Speakers' experiences and audience design: knowing *when* and knowing *how* to adjust utterances to addressees[☆]

William S. Horton[*] and Richard J. Gerrig

Department of Psychology, State University of New York at Stony Brook, Stony Brook, NY 11794-2500, USA

Abstract

In this paper, we develop an account of the types of experiences through which speakers learn to design their utterances for particular addressees. We argue that there are two important aspects of conversational situations relevant to considerations of audience design. First, speakers must become aware that audience design is necessary in the current setting. Second, they must frequently overcome other tendencies toward consistency and brevity of expression. To assess the impact of both of these factors, we conducted a referential communication experiment in which Directors described arrays of picture cards for two independent Matchers. In the early rounds, both Matchers were present and each possessed a different subset of the Directors' cards. In later rounds, only one of the two Matchers was present at a time and worked with the entire set of cards. We evaluated the degree to which Directors' descriptions showed evidence of audience design by focusing on critical rounds when the Directors described cards that the current Matcher had not previously shared. Directors generally appeared sensitive to the distinction between shared and nonshared items. Additionally, there was more evidence of adjustment at the second partner change, suggesting that the Directors had learned something about the kinds of descriptions required in this situation. Our results suggest that it is important to consider the nature of speakers' experiences of interacting in a particular situation when making claims about the presence or absence of audience design. © 2002 Elsevier Science (USA). All rights reserved.

Keywords: Conversation; Audience design; Language production; Reference

[*] Corresponding author. Fax: +631-632-7876.
E-mail address: whorton@ms.cc.sunysb.edu (W.S. Horton).

In ordinary conversation, people often tailor what they say to suit a particular audience. Consider this excerpt from Louis Bayard's novel *Fool's Errand* (1999, p. 288):

> When Patrick got home that night, his father was on the couch, eating from a box of donut holes.
>
> "Oh, Pattie," he said, his mouth half full. "That fella called."
>
> "Which one?"

"That one I can never remember his name."
Seth wasn't home, though, when Patrick called back.

In this excerpt, Patrick's father, George, makes two attempts to refer to Seth. The first attempt is inadequate, whereas the second one establishes Seth's identity. On the evidence of the novel, no one else (in the fictional world) would be able to understand "That one I can never remember his name" as an unambiguous reference to Seth; Patrick can interpret his father's reference correctly because they share a history of George's memory failures.

This referring phrase exemplifies what researchers have called *audience design* (Clark & Carlson, 1981; Clark & Murphy, 1982), which refers to how speakers construct their utterances with the intention of being understood by particular recipients. As our example suggests, the fact that speakers often design utterances for specific addressees is not controversial. What remains open to question, however, is the range of circumstances in which speakers actually engage in audience design (e.g., Brown & Dell, 1987; Horton & Keysar, 1996; Schober & Brennan, in press). In this paper, we suggest that speakers' propensity to engage in audience design will change as a function of the experiences they have in particular contexts of communication. We argue that such experiences provide speakers with evidence concerning the necessity of audience design in specific contexts and also about the appropriateness of particular types of addressee-specific adjustments. To support our claims, we present data from a referential communication experiment that was designed to allow speakers to gain experience relevant to both of these considerations.

When speakers engage in successful audience design, they are satisfying Grice's (1975) Cooperative Principle. That is, they produce utterances under the assumption that those utterances will be relevant to the topic at hand, suitably informative, truthful, and clear. With respect to reference—i.e., circumstances in which one must, for example, decide between "That fella" and "That one I can never remember his name"— the aspect of Grice's cooperative principle that most often applies is the *maxim of quantity* (p. 45):

1. Make your contribution as informative as is required (for the current purposes of the exchange).

2. Do not make your contribution more informative than is required.

According to the cooperative principle and this maxim of quantity, speakers should consider their audience when deciding how much information is necessary to identify uniquely a particular referent. Consider, for example, how one might refer to one's younger sister. All else being equal, calling her "my younger sister" to another close relative would be saying too much, while calling her "Gwyneth" to a complete stranger would be saying too little. The imperative to be cooperative establishes for each speaker the important goal of designing utterances that are optimal for each addressee. Clark and his colleagues have identified this conversational goal as the *Principle of Optimal Design* (Clark, Schreuder, & Buttrick, 1983). Given the real-time exigencies of speech production, however, it is likely that speakers often fall short of optimality. Consider this excerpt from our experimental task, in which Directors provided descriptions of objects for Matchers who attempted to select the appropriate referent:

D: Next one is um- the house without the hat.
M: House without the hat?
D: Without the hat- there's nothing sticking out to the sides.
M: There's a little triangle-
D: It's like this-
M: Does it look like a fish?
D: It looks like a fish, yea.
M: Okay

Here, the Director fails to provide a description that is suitably informative for his current audience. Even when he attempts to provide additional useful information about the referent, this too fails to be helpful.

Our aim in this paper is to outline factors that affect the optimality of audience design. We suggest, in particular, that speakers' ability to engage in audience design depends on the experiences they accrue in particular circumstances of language use. We describe two considerations relevant to those experiences.

Speakers must realize that audience design is necessary

Suppose Ann wishes Betty to pass her a book that is sitting alone, in plain view, on a nearby table. As Ann formulates the utterance, "Please pass me the book," it seems unlikely that she must

engage in any special effort to design her utterance for Betty. We use this example to make the point that many of the tasks for which people use language may not require any overt effort toward audience design. There are, that is, a wide range of conventions that govern language use in a variety of settings (Clark, 1996). For that reason, it may not always be self-evident when a situation is sufficiently beyond the bounds of convention that speakers need to adjust their utterances to particular audiences. Instead, speakers may have to learn, through experience with particular situations of language use, when audience design is necessary.

In the example we gave earlier, we illustrated one type of conversational experience that might suggest to speakers that audience design is warranted. In that excerpt, the Director produced descriptions that (presumably) were adequate for his own identification, but which proved to be inadequate from the point of view of the Matcher. From this interchange, the Director had the opportunity to learn that the object under scrutiny permitted multiple perspectives (cf. Wilkes-Gibbs, 1995).

As a further illustration of the suggestion that speakers' experiences may call attention to the need for audience design, consider the objects shown in Fig. 1. Suppose that one wished to help someone identify one of the two fish in the figure. It would be relatively easy to facilitate identification of this object by referring to standard features of fish, like the number of fins or the intensity of its color. Suppose, however, that one wished to identify one of the other two objects in Fig. 1, which are abstract shapes called tangrams. Unlike the fish, no obvious frame of reference presents itself when trying to talk about these tangrams. As a result, speakers and addressees have to negotiate a particular perspective, or *conceptual pact* (Brennan & Clark, 1996), to arrive at an appropriate referring expression. Moreover, because particular speakers and addressees will often arrive at idiosyncratic perspectives for such ambiguous stimuli, the range of possible conceptualizations is much larger for tangrams than it is for objects like fish. For example, the leftmost tangram in Fig. 1 was referred to variously in our study as "the guy in the gown," "the figure with the box that's slightly tilted," or even, quite inexplicably, "the graduate student with the weird head."

One way of describing such variability in reference is to say that the tangrams do not share the same degree of linguistic *codability* as do familiar objects like fish, which are generally strongly lexicalized. For our purposes, we intend codability to mean the extent to which native speakers agree on what to call a given object (Lachman, Shaffer, & Hennrikus, 1974). Because of this relative lack of codability, speakers and addressees must work together to establish mutually agreed-upon perspectives for objects like these tangrams (cf. Hupet, Seron, & Chantraine, 1991). We suggest that
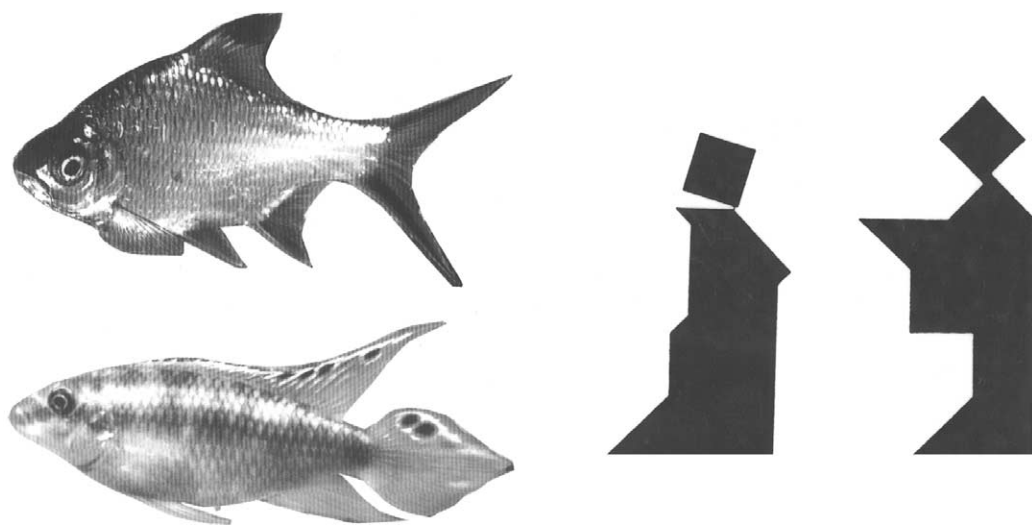

Fig. 1. Two pictures of fish and two tangrams used in the experiment.

there is a continuum in the degree to which particular conversational referents would induce speakers and addressees to work together to establish a mutually acceptable perspective; fish and tangrams represent two points on that continuum. In our view, differences in factors like codability often reveal themselves to language users through conversational interactions, which call attention to circumstances in which audience design is important.

These considerations lead to both a general and a more specific prediction. We suggest, generally, that speakers will discover the necessity of taking their addressees into account as they gain experience in particular circumstances of language use. The more specific prediction relies on differences in codability. We suggest that interactions with relatively less codable objects are more likely to lead to negotiations in which partner-specific perspectives emerge and that these negotiations will boost the relative probability of audience design with respect to such objects.

Speakers must overcome a tendency toward consistency of expression

Another factor influencing audience design processes concerns the strength of conceptual and/or lexical precedents that have been established in a particular speech situation. Because speech production is a process that occurs under time pressure, there is a general tendency toward economy. Speakers tend to produce the same types of forms over short periods of time (Bock, 1986; Bock & Loebell, 1990; Garrod & Anderson, 1987). This type of repetition of form or content may often work counter to the imperatives of audience design. Consider the experiment conducted by Brennan and Clark (1996), in which Directors were asked to describe a set of cards depicting common objects (e.g., shoes, cars) so that Matchers could reconstruct the Directors' array of cards. During one set of trials, the cards included multiple objects from the same category, such as a penny loafer, a sneaker, and a woman's pump. In this situation, Directors and Matchers swiftly adjusted toward using subordinate terms to refer to the items in each category—merely saying "shoe" would be infelicitous. In subsequent trials, however, the set of object cards was changed such that there was now only one exemplar from each category. The interlocutors, however, often used the same subordinate labels (e.g., "the

penny loafer") in the next trial, even though these expressions were now overly specific. Brennan and Clark found a similar pattern even across different Matchers: Directors who came to use terms like "penny loafer" with one Matcher often continued to use these terms with a completely new Matcher, although the target objects were now unique. Although this did not happen to the same extent as when the Matchers stayed constant, it clearly demonstrates how the strength of referential precedents can lead speakers to violate momentarily the maxim of quantity.

The notion of referential precedents also applies to the experiment we present here, in which participants engaged in another variation of the classic referential communication task. In our experiment, Directors described objects to Matchers so they could reconstruct an array of picture cards. Studies using this paradigm have consistently demonstrated that the process becomes more efficient over time. For example, Clark and Wilkes-Gibbs (1986) had pairs of Directors and Matchers engage in this type of task for six rounds. In Round 1, it took the Directors an average of 41 words to describe each figure before the Matchers were able to identify them; by Round 6 this average had fallen to eight words. This pattern demonstrates how interlocutors in a referential communication setting become increasingly efficient across rounds.

With respect to these tendencies toward consistency and increased efficiency, our experiment adds two important twists to the typical referential communication situation. First, the Directors in our experiment each worked with *two* independent Matchers. Second, in the early rounds of our task (Rounds 1–3), each Matcher had a different subset of the cards in the Director's matrix. In later rounds, the Matchers were given the entire set of cards, but only one Matcher at a time carried out the task: one did so immediately for three rounds (Rounds 4–6), while the other waited until the final three rounds (Rounds 7–9). At the transition points, Rounds 4 and 7, the Directors were faced with the task of describing certain figures for which they had a conceptual pact with the prior Matcher but not with the current Matcher. To perform optimally in these rounds, the Directors must depart from the general trend toward referential continuity and increased efficiency of expression. That is, they must develop and carry through on the awareness that extra effort is required to ensure the identification of previously nonshared objects. That this is an important

aspect of audience design is supported by results from Garrod and Clark (1993), who found that younger children were less successful in a referential communication task because they tended to stick with locally established communication schemes rather than monitoring the success of their utterances and adjusting when necessary.

Assessing audience design

To assess the extent to which Directors engage in audience design, we will restrict our attention to their utterances that are completed in advance of feedback from the Matchers. That is, we are particularly interested in the adjustments Directors make as they produce their initial descriptions for each referent. Note that this should not be taken to suggest that we believe that the Matchers have no role to play in shaping the Directors' descriptions (through, for example, responses like "House without the hat?"). On the contrary, we believe that feedback from the Matchers is a critical component of the Directors' experiences in our communicative situation (as we will document as part of our results). Even so, we wish to assess the extent to which, across our experiment, Directors become able to incorporate models of particular Matchers in advance of any feedback about the success or failure of individual descriptions. Note that we are neutral about whether speakers' adjustments to specific addressees are represented in initial utterance plans or whether they accrue as an utterance unfolds (Horton & Keysar, 1996; Keysar & Horton, 1998; Polichak & Gerrig, 1998; Schober & Brennan, in press). Instead, our approach here is to describe those factors that affect the likelihood that speakers' utterances will provide evidence of audience design, independent of the time course with which that evidence becomes incorporated into an utterance.

The current experiment

In summary, our experiment was designed to create a set of circumstances in which speakers come to realize the importance of tailoring referential expressions for particular addressees. Using a version of the standard referential communication paradigm (Clark & Wilkes-Gibbs, 1986; Krauss & Weinheimer, 1964, 1966), we gave matrices of picture cards to participants playing the role of Director. Their task was to describe these cards so that two independent Matchers would be able to reproduce the matrix. In our experiment, the Directors and Matchers participated in nine rounds of descriptions. As we mentioned earlier, in the early rounds of our task each Matcher possessed a different subset of cards in the Director's matrix. In the later rounds, however, each Matcher, for the rounds that he or she was present, was given the entire set of cards. If, in this situation, Directors are being appropriately attentive to which cards had not previously been shared with the current Matcher, then we would expect them to adjust their referring expressions to take this into account.

In addition, we used the contrast between object types exemplified in Fig. 1. As we argued earlier, pictures of common objects should (under the majority of circumstances) be more codable and should therefore require less negotiation about perspectives than the more ambiguous tangrams. As such, we believe that Directors should come to understand that audience design is most necessary—and therefore provide the most overt evidence for audience design—when they provide descriptions for the tangrams. The general aim of this experiment was to demonstrate that speakers' attempts to engage in audience design are guided by their experiences in a particular linguistic domain.

Method

Participants

Thirty-six undergraduate students (12 groups of three students) from the State University of New York at Stony Brook volunteered for this study as partial fulfillment for a psychology course requirement. All were native speakers of English.

Materials

We prepared three identical sets of picture cards (one Director's set and two Matchers' sets) as experimental stimuli. Each complete set of cards contained 18 items, consisting of nine color pictures of common living things taken from nature guides, and nine black and white abstract geometric tangram figures. Each picture was glued to a 6-in. × 4-in. index card and laminated. The set of pictures was subdivided into three categories:

three fish, three birds, and three flowers. They were chosen so that there was a reasonable amount of within-category similarity (e.g., all the flowers were pink) and so that no picture had an obvious subordinate label (e.g., we avoided easily identifiable exemplars like robins and goldfish). The tangrams were selected to fall into three general categories as well: three ''boats'' or boat-like shapes, three ''people'', and three ''rockets.'' Note, however, that these categories were mainly for purposes of the experimental design; participants could (and did) interpret the tangrams in any number of ways.

Design

The Director described the full set of 18 cards in all nine rounds. In Rounds 1–3, however, each Matcher saw a different subset of 12 cards. Each subset was created by removing one item category from the pictures and one from the tangrams (e.g., removing all three birds and all three boats for one Matcher and all three fish and all three rockets for the other Matcher). As a result, out of the complete set of 18, six cards were seen initially by the Director and both Matchers, six cards were seen initially by the Director and Matcher A only, and six cards were seen initially by the Director and Matcher B only. The categories removed to create each subset were completely counterbalanced across triads to create three versions of the experimental items. In the later rounds, each Matcher, when he or she was present (for either Rounds 4–6 only or Rounds 7–9 only), was given the entire set of 18 cards. See Table 1 for an illustration of the distribution of cards seen by each participant in each round of the experiment.

Procedure

For each triad of participants, one person was randomly selected to play the role of the Director, while the other two individuals were assigned as Matchers A and B. The Director sat at one end of a long table with a vertical barrier placed in front of his or her position such that there was enough room on the table to place the full set of 18 cards. At the other end of the table, two study carrels were placed back to back, perpendicular to the main table. The two Matchers each sat at one of the carrels, which served to block their view of the other participants and of each other's cards.

The participants were told that their goal was to get both Matchers' sets of cards arranged in the same order as the Director's set. Each person was also given an index card with a numbered grid printed on it marked with the numbers 1–18. It was explained that they should refer to this grid to help them keep track of where they were as they progressed through each round. They were told that it was very important to be accurate in placing the cards and that they should feel free to talk to each other as much as necessary to accomplish their goal.

The Director and Matchers were informed that some cards would be shared by all three persons, while others would be shared between the Director and only one or the other of the Matchers. In cases where they thought that the Director was describing a card that they did not have, the Matchers were instructed to leave a

Table 1
Sample distribution of cards possessed by each experimental participant in each round

	Director		Matcher A		Matcher B	
	Pictures	Tangrams	Pictures	Tangrams	Pictures	Tangrams
Rounds 1–3	Three birds	Three boats	Three birds	Three boats	Three birds	Three boats
	Three fish	Three people	Three fish			Three people
	Three flowers	Three rockets		Three rockets	Three flowers	
Rounds 4–6	Three birds	Three boats	Three birds	Three boats	—	—
	Three fish	Three people	Three fish	*Three people*	—	—
	Three flowers	Three rockets	*Three flowers*	Three rockets	—	—
Rounds 7–9	Three birds	Three boats	—	—	Three birds	Three boats
	Three fish	Three people	—	—	*Three fish*	Three people
	Three flowers	Three rockets	—	—	Three flowers	*Three rockets*

Note. The Matchers' cards in italics in Rounds 4 through 9 are the sets of cards that were coded as ''Nonshared'' for each Matcher in this configuration of cards.

"hole" in their arrangement of cards. At no time during the experiment were the Directors given any explicit information from the experimenter about which cards were seen by which Matcher. This could only be discovered through feedback from the Matchers.

The participants were also told that they would be carrying out the card-matching task nine times and that both Matchers would be present for Rounds 1–3, but that the Matchers would switch off for Rounds 4–6 and 7–9. Prior to Round 4, one Matcher was asked to step outside the room for the next three rounds and was then brought back in prior to Round 7, switching off with the other Matcher, who left the room for the final three rounds. The order in which Matcher A or Matcher B switched for Rounds 4–6 and 7–9 was counterbalanced across triads to avoid any confounding between picture sets and ordering.

At the beginning of each round, the experimenter shuffled the Director's set of cards and placed them on the table in a random order. The Director was told that this was the target arrangement for that round and that he or she should describe the cards for the Matchers in the order from 1 to 18. At the same time, the Matchers' cards were placed randomly in front of them. Prior to Rounds 4 and 7, the cards that had previously been missing from the Matchers' sets were added, giving them the full complement of 18 cards. The fact that all 18 cards were now present for both the Director and the Matcher in these later rounds was communicated to all participants. During each round, the experimenter wrote down the target order of the Director's cards for that round and feedback about any mismatches was given to the participants in between rounds.

Prior to the experiment, the participants took part in a brief practice round in which the director was given six cards, including both pictures and tangrams, while the Matchers were given partially overlapping sets of four cards each. This was intended to help them understand the nature of the distribution of the card sets and was also used as an opportunity to encourage the Matchers to give feedback to the Director about whether or not a match had been made. It was pointed out, for instance, that the Director often needed to know whether or not a particular description had been successful before he or she could move on. After completing the practice round, the participants were given their starting sets of experimental cards and told to begin. The entire session lasted about an hour.

Analysis and coding

We used two microphones to record the conversational interactions—one microphone was placed in front of the Director and the other was placed between the Matchers. Both microphones fed into the same tape recorder, which was paused and restarted between rounds. All audiotaped sessions were completely transcribed by the first author and a trained research assistant.

To assess audience design in our experiment, we coded the experimental transcripts with respect to a number of measures. We use combined evidence from these different measures to assess our hypotheses. To begin with, we briefly describe each measure:

1. Idea units. Earlier research on the referential communication task has typically provided global measures of communicative effort, such as the average number of words or turns Directors and Matchers need to identify each referent (e.g., Clark & Wilkes-Gibbs, 1986; Fussell & Krauss, 1992; Isaacs & Clark, 1987; Krauss & Weinheimer, 1966). In our experiment, not all Matchers had all of the cards in early rounds. As such, there was a noticeable amount of speech directed at helping the Matchers keep track of the current card position, as in this exchange:

M: Okay that's number seven, right?
D: That's number six.
M: That's number six- that's number six?
D: Yea.

Because inclusion of this sort of speech would drastically inflate simple counts of words and turns, we opted to use *idea units* as an alternative measure of effort (cf. Fussell & Krauss, 1992). To obtain this measure, we coded each content word (basically, each noun, adjective, or prepositional phrase) as belonging to a separate idea unit. For example, the topmost fish in Fig. 1 was described by one Director as *a fish, its fins are uh red, its body is blue and green, looking to the left, eyes sorta yellowish*. This utterance was coding as having nine idea units, corresponding to "fish," "fins," "red," "body," "blue," "green," "looking to the left," "eyes," and "yellowish." Our analyses will be based on the number of idea units produced by Directors in their first turns (i.e., their first utterances for each object). Idea units were coded independently by one author (WSH) and a research assistant. The small number of disagreements was resolved through discussion.

2. Hedges. When people engage in a referential communication task, utterances in early rounds

are often overtly marked to reflect the tentativeness of a particular perspective on an object (Brennan & Clark, 1996). These *hedges* most often take the form of phrases like *kinda*, *sort of*, and adjectival suffixes such as *-ish*. For example, in Round 1 one Director described the left tangram in Fig. 1 like this: *The third picture- sort of looks like uh- a person in a gown like a wedding dress it's- and uh and the head is a square that's slightly on a slant-.* By Round 9, however, the same Director was referring to this item simply as *the person in the gown.* For our purposes, hedges serve as an index of how certain Directors are that their description is adequate. For this analysis, both authors independently counted the number of hedges used by Directors in their initial description for each item (intercoder reliability was assessed using the intraclass correlation coefficient, which was calculated to be 0.91).

3. Entrainment. As a final index of audience design, we coded the extent to which Directors exhibited *lexical entrainment.* This directly assesses the consistency with which speakers describe an object for an addressee on multiple occasions (Bortfeld & Brennan, 1997; Brennan & Clark, 1996). We measured entrainment by comparing Directors' initial descriptions across pairs of rounds. Thus, in the basic coding of entrainment, initial descriptions in Round 2 were compared to those from Round 1, descriptions in Round 3 were compared to those from Round 2, and so forth. These comparisons were coded separately by both authors and any differences were resolved through discussion (intercoder reliability: $K = .89$).

With respect to entrainment, we were primarily interested in contrasting those cases in which descriptions contained the same words or were shortened across pairs of rounds with cases in which descriptions were lengthened or otherwise changed across rounds (cf. Bortfeld & Brennan, 1997). This provides a measure of the extent to which directors changed the way in which they conceptualized a given item from one description to the next. For example, in Round 3, one Director described the left tangram in Fig. 1 *as the guy at the podium facing the right, his head is not attached*, but subsequently described it in Round 4 as *a guy, the podium, with that dismembered head.* The change in the description of this figure's head from being "not attached" to "dismembered" was interpreted as a change in the director's conceptualization of this item, despite being a generally shorter description.

We have now described the set of measures we will use to evaluate our suggestion that experiential factors affect speakers' propensity to engage in audience design. Next, we begin the presentation of our results by noting the general similarity of our data to those reported in previous research with referential communication tasks. Then, we consider the results related to each of our major hypotheses.

Results

Overall patterns

Research using referential communication tasks has unfailingly demonstrated increased efficiency across rounds (e.g., Clark & Wilkes-Gibbs, 1986; Hupet et al., 1991; Krauss & Weinheimer, 1964, 1966; Schober & Clark, 1989). In terms of this overall result, our data closely parallel those past reports. For example, the mean number of idea units produced by Directors in their first turns declined from Round 1 ($M = 7.25$ idea units per item) to Round 9 ($M = 3.71$ idea units per item). To assess the reliability of this trend, as well as the other effects we report, we conducted analyses of variance with both participants ($F1$) and items ($F2$) as random variables. Those analyses confirm that directors produced reliably fewer first turn idea units as the rounds progressed (linear trend: $F1(1, 88) = 180.5$, $MSe = 1.083$, $p < .0001$; $F2(1, 128) = 456.75$, $MSe = .325$, $p < .0001$). Our other measures exhibited a similar pattern. The number of hedges in the Directors' initial descriptions generally decreased from Round 1 ($M = 1.75$ hedges per item) to Round 9 (M = 0.07 hedges per item; linear trend: $F1(1, 88) = 192.17$, $MSe = .154$, $p < .0001$; $F2(1, 128) = 372.84$, $MSe = .060$, $p < .0001$). Also, lexical entrainment increased across rounds as Directors became more consistent in their initial descriptions. Thus, there were significantly more departures from entrainment in Round 2 compared to Round 1 (82% involved a reconceptualization of the previous description) than in Round 9 compared to Round 8 (29% involved a reconceptualization; linear trend: $F1(1, 77) = 127.07$, $MSe = .123$, $p < .0001$; $F2(1, 112) = 163.91$, $MSe = .071$, $p < .0001$). These analyses confirm that the Directors became generally more efficient, certain, and constant in their first descriptions across the course of the experiment. The focus of our analyses, however, will be to

show that, against the background of this trend toward short, consistent descriptions, our participants engaged in additional effort at those points at which audience design considerations became most important.

Evidence for audience design

In the first three rounds of our experiment, each Director and Matcher shared only a subset of picture cards. In that context, evidence for audience design in subsequent rounds would be differential performance with respect to items that were *shared* vs. *nonshared*. Specifically, at Rounds 4 and 7 we expected that, in describing those items that had previously been shared with the current Matcher, Directors would use fewer idea units, fewer hedges, and that their descriptions would show a greater degree of entrainment, compared to those objects that had not been shared. Conversely, those items that had *not* been shared with the current Matcher should prompt Directors to produce more idea units, more hedges, and be more likely to depart from a previous conceptualization. To assess these predictions, we carried out analyses for each of our measures with round (Round 4 vs. Round 7), object type (picture vs. tangram), and sharedness (shared vs. nonshared) as factors. For the purposes of this and subsequent analyses, we counted as ''shared'' only those cards that were seen by the Director and *one* of the Matchers, excluding those cards shared among all three participants. This provides a cleaner contrast with the nonshared items, because the cards that are shared with only one Matcher are by definition not shared with the other Matcher.

The means for these analyses for our first two measures of audience design are presented in Table 2. As predicted, at Rounds 4 and 7 the Directors used fewer idea units in their initial

descriptions when discussing items they had shared previously with the current matcher ($M = 4.54$ idea units per item) compared to items that they had not shared ($M = 4.83$ idea units per item; $F1(1, 11) = 8.15$, $MSe = .251$, $p < .02$; $F2(1, 16) = 3.56$, $MSe = .465$, $p < .08$). The fact that Directors said more when describing the nonshared items in Rounds 4 and 7 suggests that they were designing their initial descriptions with their audience's needs in mind.

With respect to hedges, there was not a reliable difference in the extent to which Directors chose to hedge their first turn descriptions at Rounds 4 and 7 for shared versus nonshared items. There was, however, a significant interaction of sharedness and item type ($F1(1, 11) = 5.38$, $MSe = .039$, $p < .05$; $F2(1, 16) = 4.90$, $MSe = .035$, $p < .05$). As can be seen in Table 2, Directors were most likely to hedge their initial descriptions for nonshared tangrams. We suggested in the introduction that the nature of the negotiations concerning the tangrams would potentially enable Directors to be relatively sensitive to the particular referential status of these items. The hedges provide evidence in support of this special sensitivity.

Further evidence on this point comes from our entrainment measure. We compared the Directors' initial descriptions given in both Rounds 4 and 7 with those produced in Round 3. This allowed us to compare descriptions from each of the two transition rounds against the same baseline description. These data are presented in Table 3 as the percentage of reconceptualizations. We found that Directors were more likely to depart from their previous conceptualizations for nonshared items (55% of the initial descriptions for nonshared items involved reconceptualizations) than for shared items (46% involved reconceptualizations; $F1(1, 11) = 7.41$, $MSe = .026$, $p < .02$; $F2(1, 16) = 4.97$, $MSe = .030$, $p < .05$). Moreover, the proportion of

Table 2
Mean number of idea units and hedges in Directors' initial object descriptions at Rounds 4 and 7, organized by object type and sharedness

	Pictures		Tangrams	
	Shared	Nonshared	Shared	Nonshared
Round 4				
Idea units per item	4.33	4.77	4.47	4.67
Hedges per item	0.22	0.17	0.39	0.36
Round 7				
Idea units per item	4.22	4.57	5.14	5.39
Hedges per item	0.17	0.18	0.28	0.64

Table 3
Mean percentages of Directors' initial object descriptions in Rounds 4 and 7 containing reconceptualization of previous descriptions (both compared to Round 3), plus the mean percentages of descriptions containing utterance-medial and utterance-final reconceptualizations, organized by object type and sharedness

	Pictures			Tangrams		
	Shared	Nonshared	Mean	Shared	Nonshared	Mean
Round 4						
Reconceptualizations	42	44	43	44	53	49
Utterance-medial	28	33	31	36	44	40
Utterance-final	14	11	13	8	9	9
Round 7						
Reconceptualizations	39	47	43	58	75	67
Utterance-medial	22	33	28	42	47	45
Utterance-final	17	14	16	16	28	22

reconceptualizations was higher for tangrams ($M = 58\%$) than pictures ($M = 43\%$; $F1(1, 11) = 6.94$, $MSe = .074$, $p < .03$; $F2(1, 16) = 8.82$; $MSe = .043$; $p < .01$). There was also a reliable interaction between item type and round: As seen in Table 3, Directors' descriptions exhibited the highest proportion of departures from entrainment for tangrams at Round 7 ($F1(1, 11) = 5.36$; $MSe = .037$, $p < .05$; $F2(1, 16) = 5.73$, $MSe = .026$, $p < .03$). In general, however, Directors were most likely to offer reconceptualizations of previous descriptions when describing nonshared items.

Although this result for our entrainment measure is suggestive, it does not provide unambiguous support for our concept of audience design. Consider this pair of initial descriptions produced by one Director for the same item across successive rounds:

> Round 6: Picture two is the spiny flower.
> Round 7: Picture two is the spiny flower with pink needle-like petals.

The second description constitutes a reconceptualization by virtue of additional information. However, we have no way of determining whether the added clause was planned in advance of an expectation of feedback (in which case we would want to count it as audience design), or whether it arose interactively because the Matcher did not provide any immediate feedback following the first part of the utterance.[1]

To refine our analyses of entrainment, we recoded our data to differentiate between instances in which Directors' reconceptualizations were achieved by the addition of new material at the end of the entrained expression, and instances in which new material was introduced prior to or internal to the entrained expression, as in this pair of descriptions:

> Round 6: Okay the next one is the fish with the red fins.
> Round 7: And the next one is the blue and green fish with uh red and green tails and fin.

These circumstances allow for a stronger claim that the reconceptualization was planned in advance of an expectation of feedback from the Matcher.

We coded each changed description as belonging to one of the three categories:

- *Utterance final*. The Director first repeated the previous conceptualization and subsequently provided additional modifying information.
- *Utterance medial*. The Director modified the previous conceptualization with new information that either preceded or was interleaved within the old material.
- *Completely new*. The Director offered a completely different conceptualization of the same object.

Using this coding scheme, we calculated the percentage of total initial descriptions given in Rounds 4 and 7 that were either completely new or that contained utterance-medial reconceptualizations when compared to the descriptions in Round 3. The means for this measure are provided in Table 3. Note, first, that in those cases in which Directors offered reconceptualizations of

---

[1] Our first impulse was to listen to the original tapes and to see if we could deduce the speakers' intentions from, for example, intonation contours. However, because of differences among the Directors' speech patterns, we quickly came to believe that this would not be a feasible measure.

earlier descriptions, new content preceded (or was intermixed with) previously given expressions 71% of the time (e.g., for the shared pictures at Round 4, 28/42 = 0.67). Clearly, the norm was not the addition of material at the end of the utterances.

More specific results also emerged. In general, out of the total number of first descriptions in these two rounds, there was a higher proportion of utterance-medial or completely new reconceptualizations for tangrams ($M = 42\%$) than for pictures ($M = 29\%$; $F1(1, 11) = 6.77$, $MSe = .062$, $p < .03$; $F2(1, 16) = 8.11$, $MSe = .039$, $p < .02$). This suggests that the Directors were most prepared to vary their descriptions of tangrams when formulating their initial utterances. In addition, the percentage of such intrusions of new information was greater for nonshared ($M = 39\%$) than for shared items ($M = 32\%$). Although this effect of sharedness was marginal ($F1(1, 11) = 3.84$, $MSe = .037$, $p = .076$; $F2(1, 16) = 4.10$, $MSe = .026$, $p = .059$), the pattern suggests that Directors were relatively more prepared to modify their descriptions for the nonshared items. Moreover, it is important to note that this measure potentially underrepresents instances in which Directors' reconceptualizations were planned in advance of an expectation of Matcher feedback. It was likely that some subset of instances in which new material followed an entrained expression was planned as such from the outset. Thus, we believe that this relatively conservative measure provides reasonable evidence that Directors were often engaging in audience design rather than responding to the Matchers' lack of feedback.

The matchers' contributions

Our analyses thus far have focused on the factors that prompt Directors to expend effort toward audience design. However, as we suggested in the introduction, audience design in our task should be influenced by the experiences that Directors gather through negotiation with the Matchers. Indeed to the extent that Matchers provide evidence of their understanding through feedback, Directors should have the opportunity to learn how to adjust their descriptions appropriately (cf. Kraut, Lewis, & Swezey, 1982).

To assess the nature of the contributions from the Matchers, we undertook an analysis of the content of the Matchers' responses that immediately followed the Directors' initial descriptions. We coded these responses into one of the five categories:

- *Acceptance*: the Matcher indicated successful identification through a simple "Got it" or "Okay."
- *Interruption-acceptance*: the Matcher interrupted the Director's description to indicate acceptance.
- *Interruption-query*: the Matcher interrupted the Director's description to ask for additional information or to indicate some kind of difficulty.
- *Clarification request*: the Matcher requested clarification of some portion of the Director's previous description (e.g., asking "Square at the top?" after hearing "Number one is the one that has a square at the top connected to the rest of it with that tiny line.").
- *Expansion request*: the Matcher either implicitly requested an expansion of the previous description (most often by saying "umm-") or sought confirmation for a proposed expansion (e.g., asking "With four points on the sides?" after hearing "Number fourteen is like the house, the treehouse- black and white.").

The percentages for each of these categories for Rounds 4 and 7 are presented in Table 4. As is apparent from the table, the most frequent response from the Matchers was acceptance of the initial description. Including acceptances that occurred as an interruption, 71% of the Matchers' responses could be categorized as acceptances. For the majority of items in these two rounds, then, Directors were able to provide sufficient information in their initial descriptions to allow Matchers make a positive identification of the intended referent. It should also be noted that Matchers interrupted the Directors relatively infrequently: only 14 out of 288 (4.8%) first turn descriptions in Rounds 4 and 7 were interrupted by the Matcher. These data suggest that, in our task, Matchers rarely used interruptions to shape the Director's behavior.[2]

Despite the generally high rate of acceptances following the Directors' initial object descriptions,

---

[2] Although Matchers rarely interrupted, our observations were that they were often able to identify objects before the Directors finished their initial utterances. In that sense, Matchers were providing less information to the Directors than they might have provided—had they been willing to be somewhat less patient.

Table 4
Percentages of Matchers' responses following the Directors' initial object descriptions in Rounds 4 and 7 ($N = 36$ per column), classified according to the type of response (see text for an explanation of coding categories)

	Pictures		Tangrams	
	Shared	Nonshared	Shared	Nonshared
Round 4				
Acceptance	86	72	69	42
Interruption-acceptance	3	3	3	0
Interruption-query	0	0	3	3
Clarification request	0	0	0	11
Expansion request	11	25	25	44
Round 7				
Acceptance	75	58	81	50
Interruption-acceptance	14	3	8	0
Interruption-query	0	0	0	0
Clarification request	3	8	8	19
Expansion request	8	31	3	31

an inspection of Table 4 reveals that Matchers were more likely to seek additional information, either as a clarification or concerning an expansion, when the object under discussion had not been previously shared. We collapsed the three categories of nonacceptances into a single category of "queries" and found that, overall, such queries occurred reliably more often in response to descriptions for the nonshared (21.5% of the Matchers' responses) than the shared (7.6% of the responses) items ($F1(1, 11) = 64.71$, $MSe = .029$, $p < .001$; $F2(1, 16) = 22.38$, $MSe = .062$, $p < .001$). In addition, Matchers were more likely to seek further information following Directors' descriptions for the tangrams (18.4% of the Matchers' responses) than the pictures (10.8% of the responses) type ($F1(1, 11) = 9.31$, $MSe = .060$, $p < .02$; $F2(1, 16) = 9.98$, $MSe = .042$, $p < .01$). The interaction between sharedness and item type did not reach significance (both p's $> .20$), but the interaction between round and item type did ($F1(1, 11) = 6.49$, $MSe = .035$, $p < .03$; $F1(1, 16) = 7.84$, $MSe = .022$, $p < .02$); Matchers were particularly likely to request for further information about tangrams at Round 4. As a whole, these patterns of acceptances and queries suggest that Matchers provided feedback to Directors that particular categories of objects required ongoing effort.

Thus, we have seen that Matchers often provided alternative conceptualizations for certain items, most frequently by expanding upon the Directors' proposed descriptions. But did Directors then incorporate this information into their descriptions of these objects in later rounds? That

is, to what extent were the Directors' descriptions the joint product of contributions from both the Directors and Matchers? To address this question, we noted each instance in which a Matcher suggested a new conceptualization for a given item and examined whether the Director incorporated that information as part of his or her descriptions in subsequent rounds. We found that 64% of the Directors' descriptions following such feedback did *not* contain any information originally proposed by a Matcher. That is, even when Matchers provided potentially useful information about their perspectives, the Directors primarily stayed with their own perspectives in deciding how to describe these objects. Furthermore, it did not seem to matter whether an object had been previously shared or not shared with the current Matcher: Examining those cases from Round 4 and beyond in which the Director's initial descriptions did incorporate prior feedback, we found that the probability of such incorporations was 35% for shared items and 36% for nonshared items. This does not, of course, mean that the descriptions were not Matcher-specific; the entrainment measure suggests that they were. What we largely found, rather, was that Directors adjusted to Matchers by modifying the perspectives they themselves had used rather than incorporating conceptualizations proposed by the Matcher.

Our final analysis with respect to the Matchers' responses concerned the extent to which they provided feedback concerning the distribution of cards in the first three rounds. As we noted earlier, the experimenter never gave Directors any explicit

information about which Matcher had which cards. However, for 10 of the 12 triads, the Matchers themselves provided Directors with at least some information of this type, as in this example from one triad during Round 2:

> D: Okay then there's the purple and blue bird.
> MA: Mm-
> MB: I have that.
> MA: I don't have that.

Such moments provided Directors with the opportunity to learn, in a concrete fashion, about the initial distribution of cards. Even so, our analyses suggest that the differences we predicted emerged even in the triads for which Matchers provided no overt feedback. Consider our measure of first turn idea units. The two triads that lacked any explicit feedback concerning the distribution of cards yielded a 0.46 difference for shared vs. nonshared objects (4.71 vs. 5.17 idea units per item). The ten triads that included such feedback yielded a 0.20 difference (4.51 vs. 4.71 idea units per item). This informal analysis suggests that overt feedback was not necessary to produce an impact on audience design.

Learning that audience design is necessary

In the introduction, we suggested that Directors must come to understand that audience design is necessary. Our argument was that there are particular circumstances in which speakers may begin without any specific awareness about the need for audience design, but then learn through experience about the importance of design considerations. To address this issue, we carried out analyses of our data that examined the transitions between descriptions from Rounds 3 to 4 and from Rounds 6 to 7. Our reasoning was as follows: In Rounds 1–3, Directors accumulated evidence for which objects were shared and not shared with each Matcher. Because, however, both Matchers were present for the interactions for both types of objects, Directors had little opportunity to learn about the consequences of a particular object being nonshared with an addressee. It was only at Round 4 and beyond that the significance of some of the objects having been initially nonshared would become evident to the Directors. Therefore, we predicted that, all else being equal, Directors would provide more evidence of audience design at the second, Round 7 transition than at the first, Round 4 transition.

For this set of analyses, then, we considered only those descriptions for objects that were *new* to each Matcher in the critical transition rounds. In other words, for each Director, we eliminated from consideration those items that changed the status from nonshared to shared across the transitions, examining only those instances in which the Director described objects in the post-transition round (4 or 7) that had *not* been shared previously with the current Matcher. This had the advantage of allowing us to compare Directors' descriptions for the same item as it went from being described to a Matcher who had shared that object previously to being described to a Matcher for whom that object was new.

The data for these items for the first turn idea units and for the hedges are presented in Table 5. We carried out analyses of variance with transition point (3–4 vs. 6–7), sharedness (shared vs. nonshared), and object type (picture vs. tangram) as factors. The data suggest that Directors' performance across the two transitions was dissimilar. Consider the idea units produced by Directors in their first turns. During the early transition, the number of idea units fell from Rounds 3 to 4

Table 5

Mean number of idea units and hedges in the Directors' initial object descriptions, for items that were classified as shared for Rounds 3 and 6 and as nonshared for Rounds 4 and 7, organized by object type

	Pictures		Tangrams	
	Round 3 shared	Round 4 nonshared	Round 3 shared	Round 4 nonshared
First transition (Rounds 3–4)				
Idea units per item	5.67	4.77	5.86	4.67
Hedges per item	0.17	0.17	0.47	0.36
	Round 6 shared	Round 7 nonshared	Round 6 shared	Round 7 nonshared
Second transition (Rounds 6–7)				
Idea units per item	4.11	4.57	3.69	5.39
Hedges per item	0.03	0.17	0.14	0.64

(mean difference across rounds $= -1.05$ idea units per item). We suspect that this represents a continuation of the trend toward shortening of descriptions across Rounds 1–3. By comparison, from Rounds 6 to 7 the number of idea units increased (mean difference across rounds $= +1.08$ idea units per item). In our view, this change reflects a fairly impressive effort on the Directors' part to overcome the tendency (now built up over six rounds) toward shorter descriptions. This interaction of transition point and sharedness was reliable ($F1(1, 11) = 14.09$, $MSe = 1.97$, $p < .005$; $F2(1, 16) = 26.63$, $MSe = .730$, $p < .001$). Additionally, planned contrasts were carried out on the mean number of first turn idea units for the pictures and tangrams separately, at both transitions. The only reliable difference was the significant increase in idea units at the second transition point for the tangrams ($F1(1, 11) = 8.18$, $MSe = 2.11$, $p < .02$; $F2(1, 16) = 12.92$, $MSe = .730$, $p < .001$). This sensitivity to the shared status of the tangrams across the second transition point suggests that the Directors had learned the importance of adjusting their descriptions by the time they reached Round 7, and they were particularly prepared to make such adjustments with respect to the tangrams.

For the hedges, the pattern of data is similar. Directors tended to produce slightly fewer hedges in their initial descriptions in Round 4 compared to Round 3 (mean difference across rounds $= -0.06$ hedges per item), while the number of hedges increased from Rounds 6 to 7 (mean difference $= +0.32$ hedges per item). This interaction between transition and sharedness was reliable ($F1(1, 11) = 14.99$, $MSe = .058$, $p < .003$; $F2(1, 16) = 10.72$, $MSe = .059$, $p < .005$). Again, this demonstrates that by the second transition point Directors seemed to have learned that it was necessary to modify their initial descriptions, marking the nonshared items in particular as provisional through the use of hedges. Consistent with our other findings, this was particularly true

for the tangrams: Planned contrasts conducted on the tangram data revealed a reliable difference in the prevalence of hedges across the second transition point ($F1(1, 11) = 9.69$, $MSe = .155$, $p < .01$; $F2(1, 16) = 19.06$, $MSe = .059$, $p < .001$). The nature of the interactions about the tangrams apparently led the Directors to adjust their descriptions for this set of items in particular, especially at the second matcher transition.

Next, Table 6 presents the entrainment data for our comparison across the two critical transition points, reported as the percentage of reconceptualizations in Directors' post-transition initial descriptions. For this measure, we compared descriptions between Rounds 4 and 3 and between Rounds 7 and 6. In parallel with our other measures, Directors were less likely to maintain their previous conceptualization across the transition from Rounds 6 to 7 ($M = 66\%$ reconceptualizations) than at the transition from Rounds 3 to 4 ($M = 49\%$ reconceptualizations; $F1(1, 11) = 8.25$, $MSe = .040$, $p < .02$; $F2(1, 16) = 7.78$, $MSe = .032$, $p < .02$). Unfortunately, when we focus on the subset of departures from entrainment that reflect intrusion of new information (also shown in Table 6), the data become somewhat less univocal. Although we did find that utterance-medial reconceptualizations were more likely to occur when Directors described tangrams ($M = 43\%$ of the total first turn descriptions) than pictures ($M = 28\%$; $F1(1, 11) = 8.59$, $MSe = .033$, $p < .02$; $F2(1, 16) = 6.05$, $MSe = .035$, $p < .03$), the rate of utterance-medial intrusions was actually slightly smaller in the transition from Rounds 6 to 7 ($M = 32\%$ of the total first turn descriptions) than in the transition from Rounds 3 to 4 ($M = 39\%$; both F's < 1). In fact, when we calculate the conditional probability that a given reconceptualization was utterance-medial, the proportion of intrusions falls dramatically from the first transition ($M = 80\%$; i.e., $39/49 = 0.795$) to the second ($M = 49\%$: $F1(1, 6) = 5.02$, $MSe = .134$, $p = .06$; $F2(1, 14) = 7.77$, $MSe = .081$,

Table 6

Mean percentages of Directors' initial object descriptions at Round 4 (compared to Round 3) and Round 7 (compared to Round 6) coded as reconceptualizations, plus the mean percentages of descriptions containing utterance-medial and utterance-final reconceptualizations, organized by object type

	Round 4			Round 7		
	Pictures	Tangrams	Mean	Pictures	Tangrams	Mean
Reconceptualizations	44	53	49	50	81	66
Utterance-medial	33	44	39	22	42	32
Utterance-final	11	9	10	28	39	34

$p < .02$).[3] In other words, when Directors modified their descriptions of the same object across a transition, at the second transition they were more likely to do so by adding new material following an entrained description.

This result calls attention to the experience gap between the two transitions. At the first transition point, Directors had described only these objects three times previously; the descriptions were often still in a state of relative flux (i.e., it took some time for the Directors to settle upon a particular way of describing certain items). At the second transition point, however, the Directors had described these objects six times to an informed Matcher and were now confronted with having to describe these items again for a Matcher who had not seen them before. At this point, the Directors may have chosen to begin with the established descriptions not only because these descriptions were readily available, but also as a way of communicating the preferred description to the new Matcher.[4] Even so, because the Directors were aware of the need to make adjustments for the benefit of the naïve Matchers, these descriptions could have been planned with the intention of adding more information at the end. However, as we mentioned earlier, we have no principled way of discerning when utterance-final reconceptualizations were the result of a lack of feedback from the Matcher or when they were actually planned this way.

Taken together, these data are largely consistent with our suggestion that Directors are learning something across the course of the experiment. Specifically, we believe that feedback from the Matchers at the first transition is likely to have made Directors more aware at the second transition that extra effort was needed. Of course, the fact that there is a different partner at the transition to Round 7 could have prompted Directors to change their descriptions for the new Matcher, regardless of the shared status of items. However, we feel that the transition data must be interpreted in conjunction with our previous analyses, which showed that extra effort (as shown across several measures) was being made for the nonshared items in particular. This suggests that Directors were more prepared to alter their descriptions at the second transition when describing items that had not been previously shared with the current Matcher. Note also that this interpretation suggests that the Directors had encoded something about the differential distribution of cards in the first three rounds, but had to learn through the interactions at Round 4 and beyond that it was important to apply this information in formulating descriptions.

General discussion

We intended this experiment to provide support for an experiential approach to the question of speakers' ability to engage in audience design. We suggested that successful audience design has at least two important preconditions: Speakers must be given the opportunity to learn how to make the appropriate adjustments in a particular situation, and they must put forth the effort to overcome the trend toward shorter, more consistent forms of expression. We explored these points in a referential communication experiment in which we varied the codability of objects in the Directors' and Matchers' arrays (i.e., pictures versus tangrams) and also varied the distribution of shared versus nonshared experiences in talking about those objects. Our measures confirmed that, to a substantial degree, the Directors were sensitive to the distinction between items that had been previously shared and not shared with Matchers. This result is striking given the fact that shared and nonshared items were completely intermixed. Directors had to decide, on a case-by-case basis, how to tailor their description for a given object.

We hoped that our experiment would capture important features of real-world interactions, in which it is often the case that a speaker would have different patterns of shared and nonshared knowledge with different addressees. We believe that our results should generalize to those real-world situations. Speakers will take particular care to consider their addressees—to customize their utterances—in circumstances in which they are discussing topics or objects with which they have a prior history of negotiation about perspectives (Brennan & Clark, 1996; Kraut et al., 1982). Conversely, when discussing topics or objects that are relatively free from the need for negotiation, speakers may be less likely to consider their addressees as a routine element of utterance planning.

There are, however, some important ways in which our experiment differed from an ideal

---

[3] The reduced degrees of freedom in this analysis are due to the presence of empty cells.

[4] We thank Shari Speer for suggesting this interpretation.

analogy to real-world circumstances. First, although each Matcher did not have visual access to the nonshared objects, they did overhear the negotiations that passed between the Director and the other Matcher about those items in the first three rounds. As a result, the Directors may have had reason to believe that the Matchers had at least some level of awareness of the range of objects (and perspectives on those objects) that formed the set of nonshared items. This belief could have led to descriptions for these items that assumed something other than a completely naïve perspective (Wilkes-Gibbs & Clark, 1992). Presumably, such an assumption on the Directors' part would have attenuated the impact of the shared versus nonshared distinction (i.e., our results might underrepresent speakers' capacity to make this distinction). We designed the experiment as we did, however, because we wanted the shared and nonshared objects for each Matcher to be randomly intermixed as the first three rounds unfolded. We wanted our experiment to reflect the ordinary circumstances of language use, in which there is incomplete information about distributions of knowledge among interlocutors.

Note, however, that when we designed the experiment, we anticipated that Directors would be able to use the categorical structure of the stimuli to alleviate some of the memory burden inherent in the situation (e.g., they would infer that Matcher A had the boat-like tangrams and Matcher B had the people-like tangrams). We recognize the possibility that our evidence in favor of audience design may again have been attenuated because of the difficulties with encoding particular categories of items with particular partners. In retrospect, we might have explicitly called the Directors' attention to the categories within our set of objects, particularly for the tangrams. We are currently carrying out a project that directly examines this relationship between audience design and the ease of partner-specific encoding (Gerrig & Horton, 2001).

A second concern with our experimental design, with respect to real-world instances of language use, is the distinction between which Matchers knew what was relatively arbitrary. In real-world circumstances, it seems likely that distributions of knowledge would be better motivated: Speakers would adapt their referring phrases because they know, for example, that one friend shares their knowledge of ichthyology whereas another friend does not (Isaacs & Clark, 1987). Because the distributions of knowledge in

our experiment were arbitrary, we made the task somewhat more of a pure memory exercise for the Directors than might ordinarily be the case. Directors were called upon to remember who shared what, without any ready mnemonic assistance. However, as with our concerns about overhearing, these considerations suggest that we have demonstrated the transition toward more effective audience design under more adverse circumstances than might govern ordinary real-world practice, particularly with respect to the tangrams. If Directors were given more specific reasons to care about the differences in knowledge available to each Matcher, then we might begin to find more generally robust effects. Such a result would reinforce the more general point that audience design processes are sensitive to situational considerations rather than being all or nothing.

We must also acknowledge the fact that the referential communication paradigm has some limitations when it comes to investigating the role of audience design. First, the standard referential communication situation requires speakers to refer to the same objects across multiple rounds. As a result, any adjustments that occur over time may be specific to the particular set of items being described repeatedly rather than arising from more general processes. Second, because the referential communication task is inherently interactive, the contributions of one conversational partner are not independent of contributions made by the other partner. The interactive nature of this task is part of its appeal, for it has allowed researchers to investigate, in a relatively controlled setting, how conversations emerge out of such interactions. Even so, when making claims about the role of audience design in this situation, it is necessary to recognize the fact that individual utterances are not designed in a vacuum. In the current paper, we have attempted to address this concern by both examining the content of the Matchers' responses and by giving particular attention in our entrainment analyses to utterance-medial reconceptualizations.

We note, finally, that our experiment speaks neither about the time-course with which addressee-specific information is brought to bear on utterance planning nor about the automaticity of audience design. In our experiment, as with the bulk of experiments that have used referential communication tasks, we can assess only the evidence of audience design from Directors' and Matchers' overt utterances. For example, we know from our experiment that Directors were

most likely to hedge their initial referring phrases when describing the nonshared tangrams. What we do not know, however, is whether the impulse to hedge comes early or late in production. It could be the case that the initial utterance plans for all referring phrases included the intention to hedge, but that this intention was often edited out in the course of production. Alternatively, it could be the case that the intention to hedge was added to the utterance plans when it was most appropriate to do so, as for the nonshared tangrams. Although our data cannot help decide between different moment-by-moment models of language production, we believe our results do constrain the overt behavior those models should explain.

We intend the perspective we have developed in this paper to argue in favor of an experiential approach to considerations of audience design. Speakers may, in some general sense, aspire to optimality but require some criterion level of experience with a particular type of language use—with a criterion level of interaction—to approach that goal. In particular, our data demonstrate why it is necessary to divorce the question of people's intentions toward audience design from the adequacy of their efforts. Speakers may intend quite sincerely to tailor their productions for a specific audience, but lack the knowledge or resources to carry out these intentions fully. This position is entirely consistent with the view of conversation as an active, collaborative process (Clark, 1996). Discovering not only when to adjust to one's audience, but also how to adjust, is clearly an important part of being a cooperative conversationalist.

References

Bayard, L. (1999). *Fool's errand*. Los Angeles: Alyson Books.

Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology, 18*, 355–387.

Bock, J. K., & Loebell, H. (1990). Framing sentences. *Cognition, 35*, 1–39.

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

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

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

Clark, H. H., & Carlson, T. B. (1981). Context for comprehension. In J. Long & A. Baddeley (Eds.), *Attention and performance IX* (pp. 313–330). Hillsdale, NJ: Lawrence Erlbaum.

Clark, H. H., & Murphy, G. L. (1982). Audience design in meaning and reference. In J.-F. Le Ny & W. Kintsch (Eds.), *Language and comprehension* (pp. 287–299). Amsterdam: North-Holland.

Clark, H. H., Schreuder, R., & Buttrick, S. (1983). Common ground and the understanding of demonstrative reference. *Journal of Verbal Learning and Verbal Behavior, 22*, 245–258.

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

Fussell, S. R., & Krauss, R. M. (1992). Coordination of knowledge in communication: effects of speakers' assumptions about what others know. *Journal of Experimental Social Psychology, 62*, 378–391.

Garrod, S., & Anderson, A. (1987). Saying what you mean in dialogue: a study in conceptual and semantic co-ordination. *Cognition, 27*, 181–218.

Garrod, S., & Clark, A. (1993). The development of dialogue co-ordination skills in school children. *Language and Cognitive Processes, 8*, 101–126.

Gerrig, R. J., & Horton, W. S. (2001, November). *Judgments about common ground in language production*. Paper presented at the 42nd Annual Meeting of the Psychonomic Society, Orlando, FL.

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

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

Hupet, M., Seron, X., & Chantraine, Y. (1991). The effects of codability and discriminability of the referents on the collaborative referring procedure. *British Journal of Psychology, 82*, 449–462.

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

Keysar, B., & Horton, W. S. (1998). Speaking with common ground: from principles to processes in pragmatics: a reply to Polichak and Gerrig. *Cognition, 66*, 191–198.

Krauss, R. M., & Weinheimer, S. (1964). Changes in reference phrases as a function of frequency of usage in social interaction: a preliminary study. *Psychonomic Science, 1*, 113–114.

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

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

Lachman, R., Shaffer, J. P., & Hennrikus, D. (1974). Language and cognition: effects of stimulus codability, name-word frequency, and age of acquisition on lexical reaction-time. *Journal of Verbal Learning and Verbal Behavior, 13*, 613–625.

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

Schober, M. F., & Clark, H. H. (1989). Understanding by addressees and overhearers. *Cognitive Psychology, 21*, 211–232.

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

Wilkes-Gibbs, D. (1995). Coherence in collaboration: some examples from conversation. In M. A. Gernsbacher & T. Givon (Eds.), *Coherence in spontaneous text* (pp. 239–267). Amsterdam: John Benjamins.

Wilkes-Gibbs, D., & Clark, H. H. (1992). Coordinating beliefs in conversation. *Journal of Memory and Language, 31*, 183–194.