
$Speaker_{new}$–$Addressee_{new}$	$F_1(1, 18) = .58$, *n.s.*	$F_1(1, 18) = .11$, *n.s.*	$F_1(1, 18) = 2.05$, $p = .17$
vs.	$F_2(1, 84) = .77$, *n.s.*	$F_2(1, 74) = .00$, *n.s.*	$F_2(1, 16) = .40$, *n.s.*
$Speaker_{old}$–$Addressee_{new}$	min F' $(1, 49) = .33$, *n.s.*	min F' $(1, 74) = .00$, *n.s.*	min $F'(1, 22) = .33$, *n.s*
For-the-addressee effect:			
$Speaker_{old}$–$Addressee_{old}$	$F_1(1, 18) = 16.11$, $p < .01$	$F_1(1, 18) = 10.80$, $p < .01$	$F_1(1, 18) = .01$, *n.s.*
vs.	$F_2(1, 84) = 10.81$, $p < .01$	$F_2(1, 74) = 8.97$, $p < .01$	$F_2(1, 16) = 1.08$, *n.s.*
$Speaker_{old}$–$Addressee_{new}$	min F' $(1, 76) = 6.47$, $p < .05$	min F' $(1, 64) = 4.90$, $p < .05$	min F' $(1, 18) = .01$, *n.s.*

across the three narrations. Speakers mentioned fewer narrative elements in the re-telling to A1 ($Speaker_{old}$–$Addressee_{old}$) than in the re-telling to A2 ($Speaker_{old}$–$Addressee_{new}$; see Table 2); 95% CI for difference: 2.48% ± 1.28%. This was true regardless of the order in which they narrated the cartoon stories (A1–A1–A2 vs. A1–A2–A1), and there was no interaction between addressee knowledge and narration order. Perhaps surprisingly, in the test of the for-the-speaker hypothesis, speakers' knowledge status did not matter; the comparison with addressees' knowledge status held constant ($Speaker_{new}$–$Addressee_{new}$ vs. $Speaker_{old}$–$Addressee_{new}$) showed no difference in number of narrative elements mentioned; 95% CI for difference: .71% ± 1.91%. These results suggest that speakers adjust the number of narrative elements they mention according to what their addressees know rather than according to what they themselves know. There were no apparent effects of hypermnesia as revealed by the lack of a significant linear trend across the first, second and third narrations ($F_1(1, 18) = .55$, *n.s.*; $F_2(1, 84) = 1.13$, *n.s.*, min $F'(1, 38) = .37$, *n.s.*); in this task, speakers did not produce more narrative elements with repeated tellings of the same story.

Word counts of narrative elements

Speakers also used more words per narrative element when re-telling the story to a new partner ($Speaker_{old}$–$Addressee_{new}$) than when re-telling it to the same partner ($Speaker_{old}$–$Addressee_{old}$; see Fig. 2 and Table 2); 95% CI for difference: .87 ± .53. This was true regardless of the order in which they retold the story; there was no interaction between addressee knowledge and addressee order.

Even though each triad was randomly assigned to one of the two addressee orders, speakers in addressee order A1–A1–A2 happened to use more words to realize each narrative element on average than did speakers in addressee order A1–A2–A1 (the main effect of order was significant, $F_2(1, 74) = 7.48$, $p < .01$). The group of speakers who produced wordier narrative elements also tended to mention somewhat fewer narrative elements overall, as Fig. 1 illustrates, although the two groups did not differ significantly in this respect. As with the counts of narrative elements, the word counts across successive narratives showed no evidence of hypermnesia, the linear trend, $F_1(1, 18) = 1.03$, *n.s.*; $F_2(1, 84) = 1.46$, *n.s.*, min $F'(1, 47) = .60$, n.s. Once again, there was no speaker-driven effect ($Speaker_{new}$–$Addressee_{new}$ vs. $Speaker_{old}$–$Addressee_{new}$); 95% CI for difference: .11 ± .67. In this storytelling task, for word counts as well as for events realized, adaptation appeared to be driven by addressees' informational needs.

Similarity of expression: relative amount of detail

The amount of detail within narrative elements was compared in order to capture any elaborations and shortenings across narrations that were not revealed by differences in word counts. We tested the audience design hypothesis by examining whether speakers were more likely to reduce their level of detail when re-telling the story to the same partner than when re-telling it to a new partner. Indeed, compared to the baseline of their first telling of the story ($Speaker_{new}$–$Addressee_{new}$), speakers were 5.3% more likely to *reduce* the level of detail in a narrative event when re-telling it to the same partner ($Speaker_{old}$–$Addressee_{old}$), but 3.6% more likely to *increase* the level of detail in the event when re-telling it to a new partner ($Speaker_{old}$–$Addressee_{new}$). The difference in these adjustments in detail across re-tellings was reliable, $F_1(1, 18) = 7.96$, $p < .05$; $F_2(1, 73) = 3.82$, $p = .05$; min-$F'(1, 83) = 2.58$, $p = .11$; 95% CI for difference: .09 ± .07. Whether speakers narrated in the A1–A1–A2 or A1–A2–A1 order did not matter.

Similarity of expressions: perspective

This analysis captured whether speakers used the same or a different perspective to encode narrative elements across narrations. Although there is evidence that speakers maintain consistent partner-specific perspectives in referential communication (Brennan & Clark, 1996), tailoring perspectives for each addressee was not expected in this study, due in part to the relatively monologic nature of storytelling. In referential communication dialogs, speakers can adapt their conceptualizations of potentially ambiguous objects and the expressions with which they refer to them by using feedback from their new partners (Brennan & Clark, 1996). In our storytelling task there was apparently little ambiguity, and so the verbal feedback that addressees contributed was limited, with little if any need for speakers to revise or abandon a perspective once it had been presented. Moreover, the memory load imposed on speakers by telling a lengthy story was likely to be too demanding to keep track of distinctive perspectives used

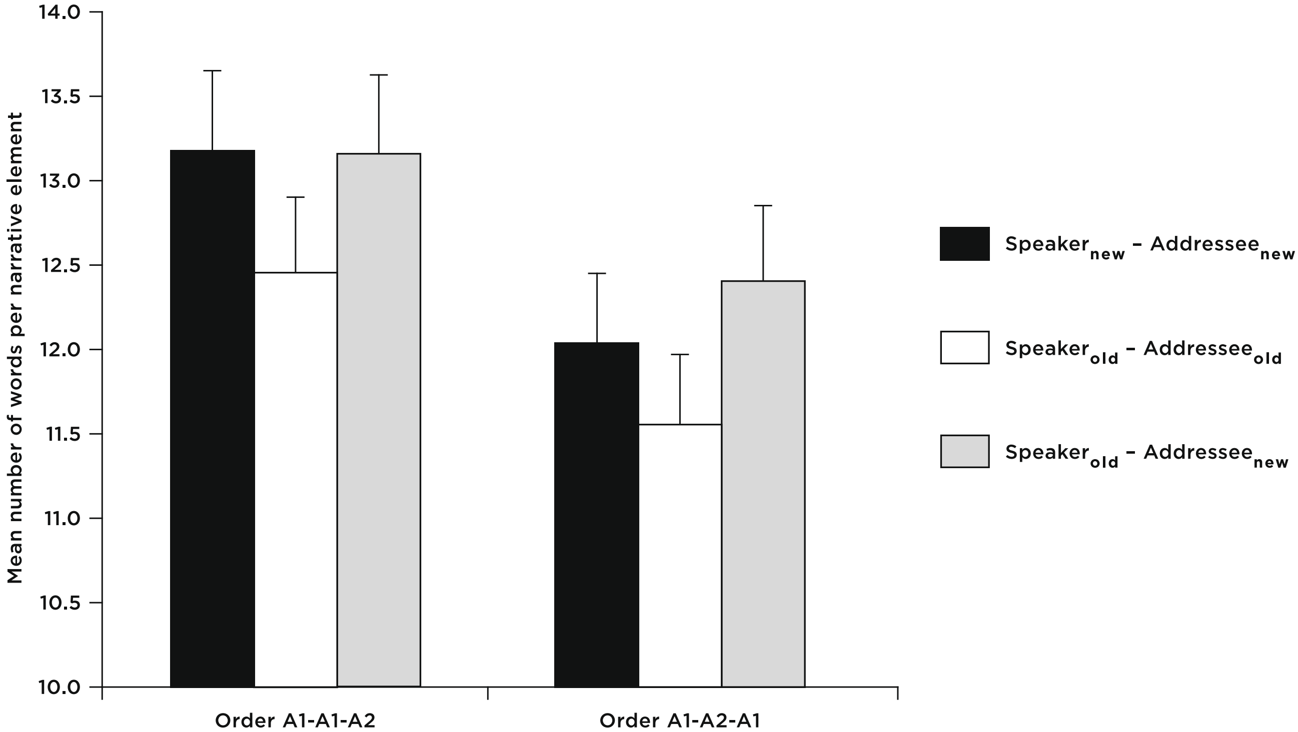

Fig. 2. Means for number of words per narrative element according to old–new status in the two narration orders. Error bars represent standard error of the mean.

for a particular element and addressed to a particular partner.

Out of the patterns of lexical perspectives that could emerge in speakers' re-tellings, we explored the occurrence of the pattern most relevant to testing the for-the-addressee hypothesis: we examined when speakers used the same perspective (regardless of the amount of detail) to one addressee, but a different perspective to the other. An example of this pattern comes from Speaker 16 who, in describing narrative element 26, used "the rock backfires" in her first telling to A1, then used "the rock ricochets back into him" in her telling to A2, and finally returned to "the rock backfires" in her re-telling to A1. In this pattern, the speaker's lexical and syntactic choices were identical in both narrations directed to A1 and differed from those choices directed to A2 ("backfires" was used as an intransitive verb and so did not take an argument, whereas "ricochets back" had the prepositional phrase "into him" as an argument). For all items with a perspective change, we compared the percentage of times lexical perspective appeared to depend on addressees' identity to the percentage expected due to chance (.50); for A1–A1–A2 order $t_1(9) = 1.91$, $p = .09$, $t_2(46) = 1.07$, *n.s.*, 95% CI for difference: .13 ± .15, and for the A1–A2–A1 order $t_1(9) = -1.10$ *n.s.*, $t_2(46) = -1.23$, *n.s.*, 95% CI for difference: .06 ± .13. When perspective changes were controlled for amount of detail, we found that speakers reused the same perspective marginally more often with the same partner than with a different partner, but only for narration order A1–A1–A2 (for this order, $t_1(9) = 3.08$, $p < .05$, $t_2(24) = 1.98$, $p = .059$, 95% CI for difference: .26 ± .19). Overall then, when re-tellings involved only one partner switch (and thereby less of a burden on memory for the perspective used with a particular partner), there was modest evidence that speakers may have tailored the lexical or syntactic perspectives within individual narrative events to addressees. It is also possible that perspective changes were not entirely addressee-driven, but that the speaker just decided that another perspective was preferable at that point (since the effect was present for the A1–A1–A2 order).

Duration

Speakers tend to shorten expressions that are more predictable (Clark & Haviland, 1977; Fowler & Housum, 1987; Lieberman, 1963; Samuel & Troicki, 1998). If articulation is flexible enough such that speakers can take into account whether an expression (or the information it encodes) is old or new to their addresses, they may shorten expressions repeated verbatim to the same addressee compared to those repeated to a new addressee. For the 89 identical triplets of expressions culled from speakers' three narrations, mean durations (normalized for number of syllables) were compared (see Table 3). It turned out that speakers did not adjust the duration of referential expressions according to who their partners were; there was no reliable effect of addressee knowledge ($Speaker_{old}$–$Addressee_{old}$ vs. $Speaker_{old}$–$Addressee_{new}$; 95% CI for difference: 1.34 ± 24.89), nor was there one of speaker ($Speaker_{new}$–$Addressee_{new}$ vs. $Speaker_{old}$–$Addressee_{new}$; 95% CI for difference: 11.91 ± 21.76, see Table 2), although the overall numerical means in Table 3 are shorten slightly with each re-telling. That is, speakers did not adjust the duration of expressions based on whether addressees were hearing

Table 3
Experiment 1, Means and *SD*s of duration normalized by number of syllables for the target 20 references.

	Speaker_{new}–Addressee_{new}	Speaker_{old}–Addressee_{old}	Speaker_{old}–Addressee_{new}
A1–A1–A2 order			
M	326.36	321.67	318.09
SD	137.45	141.01	145.32
A1–A2–A1 order			
M	386.46	374.13	374.95
SD	158.70	143.10	156.46
Total			
M	354.73	346.43	344.93
SD	150.07	143.63	152.51

them for the second time; nor did they make for-the-speaker adjustments either, as there was no shortening from the first (Speaker_{new}) to subsequent articulation of expressions (Speaker_{old}). This was true regardless of the order in which they narrated the stories.

Discussion

This pattern of findings supports the conclusion that audience design occurs for several aspects of planning the content of utterances, even in the relatively monologic task of storytelling. Speakers were less informative when re-telling stories to knowledgeable addressees than to naive addressees. First, they mentioned fewer events when a story was known to an addressee, while providing more comprehensive content when it was new to an addressee. Second, they used fewer words per event when re-telling the events of the story to the same addressee than to a new addressee. This converged with a third measure, which captured the relative amount of propositional content encoded across narrations; speakers were more likely to include less detail in referring expressions when describing an event to the same addressee than to a new one. That they were slightly more likely (3.6% of the time) to give more detail in a re-telling to a new addressee (A2) than they had in the first telling to the original addressee (A1) was not predicted; perhaps switching to a new partner served to remind them of the partner's need for information.

No evidence was found for speaker-driven attenuation upon repeated narration for any of these measures in the comparisons that controlled for addressee's information status (Speaker_{new}–Addressee_{new} vs. Speaker_{old}–Addressee_{new}). In this relatively spontaneous and memory-intensive storytelling task, attenuation appeared not to be egocentric, but driven by audience design.

The fact that speakers did not always reuse verbatim the same lexical perspectives in an addressee-specific manner is not too surprising, given the memory load that would be involved in keeping track not only of dozens of lexically or syntactically distinct expressions, but also to which partner they were addressed. That speakers were somewhat more likely to reuse the same perspective with the same partner than with a different partner in order A1–A1–A2, but not in order A1–A2–A1, is consistent with a memory load explanation. When re-tellings to the same partner are adjacent, the lexical and syntactic expressions used previously with that partner may be more available than after an intervening re-telling to a different partner (and in the latter situation, lexical perspective would have to be switched twice in order to show audience design). The storytelling task itself offered no real motivation for taking a different perspective with a different partner; such motivation might emerge only if a speaker needed to ground the identity of individual expressions for objects or actions that are ambiguous to a particular addressee (as in the partner-specific "conceptual pacts" observed by Brennan & Clark, 1996). It would be relatively costly to successfully monitor and maintain such specific lexical perspectives in an addressee-specific manner, since they are not organized by a simple characteristic of the addressee (e.g., whether the story is old or new to the addressee). Nevertheless, there was modest evidence for some audience design in lexical perspective.

In all, hypermnesia did not seem to occur for events, words, and details included in the narrations; speakers did not mention more events with each re-telling. Instead, the informativeness of speakers' stories was consistent with audience design; speakers attenuated narrative content as a function of the knowledge and needs of their addressees rather than as a simple function of repetition.

Concerning duration of expressions repeated across narrations, Experiment 1 did not find reliable evidence for adjustments either for-the-speaker or for-the-addressee (numerically, the pattern of means was consistent with the for-the-speaker hypothesis, with the first re-telling's expressions about 10 ms longer than those in subsequent re-tellings to a new addressee). This null finding is consistent with Bard et al. (2000), who did not find durational attenuation by speakers repeating the same expressions in referential communication with different addressees. However, a more robust measure of attenuation during repetition is intelligibility; Bard et al. (2000) found expressions excised from dialogs in which speakers repeated expressions to a new addressee to be reduced in intelligibility, even though other aspects of repeated expressions were not attenuated (Bard & Aylett, 2001). Therefore in Experiment 2, we investigated whether speakers' referring expressions would become less intelligible upon repetition, as well as whether referring expressions repeated to the same partner would be less intelligible than those repeated to a new partner.

Experiment 2

With clear support from Experiment 1 that audience design shapes how content is realized in narratives, we examined the influences of speaker and addressee knowledge on articulation by comparing the intelligibility of lexically and syntactically identical expressions. Experiment 2 investigated whether listeners hearing expressions out of context would rate expressions mentioned to old addressees as relatively less clear than those mentioned to new addressees; these listeners heard all three tokens for each triplet and rated them for clarity with respect to one another.[4]

Participants

Forty Stony Brook University students participated in exchange for credit toward a research requirement in a psychology course. All were native speakers of English and reported no hearing problems. None had participated in Experiment 1.

Materials

The items that listeners heard in Experiment 2 were selected from the corpus of lexically and syntactically identical expressions coded for duration in Experiment 1. Since the 89 triplets culled from narrations in Experiment 1 included expressions that were produced by more than one speaker, we identified how many of them were unique expressions. Of the 39 that were unique expressions, we excluded those that were acronyms (e.g. *TNT*) or a conjunction of noun phrases (e.g. *fork and knife*). If expressions shared morphemes, we selected the expression with the fewer morphemes (e.g., *blueprint* over *blueprints*), with the exception of *coal mine* being selected over *mine* (to avoid homophones). This yielded 27 unique expressions. Of the 27 triplets of items selected, three were excluded (*bib*, *cave*, *hats*) because two judges familiar with the cartoon stimulus could not identify any of the three tokens of each item. This resulted in a final set of 24 items.

Because some of these unique expressions were contributed redundantly by more than one speaker (for example, nine different speakers contributed triplets of *rock*), we determined which speaker's triplet to use by maximizing the number of speakers who contributed to the sample and to balance, to the extent possible, the number of triplets (both monosyllabic and multi-syllabic) from the two narration orders in Experiment 1. The final set of 24 items included 10 items from the order A1–A1–A2 and 14 from the order A1–A2–A1 (see Appendix D).

Words were excised to include their complete onsets and offsets; this was done by examining their waveforms and listening to the edited results. In order to reduce the variability of loudness and background noise across triplets, stimuli were processed through a digital to analog converter (12 bit; 16 kHz rate), filtered (using Adobe Audition's noise reduction filter), and amplified (in Goldwave). The noise reduction filter performed a Fourier analysis of an area of the recording chosen to be background noise – a pause or silence from the recording session from which the item was excised – and those constituent frequencies were removed from the original recording. The same background noise was selected for all tokens of a given triplet, since they came from the same recording session. Since there was some variability in recording quality across the experimental sessions, the volumes were adjusted to reduce this variability across items. Crucially, the same adjustments were applied to all three tokens within each item to avoid affecting their relative intelligibility; within each triplet; waveforms were adjusted equally such that none of the three tokens were clipped.

Because participants in this experiment played the audio files of the stimuli to hear them at their own pace, the audio files were coded with random bird names (e.g., *egret*, *flamingo*, *heron*) and organized in electronic folders labeled Trials 1–24.

Procedure

Listeners accessed the 24 expressions in a single randomized order; they heard all three versions of each expression. They were instructed to identify the expression represented by each triplet and rate, on an answer sheet, how clear the three versions were with respect to one another. They accessed the items on the desktop of a computer in 24 folders, each of which contained audio files with the three versions of a given expression. They were instructed to open all three audio files, play them as many times as necessary in order to identify the expression, and assess the relative clarity of the three words with respect to one another. They wrote the identified words on an answer sheet and rated them for clarity on a single scale from 1 to 5, where 1 was low 5 was high clarity. If two or even all three audio files sounded equally clear, they could be assigned the same rating.

Design and analyses

A response was coded as correct if it was within one phoneme from the original stimulus. Analyses were done on the relative clarity scores using a 3 (information status) × 2 (narration order) ANOVA. As in Experiment 1, there were two planned contrasts: to test for speaker-driven effects, we compared listeners' relative clarity ratings for expressions mentioned in Speaker_{new}–Addressee_{new} vs. Speaker_{old}–Addressee_{new} stories, and to test for addressee-driven effects, we compared relative clarity ratings for

[4] We conducted two additional intelligibility experiments in which participants heard unique tokens excised from only one of the three narrations and attempted to identify each token, while indicating their confidence in their response. The first experiment used a "words-in-noise" perceptual identification task to determine whether expressions could be identified in higher levels of noise when they came from narrations to new addressees than from narrations to old addressees. Critical words buried under a high level of white noise were played repeatedly, with each iteration including lower and lower amplitude noise; listeners were instructed to press a button at the earliest point that they could recognize the expression. However, data from 15 participants showed that accuracy was at floor, regardless of the narration the tokens had come from. The second experiment presented the tokens to listeners without noise and only once. This showed some evidence for listeners being better able to identify expressions that came from narrations to the new partners than from narrations to old partners, and for being more confident in their identifications, although this evidence did not generalize to items. The items were extremely variable, having been produced within long stretches of spontaneous speech with lengthy delay between tokens. Since in these experiment listeners heard only one version of each of the items, we sought to increase power by conducting Experiment 2, in which where naive listeners heard all three versions of each item and rated their relative clarity.

expressions mentioned in $Speaker_{old}$–$Addressee_{old}$ vs. $Speaker_{old}$–$Addressee_{new}$ stories.

Results

Overall, listeners identified expressions correctly 98% of the time ($SD = .17$), which is not surprising since they listened to all three versions of the same expression as often as they liked. The mean clarity rating for the first telling to A1 was 3.63 ($SD = 1.32$), for the re-telling to A1 it was 3.38 ($SD = 1.26$) and for the re-telling to A2 it was 3.67 ($SD = 1.24$). The same pattern held for both narration order A1–A1–A2 and for narration order A1–A2–A1. The mean relative ratings according to which narrations they were excised from are shown in Table 4.

As shown in Table 5, listeners rated expressions that came from re-tellings to the same addressee as less clear than those that came from re-tellings to a new addressee (95% CI for difference: .29 ± .07; reliably by-subjects and marginally so by-items).

The for-the-speaker comparison, on the other hand, was reliable only by-subjects, and it was not in the expected direction hypothesized by egocentricity: listeners rated expressions that had been directed to new addressees as clearer when they were old for the speaker than when they were new for the speaker (95% CI for difference: .04 ± .07). As the means in Table 4 suggest, there was also an interaction between addressee knowledge and narration order that was significant by-subjects only ($F_1(2, 78) = 15.37$, $p < .01$; $F_2(2, 44) = .79$, *n.s.*, min-$F'(2, 49) = .74$, *n.s.*). For the order A1–A2–A1, the trend was consistent with both speaker-driven and addressee-driven patterns (in which the ordered mean ratings were: 3.72, 3.58, 3.46); expressions from re-tellings to new addressees were slightly less intelligible than those from first tellings, but also slightly more intelligible than those from re-tellings to old addressees. However, for the order A1–A1–A2, when the partner switch happened only for the last narration, the expression retold to a new addressee was in fact rated as having about a 1/3 point *increase* in intelligibility over the first narration and over a 1/2 point increase over the re-telling to the old addressee, consistent with an addressee-driven pattern but not with a speaker-driven pattern.

Table 4
Experiment 2, means and *SD*s for ratings for relative clarity of lexically identical expressions excerpted from either narration order A1–A1–A2 or A1–A2–A1.

	$Speaker_{new}$–$Addressee_{new}$	$Speaker_{old}$–$Addressee_{old}$	$Speaker_{old}$–$Addressee_{new}$
A1–A1–A2 order			
M	3.50	3.27	3.81
SD	1.40	1.30	1.16
A1–A2–A1 order			
M	3.72	3.46	3.58
SD	1.25	1.23	1.29

Table 5
Experiment 2, partner-specific contrasts for ratings for relative clarity of lexically identical expressions. Significant and marginal effects are highlighted in a box or underlined.

Contrasts	Relative clarity
For-the-speaker effect:	
$Speaker_{new}$–$Addressee_{new}$ vs. $Speaker_{old}$–$Addressee_{new}$	$F_1(1, 39) = 5.49$, $p < .05$[a] $F_2(1, 22) = .12$, *n.s.* min $F'(1, 23) = .12$, *n.s*
For-the-addressee effect:	
$Speaker_{old}$–$Addressee_{old}$ vs. $Speaker_{old}$–$Addressee_{new}$	$F_1(1, 39) = 75.66$, $p < .001$ $F_2(1, 22) = 3.40$, $p = .079$ min $F'(1, 24) = 3.25$, $p = .08$

[a] *Note*: This by-subjects effect is in the opposite direction of that predicted by egocentricity.

Discussion

The intelligibility results in Experiment 2 provide evidence that the speakers in Experiment 1 did adjust their articulation of repeated vs. new information to the knowledge of their addressees, with expressions from re-tellings to the same partner being less intelligible than those from re-tellings to new partners. Using relative clarity ratings, we showed addressee-driven adjustments, with listeners rating expressions that came from re-tellings to the same addressee as less clear than those that came from tellings to new addressees.

However, concerning the for-the-speaker comparison, expressions were actually slightly (if not reliable) *less* intelligible overall when information was new to speakers than when it was old (with information status held constant at new for addressees). This pattern contrasts with the one found by Bard and colleagues; however, their corpus (Anderson et al., 1991) did not support a comparison in which the knowledge status of the speakers varied (new–old) while the status of addressees was held constant (new–new). For our narrative task, Experiment 2's evidence suggests that speakers tailored intelligibility according to audience design, especially when re-telling order involved only one partner switch such that the informational needs of the second addressee may have been particularly salient following two narrations to the same (first) addressee. The clear contrast in addressees' informational needs in one narration order (re-telling to a knowledgeable addressee and then to a naive one) may have led to the numerical boost in intelligibility over the expressions in the first narration.

General discussion

The studies reported here teased apart knowledge status for speakers from knowledge status for addressees, and the pattern of results confirms that audience design occurs at multiple grains in spontaneous utterance planning. When re-telling stories to knowledgeable addressees, speakers were less informative—mentioning fewer events, using fewer words, and including less detail—than when re-telling stories to naive addressees. And when identical tokens excised from the speaker's three narrations were played to different listeners later on, those tokens that had been addressed to knowledgeable addressees were rated as less intelligible, establishing that even a process as potentially automated as articulation can adjust to the needs of addressees.

We found no reliable evidence of egocentrism by speakers, even though speakers and addressees did not switch roles and the initiative in our storytelling task rested mainly with the speaker (it has been suggested by Haywood, Pickering, and Branigan (2005), that not switching roles or initiative may make findings of egocentrism more likely than findings of audience design). Our pattern of findings is inconsistent with proposals of modularity or dual processes (see, e.g., Bard & Aylett, 2001; Bard et al., 2000; Brown & Dell, 1987; Keysar, Barr, Balin, & Paek, 1998; Kronmuller & Barr, 2007), in which fast-acting processes (e.g., articulation) default automatically to being egocentric and encapsulated from partner-specific knowledge, leaving only more "inferential" (and resource consuming) processes to reflect partner-specific adjustments. In contrast, we found that even the articulation process is nimble enough to be influenced by addressees' needs when such needs are evident (and independent from the information's status for the speakers themselves).

The availability of cues or the salience of knowledge of the addressees' needs is, we propose, a key factor. It is particularly relevant that the situations in which audience design has been reliably documented to occur early or at a fine grain in utterance planning have involved settings in which addressees' needs are not only clear, but also simple—in fact, often captured by only two alternatives. All it takes for such situations to support audience design would be a single either/or cue that could make it extremely easy for speakers to track an audience's needs and keep them activated—a one-bit, most minimal partner model. For instance, a bilingual speaker in conversation with a monolingual limits herself (with apparent ease) to speaking the language her addressee understands (Bortfeld & Brennan, 1997); speakers who know that their actions can be monitored or that their partners have a picture of what they are discussing speak differently than those who do not (Brennan, 1990, 2005; Lockridge & Brennan, 2002); addressees who see that speakers are gazing at an object or are able to grasp it resolve ambiguous referring expressions more rapidly than addressees without this information (Hanna & Brennan, 2007; Hanna & Tanenhaus, 2004), and as we have shown here, interlocutors take into account (early in interpretation) whether they have communicated about a topic or referred to a particular object before with a particular partner (Brennan & Hanna, 2009; Matthews et al., 2008; Metzing & Brennan, 2003).

In our study, speakers' awareness of addressees' knowledge is not distinguished from mutual knowledge (as would be the case, for example, if some speakers had told the cartoon story to an addressee who had heard it first from someone else). Nonetheless, in ordinary conversation awareness of the addressee's knowledge can come from different sources that include mutual knowledge or common ground—such as linguistic co-presence, physical co-presence, or community co-membership (Clark & Marshall, 1981). We propose that audience design can take place to the extent that speakers represent relevant aspects of common ground in a simple, clear way; for example, whether a topic or referent is part of the discourse record shared with their addressee (linguistic co-presence, as in Metzing & Brennan, 2003 Matthews et al., 2008, and Nadig & Sedivy, 2002), part of their shared perceptual environment (physical co-presence, as in Brennan, 1990; Hanna & Brennan, 2007; Hanna & Tanenhaus, 2004; Lockridge & Brennan, 2002), or part of their shared sociocultural background (e.g., native language as in Bortfeld & Brennan, 1997, or expertise as in Isaacs & Clark, 1987). In this way, linguistic, contextual, and social forces can shape language planning. Another limitation is that although the spontaneous narration task used here demonstrated audience design in the planning of several kinds of linguistic constituents, this task did not lend itself to a direct examination of the time course of audience design.

It has been argued, and convincingly so, that for interlocutors to maintain and update elaborate models of one another (or of all the common ground they may share) is simply too computationally costly to support audience design automatically or early in utterance planning; inferences based on elaborate partner models should be as slow to activate and apply as any other kinds of inferences in memory (Horton & Gerrig, 2002, 2005a, 2005b; Polichak & Gerrig, 1998). But the fact remains that not all the adaptations that partners make to one another while producing and interpreting utterances unfold slowly or take the form of repairs (Hanna & Tanenhaus, 2004; Hanna et al., 2003; Metzing & Brennan, 2003; Nadig & Sedivy, 2002). On the contrary, when information about the partner's informational needs is available from an inference that has already been made, or when such information can be cued rapidly and unambiguously, speakers appear to represent and use such either/or information about the partner's needs at little or no discernable computational cost (for more discussion, see Brennan & Hanna, 2009). It is not yet clear just how simple (e.g., one-bit or more?) the partner-specific information need be in order for audience design to be computationally feasible in spontaneous spoken communication. In any event, a simple one-bit "model" of a partner such as the one implicated here (*I am telling this person this story for the first time vs. I have told this person this story before*) could well form the basis for generating a wide variety of adaptive behaviors that are not egocentric, but communicative and *for* the partner.

Acknowledgements

This material is based upon work supported by NSF under Grants IIS-0527585 and ITR-0325188. We would like to thank our colleagues from the Adaptive Spoken Dialogue Project, the Shared Gaze Project, the Dialogue Matters Network (funded by Leverhulme Trust), and the Gesture Focus Group for helpful discussion, especially Arthur Samuel, Marie Huffman, and Anna Kuhlen. We are grateful to Giorgos Tserdanelis, Elizabeth Cohen, Jenny Lee and Donna Kat for their assistance.

Appendix A

Experiment 1, narrative elements for the Road Runner cartoon ("Beep Beep").

1. Road Runner is being chased by Coyote/Road Runner and Coyote are running (in the desert)
2. Freeze shot: camera slows down on Road Runner, then on Coyote
3. Road Runner's Latin name, "Accelerati Incredibilus", is presented. Coyote's Latin name, "Carnivorus Vulgaris", is presented
4. Camera moves back to Road Runner
5. Road Runner resumes running, accelerating
6. Coyote continues running after Road Runner
7. Coyote is holding a fork and a knife and has a napkin around his neck
8. Coyote lunges forward with fork and knife, trying to stab Road Runner
9. Road Runner says "Beep Beep", takes off; Coyote misses Road Runner (Road Runner is too fast)
10. Coyote stabs the road; the utensils get stuck on the ground
11. Coyote's legs fly mid-air, still holding utensils, ending up inverted
12. Inference: Coyote tries to think of a plan
13. Coyote lifts hand to scratch head, collapses to the ground
14. Coyote props head on hands (Coyote looks frustrated)
15. Coyote gets an idea, his index finger is pointing up
16. Coyote is standing by the road next to a contraption (involving a boxing glove, which is on a spring, which is on a metal band, which is on a rock)/Coyote sets up a contraption
17. Coyote takes rock/metal band/spring/boxing glove
18. Coyote attaches boxing glove, (on spring), (on metal band), on rock
19. Coyote pushes glove (on spring) toward rock
20. Coyote clips spring on rock with a hook
21. Coyote hides behind rock
22. Coyote waits for Road Runner (holding string attached on hook)
23. Inference: Coyote's intention/the function of the contraption is to punch the Road Runner with the boxing glove
24. "Beep Beep" is heard/Road Runner comes, Coyote prepares to pull the string
25. Coyote pulls the string/lets the hook go
26. The boxing glove initially does not move.) The rock is propelled backward
27. Coyote is pushed backward by the rock and is smashed against the side of the cliff
28. Coyote panics at the sight of the boxing glove
29. The boxing glove is propelled backward and punches Coyote in the face
30. Coyote is holding/is looking at/has blueprints
31. Blueprint design/Setup of scheme: two precipices, a tightrope connecting them, a road passing between, Coyote standing on tightrope holding anvil
32. Blueprint design/Outlined steps: "Step 1: Carry anvil out onto tightwire; Step 2: Drop anvil on Road Runner; Step 3: Road Runner burger"
33. Inference: C's intention to drop the anvil on Road Runner while crossing tightrope

Appendix A (*continued*)

34. Coyote has/lifts up anvil (standing near the edge of the cliff)
35. Coyote steps on tightrope, holding the anvil
36. Inference: the anvil is heavy
37. Coyote slides down to the center of the tightrope
38. The tightrope stretches all the way to the ground, bringing Coyote all they way down to the ground
39. Coyote is standing on the ground in the middle of the road, still holding the anvil
40. Road Runner comes and stops in front of Coyote
41. Road Runner takes out his tongue and says "Beep Beep"
42. Road Runner runs away (in the direction from which he came)
43. Coyote gets angry
44. Coyote drops anvil on the ground (to go run after Road Runner)
45. Coyote is propelled up in the air by the tightrope
46. Coyote is ascending the sky
47. Coyote turns sideways as he's ascending, revealing his backpack/Coyote has backpack
48. Coyote looks unworried
49. Coyote points at his backpack (thinking/suggesting that it contains a parachute/the backpack is supposed to be a parachute)
50. Coyote begins descending
51. Coyote pulls the string on his backpack to release the parachute
52. Various items fall out of the bag and fly into the air: an axe, plates, a cup, utensils, a pan, a can
53. Coyote takes out a box of ACME aspirin
54. Coyote opens the box of aspirin
55. Aspirin pills fly into the air
56. Coyote grabs some pills and pops them into his mouth
57. Coyote looks panicked
58. Coyote waves goodbye to the camera
59. Coyote tumbles a few times in the air as he's falling
60. Coyote falls and hits the ground
61. Coyote's crash causes a cloud of dust to rise
62. Coyote is seen chasing after Road Runner. Road Runner runs into the "Old Cactus Mine"
63. Coyote moves into the mine until it becomes progres sively darker and only Coyote's eyes can be seen
64. Coyote runs back to the entrance of the mine (where there is a stand with the sign "Check caps here"), and puts on a mining helmet (with a light attached on it)
65. Road Runner appears and stops behind Coyote. Road Runner is also wearing a mining helmet
66. Road Runner takes out his tongue, says "Beep Beep" startling Coyote, and runs back into the mine
67. Coyote runs after Road Runner into the mine
68. The green light is Road Runner and the red light is Coyote
69. The green light is seen running across a tunnel and the red light is running behind it. The green and red lights are running through a grid of tunnels
70. The green light stops at an intersection on the grid
71. Close up view: Road Runner is seen entering an elevator. Road Runner takes out his tongue, says "Beep Beep", and goes down in the elevator before Coyote catches him
72. Coyote gets in another elevator and also goes down
73. Grid of tunnels view: the red light misses the green light

(*continued on next page*)

Appendix A (*continued*)

74. The Road Runner (green light) jumps across a gap. The red light (Coyote) also jumps over the gap and then Coyote (close up) is seen hanging from a cliff. Coyote gets back up and continues running
75. The green light (Road Runner) and the red light (Coyote) are seen moving around in circles. Road Runner and Coyote are seen going up and down ladders. Green and red lights continue going in circles. The green light (Road Runner) stops and hides at some point, while the red light (Coyote) continues moving in circles
76. The green light (Road Runner) and the red light (Coyote) are seen going through a zigzag path
77. The zigzag path splits into an upper path and a straight path. The green light (Road Runner) takes the upper path while the red light (Coyote) takes the straight path below/Road Runner loses the Coyote
78. The red light (Coyote) continues moving up and down in the straight path. Close up: Coyote is shown moving up and down, bumping his head on the ceiling
79. Coyote's helmet light goes off
80. It becomes dark (Coyote's eyes are seen blinking in the dark)
81. Coyote lights up a match
82. TNT boxes are seen behind the Coyote
83. The camera pans above to the top of the mine, where there are cactus trees
84. An explosion is heard
85. The cactuses fly up into the air and fall back down spelling "YIPE!

Appendix B

Episode 4 (old cactus mine) of Speaker 7 from Experiment 1 across the three narrations, with narrative elements demarcated.

According to the transcription conventions, silent pauses are marked (/) and breath pauses are marked (#). Filler are annotated as ⟨uh⟩ or ⟨um⟩. The end point of a self-interruption is marked with an asterisk (∗), and curly brackets { } contain undecipherable speech. Dysfluencies, interrupted utterances followed by repairs, fillers, the questions and contributions of the addressee, and interactive or metacommentary exchanges of the speaker with the addressee that did not add prepositional content that advanced the plot of the story are crossed out. Rising intonation is annotated with a question mark (?). Narrative elements are contained within square brackets [ ], and the number to which the element corresponds (1–85) is noted on the left. Metacomments marking the beginning and ending of episodes are also in square brackets.

Narration 1 to A1
62 [~~they~~* he followed the roadrunner into a cave #]
64 [a⟨a⟩nd/he had like the cap with the headlights on or whatever]
67 [and run in the cave]
75 [and they're just chasing each other around for a while]

Appendix B (*continued*)

77 [and then ~~he~~*/the roadrunner went a different way/in the cave than him]
82 [and he ended up going into like a/whole thing of/ ~~<uh>~~TNT?]
84 [a⟨a⟩nd he ended up/getting blown up I guess/]
Metacomment [and that was it/]

Narration 2 to A1
62 [and then he ended up/following the roadrunner into a cave]
64 [and he had the headlights on]
75 [and just chasing each other around for a while]
82 [a⟨a⟩nd/he ended up going to/a mound of TNT]
84 [and he ended up blown up/{I guess} underneath the ground]

Narration 3 to A2
~~and then he ended up chasing~~*
Metacomment [after that whole thing]
62 [he ended up chasing the roadrunner into/a cave/]
64 [and he had like the headlights on/]
65 [the green and red headlights/]
75 [and ~~<um>~~he just chased him around the cave for a while]
77 [and/roadrunner ended up/losing the coyote]
82 [a⟨a⟩nd/~~he ended up~~* the coyote ended up in like a/ mound full of like TNT/]
84 [and he got blown up underground]
Metacomment [a⟨a⟩nd that was it/]

Appendix C

Experiment 1, referring expressions selected for duration coding.

Episode 1
1. "Fork and knife"/utensils
2. Bib
3. Road/street/ground

Episode 2
4. Rock/boulder
5. String/rope
6. Boxing glove/punching glove/glove, etc.
7. Spring/spiral/coil, etc.

Episode 3
8. Backpack/book bag
9. Blueprints
10. Canyon/cliffs
11. Anvil/weight
12. Tightrope/tightwire/rope
13. Parachute
14. Aspirin(s)/pills

Episode 4
15. Mine/cave/"old cactus mine"
16. Hard hat(s)/helmet(s)/mining hat(s), etc.
17. Maze/square/tunnels, etc.
18. Match
19. TNT/explosives/dynamite
20. Cactus/cacti/cactuses

Appendix D

Experiments 2, referring expressions selected for intelligibility coding (with the narration order from which they were extracted in parentheses).

1. Handkerchief	(order A1–A1–A2)
2. Ground	(order A1–A2–A1)
3. Concrete	(order A1–A1–A2)
4. Floor	(order A1–A2–A1)
5. Road	(order A1–A1–A2)
6. Rock	(order A1–A2–A1)
7. String	(order A1–A2–A1)
8. Boxing glove	(order A1–A1–A2)
9. Spring	(order A1–A2–A1)
10. Backpack	(order A1–A2–A1)
11. Blueprint	(order A1–A2–A1)
12. Cliffs	(order A1–A2–A1)
13. Mountain	(order A1–A2–A1)
14. Anvil	(order A1–A1–A2)
15. Tightrope	(order A1–A2–A1)
16. Parachute	(order A1–A1–A2)
17. Aspirin	(order A1–A2–A1)
18. Pain pills	(order A1–A2–A1)
19. Coal mine	(order A1–A2–A1)
20. Maze	(order A1–A1–A2)
21. Cross-section	(order A1–A1–A2)
22. Dynamite	(order A1–A2–A1)
23. Match	(order A1–A1–A2)
24. Cactus	(order A1–A1–A2)

References

Anderson, A. H., Bader, M., Boyle, E., Bard, E. G., Doherty, G., Garrod, S., et al. (1991). The H.C.R.C. map task corpus. *Language and Speech, 34*, 351–366.

Ariel, M. (1990). *Accessing noun-phrase antecedents*. London: Routledge.

Arnold, J. E., Tanenhaus, M. K., Altmann, R., & Fagnano, M. (2004). The old and thee, uh, new. *Psychological Science, 15*, 578–581.

Bard, E. G., & Aylett, M. (2000). *Accessibility, duration, and modeling the listener in spoken dialogue. Gotalog 2000, fourth workshop on the semantics and pragmatics of dialogue*. Gotalog, Sweden: Gothenberg University.

Bard, E. G., & Aylett, M. P. (2001). Referential form, word duration, and modeling the listener in spoken dialogue. In *Paper presented at the twenty-third annual conference of the cognitive science society, Edinburgh, Scotland*.

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.

Bock, J. K. (1995). Speech production: From mind to mouth. In J. L. Miller & P. D. Eimas (Eds.), *Speech, language, and communication* (pp. 181–216). San Diego, CA: Academic Press.

Bock, J. K., & Levelt, W. J. M. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945–984). San Diego, CA: Academic Press.

Boersma, P., & Weenink, D. (2006). *Praat: Doing phonetics by computer (Version 5.0.11)*. <http://www.praat.org/> Retrieved 01.06.06 [computer program].

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.

Brennan, S. E. (1990). *Speaking and providing evidence for mutual understanding*. Unpublished doctoral dissertation, Stanford University, Stanford, CA.

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 [Special issue on joint action].

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 1482–1493.

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

Chafe, W. L. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C. N. Li (Ed.), *Subject and topic* (pp. 25–55). New York: Academic Press.

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior, 12*, 335–359.

Clark, H. H., & Brennan, S. E. (1991). Grounding in Communication. In L. B. Resnick, J. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 127-149).

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., & Marshall, C. R. (1981). Definite reference and mutual knowledge. In A. K. Joshi, B. Webber, & I. Sag (Eds.), *Elements of discourse understanding* (pp. 10–63). Cambridge: Cambridge University Press.

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.

Cooper, W. E., & Paccia-Cooper, J. (1980). *Syntax and speech*. Cambridge, MA: Harvard University Press.

Dahan, D., Tanenhaus, M. K., & Chambers, C. G. (2002). Accent and reference resolution in spoken language comprehension. *Journal of Memory and Language, 47*, 292–314.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review, 93*(3), 283–321.

Dell, G. S., & Brown, P. M. (1991). Mechanisms for listener-adaptation in language production: Limiting the role of a model of the listener. In D. Napoli & J. Kegl (Eds.), *Bridges between psychology and linguistics* (pp. 105–129). San Diego: Academic Press.

Dell, G. S., & O'Seaghdha, P. G. (1992). Stages of lexical access in language production. *Cognition, 42*, 287–314.

Fowler, C. A., & Housum, J. (1987). Talkers signaling 'new' and 'old' words in speech and listeners' perception and use of the distinction. *Journal of Memory and Language, 26*, 489–504.

Gordon, P. C., Grosz, B. J., & Gilliom, L. A. (1993). Pronouns, names, and the centering of attention in discourse. *Cognitive Science, 17*, 311–347.

Gregory, M. L., Jurafsky, D., & Healy, A. F. (2001). The role of the hearer in durational shortening. In *Poster presented at the seventh annual meeting of architectures and mechanisms for language processing (AMLaP), Saarbreucken, Germany*.

Gregory, M. L., Healy, A. F., & Jurafsky, D. (2002). *Common ground in production: Effects of mutual knowledge on word duration*. Unpublished manuscript.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.). *Syntax and semantics: Speech acts* (Vol. 3, pp. 41–58). New York: Academic Press.

Grosz, B. J., Joshi, A. K., & Weinstein, S. (1995). Centering: A framework for modeling the local discourse. *Computational Linguistics, 21*, 203–225.

Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions. *Language, 69*(2), 274–307.

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. (2004). Pragmatic effects on reference resolution in a collaborative task: Evidence from eye movements. *Cognitive Science, 28*, 105–115.

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

Haywood, S. L., Pickering, M. J., & Branigan, H. P. (2005). Do speakers avoid ambiguities during dialogue? *Psychological Science, 16*(5), 362–366.

Horton, W. S., & Gerrig, R. J. (2002). Speaker's experiences and audience design: Knowing when and knowing how to adjust utterances to addressees. *Journal of Memory and Language, 47*, 589–606.

Horton, W. S., & Gerrig, R. J. (2005a). Conversational common ground and memory processes in language production. *Discourse Processes, 40*, 1–35.

Horton, W. S., & Gerrig, R. J. (2005b). The impact of memory demands on audience design during language production. *Cognition, 96*, 127–142.

Horton, W. S., & Keysar, B. (1996). When do speakers take into account common ground? *Cognition, 59*, 91–117.

Hudson, S. B., Tanenhaus, M. K., & Dell, G. S. (1986). The effect of the discourse center on the local coherence of a discourse. In *Proceedings of the 8th annual conference of the cognitive science society* (pp. 96–101). Hillsdale, NJ: Lawrence Erlbaum Associates Inc..

Hudson-D'Zmura, S. (1988). *The structure of discourse and anaphor resolution: The discourse center and the roles of nouns and pronouns*. Unpublished Ph.D. dissertation, University of Rochester.

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

Keysar, B. (1997). Unconfounding common ground. *Discourse Processes, 24*, 253–270.

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., & Horton, W. S. (1998). The egocentric bias of language use: Insights from a processing approach. *Current Directions in Psychological Science, 7*, 46–50.

Kraljic, T., & Brennan, S. E. (2005). Prosodic disambiguation of syntactic structure: For the speaker or for the addressee? *Cognitive Psychology, 50*(2), 194–231.

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.

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

Lieberman, P. (1963). Some effects of context on the intelligibility of hearing and deaf children's speech. *Language and Speech, 24*, 255–264.

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

Matthews, D. E., Lieven, E. V. M., & Tomasello, M. (2008). What's in a manner of speaking? Children's sensitivity to partner-specific referential precedents. In *Proceedings of the LONGDIAL workshop on the semantics and pragmatics of dialog, London, UK*.

Metzing, C., & Brennan, S. E. (2003). When conceptual pacts are broken: Partner-specific effects in the comprehension of referring expressions. *Journal of Memory and Language, 49*, 201–213.

Nadig, A. S., & Sedivy, J. S. (2002). Evidence of perspective-taking constraints in children's on-line reference resolution. *Psychological Science, 13*, 329–336.

Payne, D. G. (1987). Hypermnesia and reminiscence in recall: A historical and empirical review. *Psychological Bulletin, 101*, 5–28.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences, 27*, 167–226.

Polichak, J. W., & Gerrig, R. J. (1998). Common ground and everyday language use: Comments on Horton and Keysar (1996). *Cognition, 66*, 183–189.

Rauscher, F. H., Krauss, R. M., & Chen, Y. (1996). Gesture, speech and lexical access: The role of lexical movements in speech production. *Psychological Science, 7*, 226–231.

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

Schober, M. F. (1993). Spatial perspective-taking in conversation. *Cognition, 47*, 1–24.

Stalnaker, R. (1974). Pragmatic presuppositions. In M. K. Munitz & P. K. Unger (Eds.), *Semantics and philosophy* (pp. 197–213). New York: University Press.