COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 46 (2022) e13077

ISSN: 1551-6709 online
DOI: 10.1111/cogs.13077

Encoding Motion Events During Language Production: Effects of Audience Design and Conceptual Salience

Monica Lynn Do,[a] Anna Papafragou,[b] John Trueswell[a]

[a]*Department of Psychology, University of Pennsylvania*
[b]*Department of Linguistics, University of Pennsylvania*

Received 28 April 2021; received in revised form 15 November 2021; accepted 25 November 2021

Abstract

We investigate the extent to which pragmatic versus conceptual factors can affect a speaker's decision to mention or omit different components of an event. In the two experiments, we demonstrate the special role of pragmatic factors related to audience design in speakers' decisions to mention conceptually "peripheral" event components, such as sources (i.e., starting points) in source-goal motion events (e.g., a baby crawling from a crib to a toybox). In particular, we found that pragmatic factors related to audience design could not only drive the decision to omit sources from mention, but could also motivate speakers to mention sources more often than needed. By contrast, speaker's decisions to talk about goals did not appear to be fundamentally driven by pragmatic factors in communication. We also manipulated the animacy of the figure in motion and found that participants in our studies treated both animate and inanimate source-goal motion events in the same way, both linguistically and in memory. We discuss the implications of our work for message generation across different communicative contexts and for future work on the topic of audience design.

Keywords: Audience design; Goal bias; Language production; Message generation; Motion events

1. Introduction

When describing an event in the world, speakers do not simply mention everything about that event; instead, they make choices about what to say and what to omit. One of the most important factors affecting these choices is the perceived knowledge state and informational needs of their addressee. In a process known as *audience design*, speakers often adjust their utterances to facilitate comprehension for listeners at large. It is well-known, for instance, that speakers tend to mention things which are atypical or unlikely to be inferred by a generic

This research was supported by the National Science Foundation the grant 1632849 to Anna Papafragou.
Correspondence should be sent to Monica Lynn Do, Department of Linguistics, University of Chicago, 1115 E. 58th Street, Rosenwald Hall Room 229C, Chicago, IL 60637, USA. E-mail: monicado@uchicago.edu

addressee within their cultural or linguistic community (Brown & Dell, 1987; Clark & Marshall, 1981; Fussell & Krauss, 1992; Grigoroglou & Papafragou, 2019a). In addition, speakers are able to adjust their utterances to meet the moment-by-moment needs of their specific interlocutor(s). They do this by taking into account their listener's visual perspective (Grigoroglou & Papafragou, 2019b; Lockridge & Brennan, 2002; Brown-Schmidt & Heller, 2018), as well as prior interactions with their interlocutor (e.g., Brennan & Clark, 1996; Clark & Wilkes-Gibbs, 1986; Galati & Brennan, 2010, 2014; Heller, Gorman, & Tanenhaus, 2012; Metzing & Brennan, 2003; Horton, 2007; Horton & Gerrig, 2005; among others). Speakers also consider the attentional states of their addressee, recounting events in more detail when the addressee appears to be engaged as opposed to bored or inattentive (Grigoroglou & Papafragou, 2019a; Kuhlen & Brennan, 2010; Pasupathi, Stallworth, & Murdoch, 1998).

Most audience design studies have used the referential communication paradigm, wherein speakers name or describe an object to their addressee by referencing the object's size, color, shape, or pattern (Brennan & Clark, 1996; Clark & Wilkes-Gibbs, 1986; Horton, 2007; Horton & Gerrig, 2002, 2005; Horton & Keysar, 1996; Isaacs & Clark, 1987; Brown-Schmidt & Tanenhaus, 2006; Deutsch & Pechmann, 1982; Hanna, Tanenhaus, & Trueswell, 2003; Pechmann, 1989; Wardlow Lane & Ferreira, 2008; Wardlow Lane, Groisman, & Ferreira, 2006). However, a small but growing set of studies has begun adapting this paradigm to look at language production, and specifically at audience design, in the context of events (e.g., Brown & Dell, 1987; Grigoroglou & Papafragou, 2019a, 2019b; Lockridge & Brennan, 2002; Do, Papafragou, & Trueswell, 2020). In particular, because the representation of events incorporates not only different components but also structured relationship between those components, studying how people describe events at multiple levels of specificity can reveal how pragmatic factors associated with audience design combine with a speaker's conceptual representation of the event to determine what to say and what to omit.

In a recent study, Do et al. (2020) investigated the interaction between pragmatic and conceptual factors and the consequences of that interaction for what participants chose to mention when describing source-goal motion events, such as the event of a baby crawling from a crib to a toybox. At their core, these source-goal motion events involve a figure moving from one location—known as the *source* or starting point of the motion event—to the *goal*, or end point, of the event (Talmy, 1985).[1] There is ample evidence, though, that goals may be more prominent than sources in people's linguistic and nonlinguistic representations of these kinds of motion events. When talking about motion events, for instance, speakers in a number of typologically different languages, including Greek (Johanson, Selimis, & Papafragou, 2019), Arabic (Regier & Zheng, 2007), Japanese (Ihara & Fujita, 2000), and English (Do et al., 2020; Lakusta & Landau, 2005, 2012; Papafragou, 2010), overwhelmingly choose to mention the goal while omitting the source. This is true even in the case of deaf homesigners who have never had any exposure to a fully conventionalized language (Zheng & Goldin-Meadow, 2002). Furthermore, goals of motion are often accompanied by morphological markings and appear to come with a more diverse, semantically precise set of meanings than sources (Kabata, 2013; Narasimhan, Kopecka, Bowerman, Gullberg, & Majid, 2012).

At the same time, studies from the field of event cognition have shown that the bias toward the goal is present even when there is no communicative task involved. For instance,

preverbal infants attend to goals more than sources (Lakusta, Wagner, O'Hearn, & Landau, 2007; Tatone, Geraci, & Csibra, 2015; Woodward, 1998; inter alia). Likewise, both adults and children remember goals more accurately than sources, even when they are prevented from encoding these events linguistically (Lakusta & Landau, 2012; Papafragou, 2010, Regier, 1996; Regier & Zheng, 2007). And, studies have shown that the bias toward the goal is likely to stem from the way that these events are conceptually represented, rather than by simple effects of recency (Lakusta & Carey, 2015; Lakusta & Landau, 2012; Regier, 1996; Regier & Zheng, 2007).

Against this backdrop, Do et al. (2020) asked how the conceptual representation of a source-goal motion event (specifically, the conceptual prominence of goals compared to sources) interacts with pragmatic factors in communication while speakers are generating a message for language production. They found that speakers almost always mentioned the conceptually privileged goals, but tended to omit conceptually "peripheral" sources *unless* mentioning them was pragmatically informative for their addressee. Specifically, when describing source-goal events, such as the event of a butterfly$_{\text{FIGURE}}$ flying from a lamppost$_{\text{SOURCE}}$ to a chair$_{\text{GOAL}}$, to a physically copresent addressee who had already seen the figure of motion located at the source landmark (Common Ground condition), speakers largely omitted the source from their descriptions of the event; but when describing events to an addressee who had seen no part of the event at all (No Common Ground condition), speakers included the source in the vast majority of utterances.

Do et al. (2020) findings thus demonstrate the interplay of pragmatic and conceptual factors in message generation and raise a number of questions for further research. A first issue is whether the decision to mention sources in the No Common Ground should be attributed specifically to the conceptual status of the source itself, or to a more general tendency to mention more about the scene (e.g., because the addressee knew nothing about it). Since Do et al. (2020) manipulated the communicative status of the source via the addressee's visual access to the speaker's computer screen, addressees in the Common Ground condition were able to see everything—including the backdrop, figure in motion, source landmark, and goal landmark—except the motion to the goal. By contrast, addressees in the No Common Ground condition, who had no visual access to the speaker's screen, saw nothing at all. Even though this manipulation allowed Do et al. (2020) to successfully vary the pragmatic status of the source, an unintended consequence was that it also varied the pragmatic status of everything else in the scene. This raises the question of whether a pragmatic manipulation that more narrowly targets the source of the event would yield the same results.

A second, more general issue concerns the many factors—pragmatic and conceptual—that can affect a speaker's tendency to omit the more peripheral components of an event, such as the source of motion. In particular, the results of Do et al. (2020) pointed to a *pragmatic* asymmetry between sources and goals that ultimately affected speakers' tendency to omit sources: One reason why sources tend to be omitted is because sources, unlike goals, typically constituted already-known information for the listener. They also, however, pointed to the possibility of a second, *conceptual* asymmetry between sources and goals (see Do et al., 2020 for relevant discussion). Specifically, while prior work has visually packaged sources as part of the prior state of affairs, goals have usually been presented as a departure from that

initial state (e.g., Do et al., 2020; Lakusta & Landau, 2005, 2012; Papafragou, 2010; Regier & Zheng, 2007). In viewing source-goal motion events, then, it is possible that—in addition to construing sources as pragmatically informative to the *addressee*—speakers may also have construed sources as being conceptually less important than goals in their *own* representation of "what happened." Although Do et al. (2020) included a memory task to investigate the conceptual prominence of sources versus goals in these events, the task was administered to the same participants who had previously offered descriptions of motion events in one of the two Common Ground conditions. As a consequence, they were unable to fully separate pragmatic from conceptual factors contributing to the tendency to omit the source. In the current work, we compare the pattern of results obtained from speakers' descriptions of source-goal motion events with the results of a completely separate memory study.

Third, the role of animacy in either language or memory for source-goal motion events is not yet fully understood. In particular, a number of studies have consistently pointed to the salience of the goal in both language and memory for events involving *animate* figures in motion (e.g., Lakusta & Carey, 2015; Lakusta & Landau, 2005, 2012; Lakusta, Spinelli, & Garcia, 2017; Do et al., 2020; Lakusta et al., 2007; Papafragou, 2010; Regier, 1996; Regier & Zheng, 2007). In the case of inanimate figures, however (e.g., a ball rolling from a crib to a toybox), results are less clear. Although some studies have suggested that goals do not occupy a privileged status in nonlinguistic representations of inanimate events (Lakusta & Carey, 2015; Lakusta & Landau, 2012; Lakusta, Muentener, Petrillo, Mullanaphy, & Muniz, 2016), others have found a bias toward the goal in language but not nonlinguistic motion representation. Lakusta and Landau (2012), for instance, showed that when talking about either animate or inanimate motion events, adults and 3- to 4-year-old children showed a bias for the goal. Using the same stimuli, though, they found that participants only showed a bias for the goal in *memory* when the figure in motion was animate; when the figure was inanimate, accuracy for both source and goal changes was equally poor. Likewise, Lakusta and Carey (2015) used a change detection paradigm to investigate the goal bias among 12-month-olds and found that the bias toward the goal, which existed for animate figures in motion, disappeared when the figure in motion appeared to be inanimate. A similar pattern was obtained for preverbal 16-month-old infants in work by Lakusta et al. (2016). Given work by Do et al. (2020), one possible explanation for the asymmetry in language versus memory for inanimate motion events is that goal mention in the case of inanimate figures may be primarily driven by pragmatic factors in communication, while goal mention in the case of animates is driven by both pragmatic factors and factors tied to the conceptual status of the goal (Do et al., 2020; see also Lakusta & Landau, 2012). However, because prior work has not explicitly investigated the role of pragmatic factors in language for figure-inanimate events, this possibility has not yet been tested.

Finally, it is well-known that a speaker's ability to engage in audience design is limited not only by speaker-internal pressures (e.g., Horton & Keysar, 1996; Arnold & Griffin, 2007; Engelhardt, Bailey, & Ferreira, 2006; Horton, 2007; Horton & Gerrig, 2002; Pechmann, 1989; Wardlow Lane & Ferreira, 2008; Wardlow Lane et al., 2006), but also by speaker-external constraints, such as the availability of cues to the listener's knowledge state (e.g., Fussell & Krauss, 1989; Fussell, Kraut, & Siegel, 2000; Kraut, Miller, & Siegel, 1996; Krauss &

Weinheimer, 1964, 1966; Kuhlen & Brennan, 2010). Yet, work to date on the interaction between pragmatic and conceptual factors in language production has tended to focus on a single kind of pragmatic manipulation—the presence or absence of (visually established) common ground with a physically copresent addressee (e.g., Do et al., 2020; Grigoroglou & Papafragou, 2019a, 2019b; Lockridge & Brennan, 2002). In these contexts, the types of verbal and visual cues that speakers rely on to accurately assess and model the knowledge state of a physically copresent addressee are well-known and tend to be easily available: these include both verbal cues, such as back-channeling (*uh-huh, mhmm*), as well as nonverbal cues, such as the listener's eye-gaze, nodding, gesture, and facial expressions that can help to establish common ground (Brennan & Clark, 1996; Clark, 1996; Clark & Marshall, 1981; Giles & Powesland, 1975; Shockley, Richardson, & Dale, 2009, inter alia). An open question, then, is how mention of different event components, such as sources and goals, might be affected by the contexts where cues to the addressee's knowledge state may be missing, less detectable, or less reliable—for instance, when speakers and addressees do not share physical common ground (Fussell et al., 2000; Kraut et al., 1996; Kraut, Fussell, Brennan, & Siegel, 2002; Kraut, Fussell, & Siegel, 2003).

In the current studies, we adapt the paradigm of Do et al. (2020) to address the issues mentioned above and further understand how pragmatic and conceptual factors can jointly affect the linguistic and nonlinguistic encoding of motion events. We depart from that earlier study in four ways. First, we manipulate the pragmatic status of the source in a more targeted way (Experiment 1a). Second, we tease apart the pragmatic from the conceptual factors underlying the tendency to omit sources from mention by introducing a separate memory task (Experiment 1b) that allows us to take a precise look at what speakers consider salient about an event outside of the demands of language production. Third, we manipulate the animacy of the figure in motion (animate vs. inanimate) to understand the role that animacy can play in speakers' linguistic and conceptual representations of motion events (Experiments 1a and 1b). Finally, we ask how the interplay between conceptual and pragmatic factors is affected when cues to the knowledge state of the interlocutor are made less readily available (Experiment 2).

2. Experiment 1a: Language production

To understand the pragmatic factors that can affect speaker's descriptions of source-goal motion events, Experiment 1a asked participants to describe source-goal motion events to a physically copresent confederate addressee. This study built on the design of Do et al. (2020), but was modified so that the common ground status of *only* the source—in particular whether it was previously known versus unknown to both speaker and addressee—varied across conditions. Every other aspect of the motion event (e.g., backdrop and figure in motion) was already known to both the speaker and the addressee. If the findings from Do et al. (2020) are due to the pragmatic status of the *source* rather than a more general tendency to mention more about the event when the listener has seen less of it, we should replicate their results in this more direct manipulation: speakers should mention the source of the motion

more frequently when the source is previously unknown to the addressee (Source Unknown condition) than when it is already known (Source Known condition).

In addition, Experiment 1a investigated the role of pragmatic factors in the context of animate (e.g., a baby) versus inanimate (e.g., a ball) figures in motion. Unlike animate events, where the goal is privileged in both linguistic and nonlinguistic representations of motion events (Lakusta & Carey, 2015; Lakusta & Landau, 2012; Lakusta et al., 2016), prior studies have found a linguistic bias for the goal even when the goal is not considered conceptually prominent in inanimate events (Lakusta & Landau, 2012). To see whether the asymmetry between the goal bias in language versus in memory for inanimate events—in particular, whether the presence of the goal bias in language for inanimate events—can be attributed to pragmatic factors specific to language production, we extend the paradigm used in Do et al. (2020) to source-goal motion events with inanimate figures in motion.

2.1. Methods

2.1.1. Participants

Twenty-four native speakers of American English, who were at least 18 years old, were recruited from the University of Pennsylvania and given course credit for their participation. The number of participants was similar to the number of participants tested in each group of Do et al. (2020) study.

2.1.2. Materials and design

Experiment 1a was a 2×2 within-subjects design with factors Source Status and Animacy. The status of the source was varied across two conditions (Fig. 1). In the Source Known condition, the figure is visible from the frame of the video clip, prepositioned at the source of the motion event. In this condition, then, the source landmark (i.e., the object that the figure was moving from) was fully known to both speaker and addressee from the outset of each trial, and thus pragmatically uninformative to talk about. By contrast, figures in the Source Unknown condition are not visible at all from the opening scene; they emerged from behind or inside the source object only after the video started. In this latter condition, then, sources of motion were unknown until after the motion had started and consequently, informative to mention to an addressee who is not able to see the event unfold.

Animacy was also manipulated within subjects. Animated video clips featured either an animate (e.g., baby) or inanimate (e.g., ball) entity moving from a source (e.g., crib) to a goal (e.g., toybox), which was typically located on the other side of the screen (Fig. 1). The path of motion remained the same across Animate and Inanimate conditions; however, for practical reasons, the manner of motion varied for a small subset of items (e.g., crawling vs. rolling). Video clips were counterbalanced such that figures moved left-to-right in half the items, and right-to-left in the other half.

Twelve filler items were also included in Experiment 1a. They likewise depicted animate or inanimate figures in motion. However, these did not involve motion from a source to a goal (e.g., alarm clock shaking back and forth; flamingos dancing in place). In some cases,

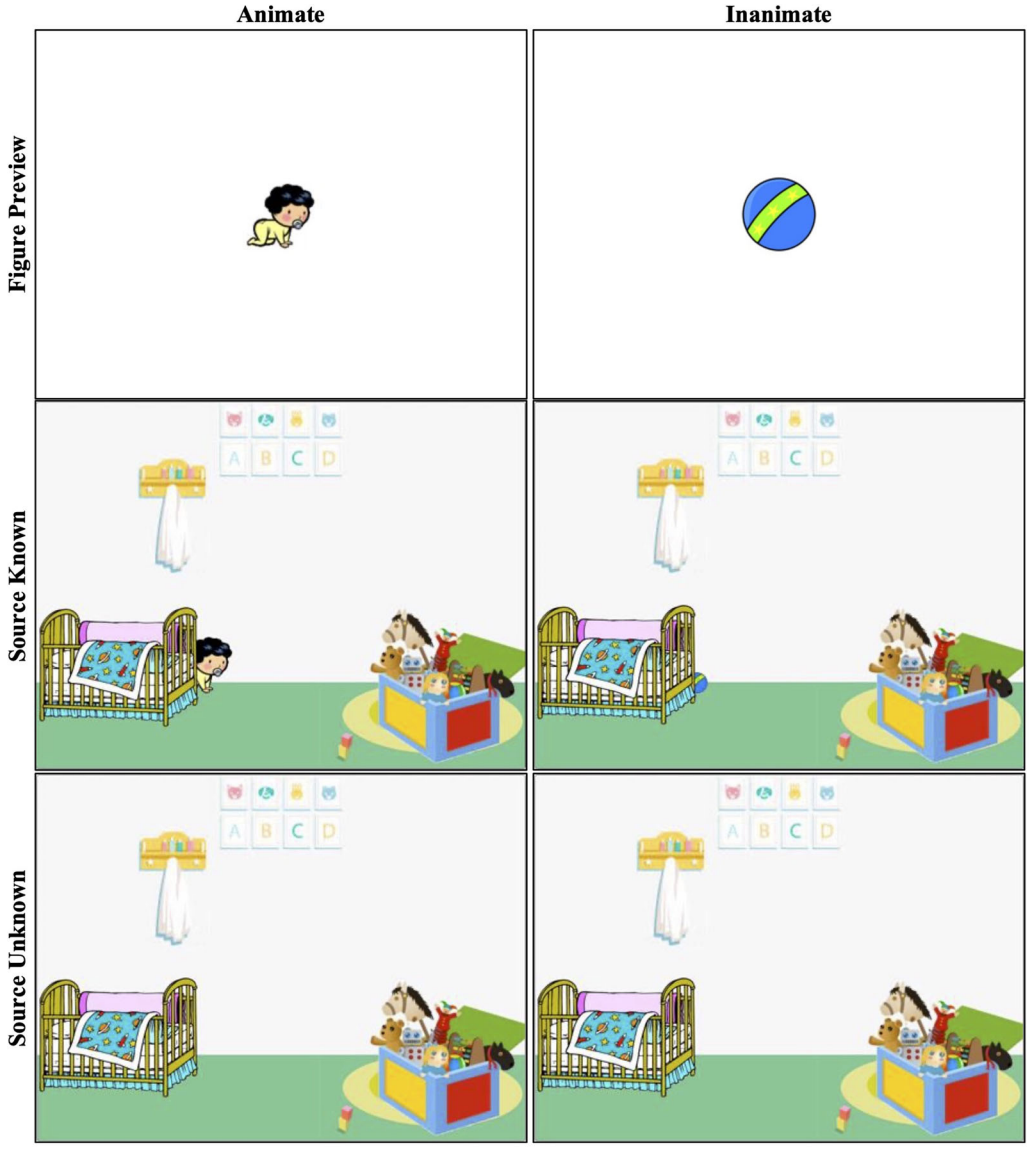

Fig 1. The Figure Preview screen introducing the animate (i.e., baby) and inanimate (i.e., ball) Figure in motion and opening scenes of the "baby / ball going from a crib to a toybox" event in each condition of Experiment 1a. In the Source Known conditions, the baby and ball are visible prior to the start of the motion to the toybox. In the Source Unknown conditions, neither the baby nor the ball are visible from the opening scene; they only become visible to participants once they emerge out from behind the crib after the motion starts. Participants see the Figure Preview screen before each trial in every condition.

the backdrop of the filler items was configured to match the target items—that is, they were designed so that participants might expect a source-goal motion event.

2.1.3. Procedure

Prior to the experiment, participants and confederates worked together to complete a *Tower of Hanoi* task that had no bearing on the research question at hand. The purpose of this introductory task was simply to allow confederates to establish some rapport with participants and demonstrate their engagement in the experiment. During the experiment, confederates maintained engagement with the speakers by maintaining eye-contact and through verbal back-channeling cues (e.g., *yup*, *mhmm*, *ok*).

Both participant and confederate addressee were seated side-by-side in front of a monitor that was visible to both speaker and addressee. Participants were told that they would watch some short video clips and describe what happened in those clips to their listener (the confederate addressee). Participants were told that their addressee would be answering simple questions based on their descriptions of the videos. Confederates were provided with a separate monitor on which to "answer" their questions; this screen was not visible to participants.

Each trial started with a separate, still image of the figure for the upcoming clip centrally displayed against a white background on the participant's monitor (Fig. 1). This was done to ensure that the figure would be known to both speaker and addressee, even in Source Unknown trials, where the figure was not immediately visible from the opening scene. Once participant and addressee had both seen the figure on the participant's screen, participants pressed the SPACE BAR to advance to the first frame of the video clip. Unlike in Do et al., both speaker and confederate addressee were then allowed to inspect this first frame of the motion clip. Once they had finished inspecting the opening scene, the confederate turned the participant's monitor away so that the screen was only visible to the participant. Participants pressed SPACE to start the clip. At the end of each video, they described what happened to their addressee. Once confederates finished answering the question that appeared on the second monitor, participants returned the shared screen to the starting position. The experiment lasted approximately 45 min.

2.2. Results

All stimuli, data, and analyses are available on the first author's OSF (https://osf.io/f5bm3/). Each utterance was coded for source (Y/N) and goal (Y/N) mentions, independently. For both sources and goal, the majority of mentions occurred in either a prepositional phrase (Source: *from the crib*; Goal: *to the toybox*; etc.) or a particle + preposition string (e.g., Source: *from behind the crib*; Goal: *over to the toybox*). However, there were a small percentage of cases in which sources were mentioned in the verb + NP structure (e.g., *left the lamppost*).

Analyses were performed using logistic mixed effect models (Baayen, Davidson, & Bates, 2008) with the lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2017) in R (R Core Team, 2017). The Mention Type (Source vs. Goal), Source Status (Source Known vs. Source Unknown), and Animacy Type (Animate vs. Inanimate) independent variables were all contrast coded using the –0.5 and 0.5 schema.

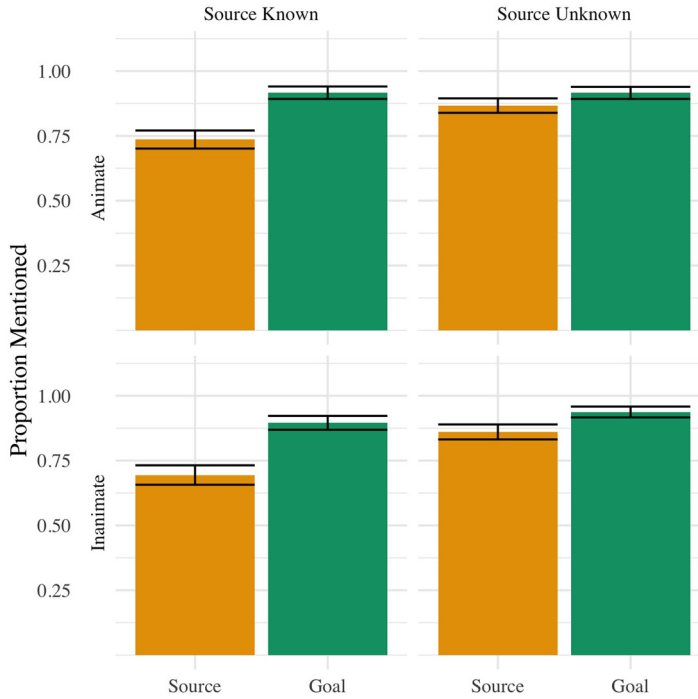

Fig 2. Proportion of source versus goal mentions by condition for Experiment 1a. Error bars indicate +/– 1 standard error.

Fixed effects were specified based on our questions of interest in our study. To see whether goals were mentioned significantly more than sources, we included Mention Type as a fixed effect. Because we were also interested in whether sources would be mentioned more frequently in the Source Unknown versus Source Known conditions, a Mention Type × Source Status interaction was also included. Finally, we also included a Mention Type × Animacy Type interaction to test whether goals were mentioned more frequently when the figure was animate versus inanimate. These same factors were included in the by-subject and by-item random effect structures. We started with the maximal random effect structure and simplified if the model failed to converge or if random effects did not contribute significantly (Bates, Kliegl, Vasishth, & Baayen, 2015; Matuscheck, Reinhold, Vasishth, Baayen, & Bates, 2017; chi-square > .05).[2]

The mean proportions of source versus goal mentions for each condition are shown in Fig. 2. Speakers were significantly less likely to mention sources (79%) than they were to mention goals across all conditions (92%); this was confirmed by the presence of a statistically significant main effect of Mention Type ($\beta = 1.42$, SE = .22, $|z| = 6.51$, $p < .001$). Importantly, as predicted, we also detected a significant interaction between Mention Type and Source Status such that speakers were more likely to mention sources in the Source Unknown condition (86%) than in the Source Known condition (72%; $\beta = 1.34$, SE = .26,

$|z| = 5.18$, $p < .001$). Unsurprisingly, there was no effect of Source Status on the mention of goals ($p > .3$)

Statistically, there were no significant interacts involving Animacy ($p > .5$), meaning that Figure animacy had no effect on either the rate of goal or source mentions in the animate versus inanimate conditions. Speakers mentioned the goals of source-goal motion events to the same extent (92%), whether the moving figure was animate or inanimate.

2.3. Discussion

Prior work by Do et al. (2020) has suggested that compared to conceptually prominent entities, such as goals, the decision to either mention or omit conceptually "peripheral" components, such as sources, is more likely to be driven by pragmatic factors in communication. Although that study found initial evidence to this effect, the design of that study left open the possibility that their results may have been driven by a general bias to mention more about events overall, rather than a source-specific effect. Using a more targeted design which varied only the pragmatic status of the source (Source Known vs. Source Unknown) conditions, we found that speakers were more likely to mention sources in the Source Unknown condition, where the source was previously unknown to the addressee, than they were in the Source Known condition, where the source was known information. In other words, speakers were more willing to mention sources when doing so provided new information to their addressee. Our results thus reaffirm the importance of pragmatic factors in driving decisions about conceptually "peripheral" components of an event (Grigoroglou & Papafragou, 2019a, 2019b; Lockridge & Brennan, 2002) and suggest that the original results observed by Do et al. (2020) were not simply a consequence of a general mention-more bias.

Turning to animacy, we took as our point of departure prior work showing that goals are conceptually privileged in the case of animate, but not inanimate events (e.g., Lakusta & Landau, 2012, Lakusta & Carey, 2015; Lakusta et al., 2016). Findings from those studies point to the possibility that the goal bias observed in language for *inanimate* source-goal motion events may be driven by pragmatics, alone, while goal mention in the case of animate events may be driven by both conceptual and pragmatic factors. We tested this possibility in Experiment 1a using carefully matched animate and inanimate figures in motion and controlling for pragmatic context. Somewhat surprisingly, though, we did not find any evidence of a difference in the rates of goal mention between the Animate versus the Inanimate conditions.

Having investigated the pragmatic factors that may have influenced both source and goal mentions in Experiment 1a, we turn to Experiment 1b, where we probe participants' memory for these same events to further investigate the conceptual factors that might also influence message generation for sources versus goals in source-goal motion events.

3. Experiment 1b: Memory

Although Experiment 1a confirmed that the decision to mention or omit sources of motion can be directly affected by pragmatic factors in communication, an open question is whether

there might also be other factors—such as the visual packaging of sources, as part of the initial state of affairs, versus goals, as part of a departure from that initial state—that can independently influence the tendency to omit sources from mention. To investigate this question, we asked a completely separate group of participants (cf. Do et al., 2020) to watch and remember the same events in Experiment 1a as best they could. Because no linguistic task was involved, we were able to effectively remove effects stemming from audience design and see whether the construal of sources versus goals can also affect a speaker's decision to mention or omit "peripheral" event components like sources.

If the tendency to omit sources is not purely pragmatic (i.e., if it can also be attributed to the construal of sources as part of the initial state of affairs), then we might expect differences between the Source Known and Source Unknown condition to obtain even when pragmatic factors associated with audience design are removed. This is because sources in the Source Unknown condition, but not the Source Known condition, can (like goals) be construed as departing from the initial state of affairs. If, on the other hand, the factors affecting speakers' decisions to omit the source are largely pragmatic, then the difference between the Source Unknown and Source Known conditions observed in Experiment 1a should disappear in a nonlinguistic memory task, where there is no interlocutor and no communicative act. In particular, speakers should remember sources just as well in the Source Unknown as in the Source Known conditions.

By the same logic, Experiment 1b also allows us to see whether the linguistic bias toward the goal in Inanimate Conditions really is, as we hypothesized, strictly a consequence of the pragmatic status of the goal (i.e., mentioning the goal is informative because it is not known to the addressee) or whether there may also be conceptual factors at play. This is especially important given the unexpectedly similar pattern of goal mentions in the Animate versus Inanimate conditions of Experiment 1a. If goal mentions in the *Inanimate* condition of Experiment 1a were driven purely by pragmatic factors related to informativity, then the bias toward the goal should disappear entirely when there is no addressee and no communicative intent. But, if there are separate conceptual factors underlying the goal bias in Inanimate conditions, they should be apparent here, once pragmatic factors associated with communication are removed.

3.1. Methods

3.1.1. Participants

We recruited 120 native speakers of American English from Amazon Mechanical Turk to participate in the study. They were paid $5 for 30 min of participation. We excluded participants ($n = 14$) who performed below chance in the filler items, which served as catch trials for this experiment. Data from 106 remaining participants were analyzed.

3.1.2. Materials and design

We used the same 24 target and 18 filler motion clips that had been used in the previous experiment. In addition, we also created a Source Change and a Goal Change variant of the original motion clip for each target item (Fig. 3). In these change videos, the original source and goal objects were swapped with a new item of the same kind (e.g., a different image of

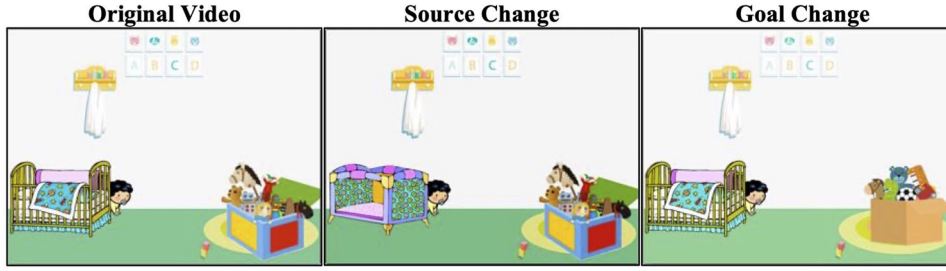

Fig 3. Sample images from the "baby crawled from the crib to the toybox" video (Experiment 1b) with Source and Goal changes. For demonstration purposes, changes in each condition are indicated here using a red arrow. Participants were not shown any arrow and did not receive any indication about.

a crib or a toybox). No new videos were created for filler items. These served as our baseline condition.

In addition to Source Status and Animacy, Change Type (Source Change vs. Goal Change) was also included as a within-subjects factor, yielding a 2×2×2 design. We rotated these conditions across eight lists using a standard Latin Squared design.

3.1.3. Procedure

Video clips were divided into eight blocks of four items each. Each of these blocks consisted of a study phase and a test phase. During the study phase, participants were instructed to remember each video as best they could. They were first shown a still image of the figure (the same image used in Experiment 1a), which remained on the screen for 3 s. After the still image disappeared, participants were shown the same motion clips used in Experiment 1a. Only after the video had finished playing were they allowed to press a button that advanced them to the next trial in the study phase. Participants were only allowed to view each still image and each video once.

During the test phase, participants were told that they would see a second set of four video clips. (Still images were not shown during the test phase.) Their task was to decide whether this second set of video clips was exactly the same as the ones they had seen during the study phrase. They were also told that some videos seen in the test phase would be the same and some videos would be different, but they were not given any information about the types of changes that they might encounter. Participants pressed "Yes" if they believed the videos were exactly the same and "No" otherwise. Test videos for target items were drawn from either the Source Change or Goal Change conditions; test videos for filler items (our No Change baseline) were exactly the same as the videos shown during the study phase.

3.2. Results

Results were coded for accuracy (Y/N) and analyzed using a logistic mixed effect model. Given our predictions, we were interested in a main effect of Change Type (Source Change vs. Goal Change), as well as its interactions with Source Status and Animacy. These factors were

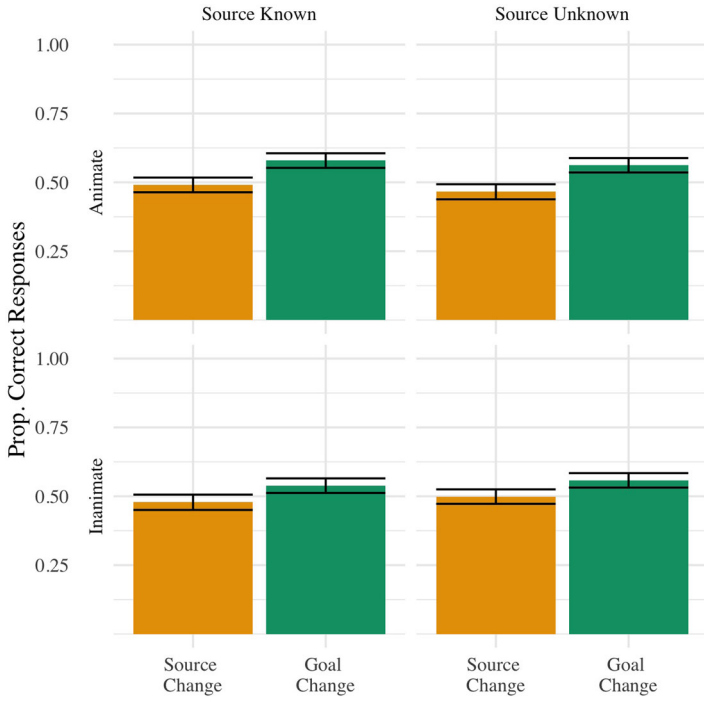

Fig 4. Mean proportion of correct responses for Source and Goal changes in each condition of Experiment 1b. Error bars indicate +/– 1 standard errors.

contrast coded and entered in the model as before; they were also included in the by-subject and by-item random effect terms. Model comparison was done as before.[3]

As shown in Fig. 4, participants were significantly worse at detecting changes to the source (48%) than they were to the goal (56%) in all conditions. This was confirmed statistically by a main effect of Change Type ($\beta = .39$, SE $= .16$, $|z| = 2.48$, $p < .05$). A separate analysis showed that participants did not differ significantly from chance in any Source Change conditions (p's $> .3$), but did perform significantly better than chance (p's $< .05$) in nearly all Goal Change conditions (in the Inanimate Source Known condition, the difference from chance was marginal, $p = .06$).

There was no evidence that participants were better at detecting changes to the source in the Source Unknown (49%) than in the Source Known (48%) condition as the interaction between Change Type and Source Status did not reach significance ($p > .9$). Unrelated to our predictions, participants also showed no difference in their ability to detect changes to the goal in the Source Known versus Source Unknown conditions (56% in both; $p > .8$).[4]

We also found participants performed comparably in the Animate and Inanimate conditions: Accuracy for sources was roughly 48% in both conditions, while accuracy for goal changes was roughly 56% in both Animate and Inanimate conditions. Accordingly, the Change Type × Animacy interactions failed to reach significance (p's $> .3$).

3.3. Discussion

Prior work by Do et al. (2020) showed that pragmatic and conceptual factors can affect the conceptually "peripheral" versus conceptually "core" components of an event differently. Given this, the aim of Experiment 1 was to take a closer look at the extent to which these two factors affect the decision to mention/omit conceptually "peripheral" sources (i.e., starting points) of motion versus conceptually "core" goals (i.e., endpoints) of motion. Although prior work (Do et al., 2020; Ihara & Fujita, 2000; Johanson et al., 2019; Lakusta & Landau, 2005, 2012; Papafragou, 2010; Regier & Zheng, 2007) has shown that speakers overwhelmingly choose to omit sources from mention, Experiment 1a showed that speakers were more willing to mention them when they are considered informative to the listener: Speakers mentioned sources more frequently in the Source Unknown conditions, where the source was initially unknown and uninferrable, than in the Source Known conditions, where the source was already known to the addressee.

In Experiment 1b, we used a nonlinguistic memory task to see whether there were also conceptual factors contributing to the tendency to omit sources from mention. Because the source of motion has typically been presented as part of the initial state of affairs (e.g., Do et al., 2020; Lakusta & Landau, 2005, 2012; Papafragou, 2010; Regier & Zheng, 2007), while goals of motion are visually packaged as part of the departure from that initial state of affairs, one possibility was that this construal of sources may additionally contribute to the tendency to omit sources from mention. The results of Experiment 1b provided evidence against this possibility: Accuracy in the memory task was the same regardless of whether sources were presented as part of the initial state of affairs (Source Known condition) or whether sources—like goals—were construed as a departure from that initial state (Source Unknown condition). In other words, while sources in the Source Unknown condition of Experiment 1a may have been privileged pragmatically (e.g., because they provided new information to the addressee), there was no evidence that speakers considered these same sources to be privileged independent from language production—that is, in speakers' own nonlinguistic representations of the event.

In addition to corroborating the results of Do et al. (2020), the results of Experiments 1a and 1b shed further light on not only the interaction between event structure and pragmatic factors during message generation, but also on the nature of the goal bias independent from language. In particular, the divergence in the pattern of results that we saw in language (Experiment 1a) versus in memory (Experiment 1b) suggests that the tendency to omit sources from mention was more likely due to pragmatic factors related to audience design than by the way that sources (but not goals) have typically been construed in motion events—a question left open by prior work. In spite of the important role that it played in language, the same Source Known versus Source Unknown manipulation had virtually no effect on the way that sources were represented outside of language (Experiment 1b). Once the task of language production and the pragmatic factors associated with audience design were removed, the difference between the Source Known and Source Unknown conditions that we had seen in language disappeared, confirming that the decision to mention sources more frequently in Source Unknown conditions of Experiment 1a was a pragmatically driven one.

At the same time, the results of the memory study suggest that the bias toward the goal is fundamentally about the relative prominence of the goal in people's representation of source-goal motion events and not about the way that these events have typically been presented in prior studies. More broadly, our results show that the prominence of the goal in both linguistic and nonlinguistic cognition can extend even to situations where the viewer may not already know where the figure in motion is starting from.

We also investigated the extent to which pragmatic and conceptual factors could affect the mention of conceptually "core" event elements, such as goals by manipulating the animacy of the figure in motion. Importantly, prior studies have noted an important asymmetry between animate and inanimate motion events (Lakusta & Landau, 2012). When the figure in motion is animate, the relationship between linguistic and nonlinguistic representations of source-goal motion events is clear: in both cases, there is a strong bias toward the goal. When the figure in motion is *inanimate*, however, the bias toward the goal—while still present in language—disappears when no linguistic task is involved (Lakusta & Carey, 2015; Lakusta & Landau, 2012). Given this, we wanted to test the possibility that the bias toward the goal in the case of inanimates was purely pragmatically driven.

Using tightly controlled stimuli that minimized the differences between our animate and inanimate motion events, we found that participants in Experiment 1a mentioned goals at similar rates in both Animate and Inanimate conditions. In Experiment 1b, however, we found somewhat surprisingly that participants were similarly accurate at detecting changes to the goal in both the Animate and Inanimate conditions. Contra prior studies (Lakusta & Carey, 2015; Lakusta & DiFabrizio, 2017; Lakusta & Landau, 2012), one surprising outcome from the present work is that goals of *inanimate* motion events can also be considered conceptually prominent even when pragmatic factors are removed. We discuss the role of animacy, as well as the possible reasons for the difference between our work and others', in more detail in the General Discussion.

4. Experiment 2: Language production under uncertainty

Experiment 2 asked how readily the findings from Experiment 1a might generalize to contexts where reliable cues to the knowledge state of the addressee may be less available to the speaker. In particular, if the tendency to mention sources more frequently in the Source Unknown than in the Source Known condition is fundamentally driven by pragmatic factors related to audience design, then curtailing the speaker's ability to effectively engage in audience design should have a direct effect on the difference between these two conditions.

To that end, the design and materials in Experiment 2 were exactly the same as in Experiment 1a, except that instead of describing events to a physically co-present addressee, participants described events to an addressee who appeared inside a picture-in-picture window at the top of their display. As in Experiment 1a, this addressee saw an initial presentation of the figure as well as the opening scene of the video clip on a shared display (i.e., the television screen) but did not see the rest of the video unfold. Because the conceptual structure of the event as well as the knowledge state of the addressee were exactly the same in Experiments 1a and 2, the only difference was *how* speakers were able to ascertain the knowledge state

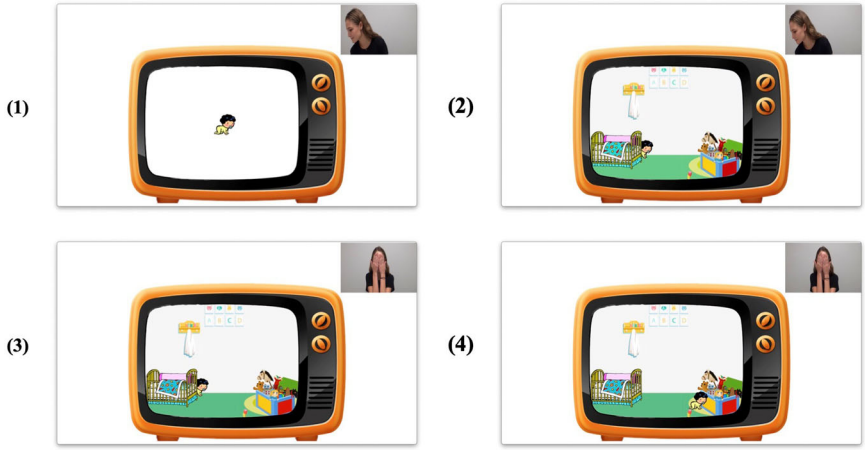

Fig 5. Sample sequence for the "baby ran from the crib to the toybox" item (Experiment 2). At the first chime, participants saw the interlocutor look at the still image of the figure in motion (Panel 1). After a second chime, participants saw the interlocutor inspect the first scene of the motion clip (Panel 2). Once the interlocutor was done inspecting the opening scene of the motion event, she placed her hands over her eyes (recall that the addressee was not allowed to see the motion event) and indicated that she was ready for the trial to begin (Panel 3). At this point, participants pressed the SPACE BAR to start the motion clip. Once the motion clip had ended, participants were instructed to describe what happened in the motion event to the interlocutor, whose hands remained over her eyes until the end of the trial (Panel 4).

of their addressee. In Experiment 1a, speakers were able to model the knowledge state of the addressee via turn-taking with the shared monitor and through other verbal and nonverbal cues. In Experiment 2, however, speakers were only able to establish the knowledge state of the interlocutor by keeping track of what she could see from her picture-in-picture window. Cues to the knowledge state of the interlocutor were thus less transparent and less reliable in Experiment 2 than in Experiment 1a.

4.1. Methods

4.1.1. Participants

We recruited 26 people from Amazon Mechanical Turk to participate in Experiment 2. We analyzed data from 23 of those recruits (three were excluded because they were not native speakers of American English or because they failed to complete the study). All were paid $5.50 for roughly 30 min of participation.

4.1.2. Materials and design

Unlike in Experiment 1a, participants spoke to an addressee who appeared inside of a picture-in-picture window at the top of the display. Both participants and addressees watched the animated motion clips unfold on a television screen that appeared on the participant's display (Fig. 5).[5] The target and filler motion events shown inside the television screen were exactly the same as the ones used in Experiment 1a.

4.1.3. Procedure

The experiment was deployed through PCIbex (Zehr & Schwarz, 2018). The general structure of Experiment 2 was the same as in Experiment 1a; however, some adjustments had to be made in order to accommodate the picture-in-picture addressee (Fig. 5). For instance, although each trial began with the still image of the figure that would appear in the upcoming video, the appearance of this figure in Experiment 2 was accompanied by a chime that signaled the appearance of the figure on the shared display (i.e., the television screen). Upon hearing this chime, the interlocutor "peered out" of her picture-in-picture window to inspect the image on the television set. When she had finished looking at the image, she sat back in her window. Then, the first frame of the video clip appeared, accompanied by a second chime, and the interlocutor "peered out" of her window again. After she had finished looking around and inspecting the first frame of the video clip, she sat back in her window, placed her hands over her eyes, and verbally indicated that participants could begin playing the clip (e.g., *ok, I'm ready, yup*). Participants then pressed a button to begin the animated clip. The interlocutor kept her hands over her eyes until the end of the video clip. Once the clip ended, participants described what they had seen in the clip and then pressed a button to advance to the next trial.

4.2. Results

Results were contrast coded and analyzed using logistic mixed effect regression models as in Experiment 1a.[6]

As can be seen in Fig. 6, sources (82%) were mentioned less often than goals (91%) in all conditions, resulting in a statistically significant main effect of Mention Type that was not modulated by any other factors ($\beta = 1.11$, SE $= .55$, $|z| = 2.03$, $p < .05$). Contrary to what we predicted and unlike what was observed in Experiment 1a, however, there were no main or interaction effects involving Source Status, suggesting that there were no statistically significant differences in the rate of source ($p > .9$) mentions in the Source Known (79%) than in the Source Unknown (85%) conditions. Though unrelated to our predictions, there were also no significant effects involving goal mentions in Source Known versus Source Unknown conditions ($p > .05$). As in Experiment 1a, we also saw no significant main or interaction effects involving Animacy. In particular, the rate of goal mentions did not differ significantly in the Animate (92%) versus Inanimate (90%) conditions ($p > .3$). There were also no significant differences in the rate of source mentions across animacy conditions ($p > .3$).

4.2.1. Comparison of Experiment 1a and Experiment 2

We compared the results of Experiment 1a and Experiment 2 directly using the same logistic mixed effect regression and model comparison methods as before. Experiment (in person addressee vs. picture-in-picture addressee) was added as a fixed effect factor and as a factor in the by-item random effects.[7]

As expected, sources were mentioned significantly less frequently than goals across all conditions of both experiments: We detected a significant main effect of Mention Type ($\beta = 2.16$,

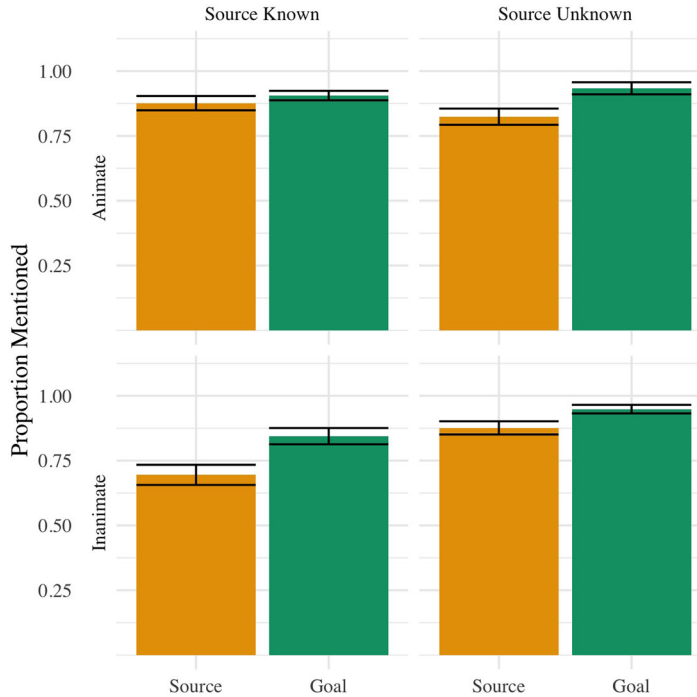

Fig 6. Proportion of Source and Goal mentions in each condition of Experiment 2. Error bars show +/– 1 standard error.

SE = .49, |z| = 4.39, $p < .001$). We also detected a significant Mention Type × Source Status interaction such that sources were mentioned significantly more in the Source Unknown than the Source Known condition ($\beta = .95$, SE = .23, |z| = 4.13, $p < .001$). However, this was modulated by a significant Mention Type × Source Status × Experiment interaction ($\beta =$ –.91, SE = .45, |z| = 2.02, $p < .05$), suggesting that the Source Status effect was dependent on Experiment: Sources were mentioned significantly more in the Source Unknown than the Source Known condition in Experiment 1a but, in line with the above results, no such difference emerged in Experiment 2. There were no other significant effects involving Source Status.

Turning our attention to effects involving Animacy, we did not find differences in the rate of goal mentions in either the Animate or Inanimate conditions across experiments (p's > .1).

4.3. Discussion

In Experiment 2, we reasoned that if decisions about whether to mention or omit "peripheral" components of an event—such as sources—are largely driven by pragmatic pressures related to audience design, then source mentions should be affected by speakers' ability to actually determine the knowledge state of their addressee. We, therefore, made the cues to

the addressee's knowledge state less transparent by asking participants to describe the same events in Experiment 1a to an addressee visible through a picture-in-picture window on the speaker's screen, rather than to a physically copresent addressee.

Although speakers described the exact same motion events and addressees knew the same amount of information about the event in Experiment 2 as in Experiment 1a, we found a different pattern of results across these two experiments—at least when it came to mentioning the source. Specifically, while speakers in Experiment 1a mentioned the source roughly 15% more often in the Source Unknown than in the Source Known condition, speakers in Experiment 2 mentioned sources to the same extent regardless of whether sources were already known to the addressee (Source Known condition) or not (Source Unknown condition). In other words, consistent with what is expected under a pragmatic account of source mention, we found that the decision to mention or omit sources did, indeed, depend on the speakers' ability to adapt to the pragmatic needs of their interlocutors. In conjunction with the results of Experiment 1, these results provide additional support for the possibility that the decision to mention or omit less prominent, "peripheral" event components is primarily driven by pragmatic factors in communication.

Although unpredicted, an additional pattern worth noting is the tendency for speakers in Experiment 2 to be *overinformative* in their utterances. In particular, even in the Source Known conditions, where the source was already known and pragmatically uninformative to their addressee, speakers in Experiment 2 mentioned sources at rates comparable to what was seen in the Source *Unknown* conditions of Experiment 2 and Experiment 1a. Given that the main aim of Experiment 2 was not to investigate the factors that give way to overinformativity, further research is certainly required. Nevertheless, the pattern of overinformativity seen in Experiment 2 hints at the role that uncertainty about an interlocutor's knowledge state may play in determining speakers' choices about under versus overinformativity. One possibility, for instance, is that the pattern of overinformativity observed in Experiment 2 reflects the type of adaptation used by speakers who are less able to reliably track the knowledge state of their interlocutor. In Experiment 1a, the in-person confederate played an active, collaborative role in establishing whether the source was known versus unknown: The confederate inspected the image on the participant's screen and in fact, it was the confederate who turned the participant's screen away after inspecting the display. These interactions between speaker and interlocutor in Experiment 1a meant that cues to the interlocutor's informational need were clear and easy to monitor from trial to trial. In Experiment 1a, then, speakers may have been more willing to omit the source because they could assess, with a reasonable degree of certainty, the informational needs of the listener.

In Experiment 2, by contrast, cues—both verbal and nonverbal—to the informational needs of the listener were both less collaborative and less transparent: Although the interlocutor behaved as though she was looking around at the images, differences in the speaker's versus the interlocutor's perspective may have made it difficult to readily keep track of what the interlocutor had really seen of the event, especially since speakers had to divide their attention between the event being described and the addressee on the display (see Arnold & Griffin, 2007 for related discussion). Although the interlocutor always appeared to be looking at the television screen and always verbally indicated when she had finished inspecting the scenes,

speakers may still have been unclear about precisely how much of the scene their interlocutor really saw. These less transparent cues, then, may have reduced speakers' certainty about the knowledge state of their addressee and in turn, motivated the pattern of overinformativity that we saw in the Source Known condition of Experiment 2.

This account of our results is consistent with the type of overinformativity found in other studies in which speakers were asked to communicate with an unidentified interlocutor (Fussell & Krauss, 1989; Arts, Maes, Noordman, & Jansen, 2011). In both cases—that is, in Experiment 2 and in the context of an unidentified interlocutor—it may have been uncertainty about the addressee's knowledge state that motivated speakers to take a "better safe than sorry" approach and say more about an intended referent to avoid miscommunication. If so, then this would also lend support to recent studies showing that speakers can, in some circumstances, be "strategic" in their choice to be overinformative (Arts, Maes, Noordman, & Jansen, 2011; Paraboni & van Deemter, 2014; Rubio-Fernandez, 2016, 2019; see also related work by Yoon & Brown-Schmidt, 2018, 2019). We consider the issue of overinformativity further in the General Discussion.

Unlike sources, goals did not appear to be affected by changes in pragmatic context across experiments. In fact, the pattern of goal mentions in Experiment 2 replicated the results seen in Experiment 1a. Again, we found that rates of goal mention were the same across Animate and Inanimate conditions, suggesting that speakers in Experiment 2, as in Experiment 1a, treated animate and inanimate source-goal motion events in largely the same way (cf. Lakusta & Carey, 2015; Lakusta & Landau, 2012).

5. General discussion

When generating a message, speakers have to decide what to mention or not mention about an event. This choice is influenced by conceptual factors: Speakers are more likely to mention the conceptually more prominent "core" elements of an event, such as the goal (i.e., endpoint) of a motion event, compared to conceptually less prominent "peripheral" elements of an event, such as the source (i.e., starting point) of a motion event (Lakusta & Landau, 2005, 2012). And, it is also influenced by pragmatic factors related to audience design: Speakers take into account the informational needs and the knowledge state of their addressees by mentioning information that may be unknown or uninferable to addressees, and omitting information that they believe is already known to their interlocutor (e.g., Brown & Dell, 1987; Clark, 1992, 1996; Clark & Wilkes-Gibbs, 1986; Lockridge & Brennan, 2002).

Recently, a growing body of research (e.g., Do et al., 2020; Grigoroglou & Papafragou, 2019a, 2019b) has begun asking how the different components of an event—"core" versus "peripheral"—can be differentially impacted by conceptual factors versus pragmatic factors related to audience design. The work presented here extends that line of research by investigating the extent to which pragmatic versus conceptual factors can affect speakers' decisions to omit or mention sources versus goals when describing animate and inanimate motion events.

5.1. Pragmatic and conceptual factors in message generation

Using a test that more stringently targeted sources as parts of the conversational common ground between speaker and addressee (cf. Do et al., 2020), we show that the decision to mention conceptually "peripheral" event components, such as sources, appears to be primarily contingent on the pragmatic status of the source: In Experiment 1a, sources were mentioned more often when they were unknown to an interlocutor (Source Unknown condition); but differences between the Source Known and Source Unknown conditions disappeared when was no interlocutor present—in other words, when pragmatic factors were eliminated. Further evidence for the role of pragmatic factors in driving decisions about source mention came from Experiment 2, which showed that changes to pragmatic context could then motivate speakers to mention sources of motion even more often than needed. Taken together with work by Do et al. (2020) and work by Grigoroglou and Papafragou (2019a, 2019b) showing similar effects with instruments, the results of the current study highlight the important distinction between the ways that different types of factors can influence the linguistic mention of the different parts of an event—in particular, the special role that pragmatic factors related to audience design can play in linguistically promoting otherwise "peripheral" event roles, such as sources.

The same does not appear to be true, though, for the conceptually prominent, "core" elements of an event, such as the goal. Unlike with sources, the decision to mention goals did not appear to be dependent on pragmatic factors in communication: The bias toward the goal was apparent for Animate and Inanimate events, even in a completely nonlinguistic task, where pragmatic factors were removed. Moreover, the rate of goal mentions in both the Animate and Inanimate conditions (unlike source mentions) did not correspond to changes in the pragmatic context across Experiment 1a and Experiment 2. These results are consistent with prior studies showing that speakers tend to mention goals not necessarily for pragmatic reasons (Do et al., 2020) but because goals are especially prominent in people's representation of motion events, whether they intend to talk about those events or not (Lakusta & Landau, 2012). Nevertheless, the extent to which these factors can ultimately affect message generation for conceptually "core" elements of an event remains an important issue for future investigation. While goals have thus far appeared to be resilient to pragmatic factors associated with audience design, it is unclear whether pragmatic factors *alone* cause speakers to omit the goal from their utterances, or whether omission of conceptually privileged event components necessarily depends on conceptual factors.

5.2. Effects of (In)animacy on source-goal representations

Given prior work pointing toward a difference in the way that participants conceptually represent figure-animate and figure-inanimate motion events (Lakusta & Carey, 2015; Lakusta & Landau, 2012), an interesting puzzle raised by our work is why participants in our nonlinguistic memory study, unlike in other studies, appeared to encode animate and inanimate motion events in a similar way.

On this question, one possibility is that the differing results may stem from differences in experiment design. Specifically, whereas Lakusta and Landau (2012) compared different

kinds of animate (e.g., a person walking from a chair to a box) versus inanimate (e.g., a piece of paper being blown off the arm of a sofa onto the seat cushion), we aimed to create as direct a comparison as possible between our animate versus inanimate events. This meant using cartoon clips that minimized differences in paths and manners of motion between animate and inanimate figures and manipulating animacy within-subjects. And while our design did undoubtedly provide tighter control of the stimuli, it is possible that the distinction between animate versus inanimate figures became heavily attenuated and the surface characteristics of animate behavior became harder to distinguish from those of inanimate movement.

Some support for this possibility comes from other studies, which have likewise shown that goals can be conceptually prominent in the case of inanimate figures when those inanimates are imbued with animate, agent-like properties (Csibra, Bíró, Koós, & Gergely, 2003; Luo & Baillargeon, 2005; Papafragou, 2010). If this account is on the right track, then a natural question is how central perceptual cues, such as the path or manner in which a figure moves, are to the perception (or construal) of animacy (see van Buren, Gao, & Scholl, 2017 for related discussion).

A second aspect of our experiment design that may also have contributed to the indistinction between our animate versus inanimate figures in motion may have been the within-subjects manipulation employed in our study, but not by Lakusta and colleagues. Videos in our study, for instance, largely depicted both animate and inanimate figures moving across the center of the screen while the rest of the video remained unchanged. Seeing animate and inanimate events in this way may have encouraged speakers to use higher-level inferencing to reason about which aspects of the motion clip they were "supposed" to talk about (Rissman, Woodward, & Goldin-Meadow, 2018): Why, for instance, would a ball be shown as the visual focus of an event if it should not be considered important to talk about? By presenting participants with one highly similar event after another, then, we may have further encouraged participants to treat animate versus inanimate motion events in the same way.

Nevertheless, our results do shed light on the nature of the homology between speakers' linguistic and nonlinguistic representation of events. In particular, Lakusta and Landau (2012) showed that one reason why the homology between language and memory for source-goal motion events could be disrupted in the context of inanimates was because of differences in the way that animate and inanimate source-goal motion events are conceptually represented. Consistent with their suggestion, our studies showed that when animate and inanimate events *are* treated in the same way conceptually, that very tight correspondence between people's linguistic and nonlinguistic representations can re-emerge.

5.3. Audience design under uncertainty

A number of studies have already pointed to some of the ways that speaker-internal pressures may lead to overinformativity: Speakers are known to be overinformative (i) when attentional demands are high (Arnold & Griffin, 2007); (ii) when they are unable to fully suppress irrelevant but highly salient information (Wardlow Lane & Ferreira, 2008; Wardlow Lane et al., 2006); or (iii) when they fail to differentiate their own perspective from that of their interlocutor (Engelhardt et al., 2006). Unlike in those studies, though, speakers in our study were not directly limited by their own *capacity* to attend to their interlocutor's knowledge

state. Rather, it was the (un)reliability of the pragmatic cues to the listener's knowledge state that led to the production of overinformative utterances in the Source Known condition of Experiment 2. Though unexpected, the results of Experiment 2 thus show that the reliability of a pragmatic cue may be just as important as the speaker-internal factors motivating overinformativity.

At the same time, the contrast between the results of Experiment 1a and Experiment 2 also highlights important theoretical questions about the production of over versus underinformative utterances. For instance, while it is well known that the process of audience design relies on adequate feedback by the interlocutor (Brennan & Clark, 1996; Clark, 1996; Clark & Marshall, 1981; Krauss & Fussell, 1996; Krauss & Weinheimer, 1966), an open question is how that lack of feedback triggers the decision to be over versus underinformative. Some studies (Kuhlen & Brennan, 2010; Pasupathi et al., 1998) have shown that when lack of feedback is tied to inattentiveness or to the disingenuous needs of a listener, this may cause people to say less about an event. On the other hand, though, the results from Experiment 2 and other studies (Fussell & Krauss, 1989; Krauss & Fussell, 1996; Krauss & Weinheimer, 1964, 1966) demonstrate that speakers may also produce lengthier utterances in the absence of real-time feedback from their interlocutors. Given that speakers do appear to adapt differently to similar pragmatic cues, an important question for future work is not simply the types of cues (un)available to the speaker, but also *how* those cues are interpreted by the speaker during language production (Kuhlen & Brennan, 2010).

Finally, while the aim of this work was not to investigate the nature of real-time, virtual communication (e.g., through Zoom, Skype, or Facetime), our findings do have implications for audience design in the virtual medium. Our results suggest, for instance, that speakers who are communicating to a virtual addressee may be similarly unsure about the knowledge state of their addressee. In addition to uncertainty about what an addressee can or cannot see, gestures made by an interlocutor may be obscured or occur off-screen and verbal back-channels can often be reduced or delayed (e.g., Kraut, Fussell, & Siegel, 2003; Kraut, Fussell, Brennan, & Siegel, 2002; Kraut, Miller, & Siegel, 1996; Fussell, Kraut, & Siegel, 2000). We expect, then, that speakers in these contexts may also be particularly susceptible to a specific kind of audience design failure—namely, overinformativity. At the same time, our results point to one reason why virtual interactions seem more tiresome and effortful than interactions involving physically copresent interlocutors. In particular, tracking the moment-by-moment knowledge state and making listener-specific adjustments when communicating with a physically copresent, fully cooperative interlocutor is already cognitively resource-intensive (e.g., Brown-Schmidt & Duff, 2016; Horton & Gerrig, 2005; Nadig & Sedivy, 2002). When those cues to the listener's knowledge state are not as readily available, the task of engaging in audience design may become even more difficult.

This may be of import methodologically in areas of research where the use of virtual interlocutors may already be (or may become) standard. In studies of pragmatics with young children, for instance, it may also be important to take into account the added difficulty and increased effort required by speakers to monitor, reliably assess, and adapt to the needs of a picture-in-picture and/or virtual interlocutor (see also Grigoroglou & Papafragou, 2019a for a similar discussion).

6. Conclusion

We used source-goal motion events to investigate the extent to which pragmatic versus conceptual factors can inform the process of message generation. We highlight the special role of pragmatic factors related to audience design in governing speakers' decisions to be under or over informative when talking about conceptually "peripheral" event components, such as sources. To investigate the way pragmatic and conceptual factors affect mention of conceptually "core" elements, such as the goal of a motion event, we manipulated animacy of the figure in motion. Contra prior studies (Laksuta & Landau, 2012; Lakusta & Carey, 2015), participants in our studies treated both animate and inanimate source-goal motion events in the same way, both linguistically and in memory. We discuss the implications of our work for message generation across different communicative contexts and for future work on the topic of audience design.

Notes

1 Source/goal events can also include a transfer of possession (Lakusta & Landau, 2005; Tatone, Geraci, & Csibra, 2015), an act of attachment/detachment (Lakusta & Landau, 2005; Narasimhan, Kopecka, Bowerman, Gullberg, & Majid, 2012), and even events in which the figure undergoes a change of state. We will not be discussing these cases here.

2 The final model for Experiment 1a is listed here: glmer(data=AnimXPred_InLab_Long, MentionCode ∼ MentionType + MentionType:InformativityType + MentionType: AnimacyType + (1|SubjID) + (1|ItemID), family=binomial, control=glmerControl(optimizer = "bobyqa", optCtrl = list(maxfun=2e5))). Note that neither the Animacy × Source Status nor the Animacy × Source Status × Mention Type interactions were included as fixed or random effects. We chose to exclude these interactions because—given that Animacy is only expected to affect goal mention, while Source Status is only expected to affect source mention—we are not aware of any theory of event structure, pragmatics, or language production that would predict an interaction between Animacy and Source Status. As such, we are not sure that these could be meaningfully interpreted. We, therefore, opted to exclude interactions between Animacy and Source Status in the interest of maximizing statistical power. Likewise, main effects involving Source Status or Animacy, which do not take into account whether the entity being mentioned was the source or goal, respectively, were also considered similarly uninformative, so both were also excluded as predictors.

3 The final model for Experiment 1b is provided here: glmer(data=MemoryTarg, ResponseCorrect ∼ ChangeType + ChangeType:InformativityType + ChangeType:Animacy Type + (1|SubjID) + (1+ ChangeType + ChangeType:InformativityType|ItemID), family=binomial, control=glmerControl(optimizer = "bobyqa", optCtrl = list(maxfun=2e5))).

4 One possibility was that differences between Source Known and Source Unknown conditions were obscured because of the difficulty of the task. To investigate this, we analyzed results from just the top 50th percentile of performers in the Goal Change condition —

all of whom scored above chance on both Source Change (p's $< .02$) and Goal Change conditions (p's $< .001$). Even among "high-performing" participants, we failed to find any effect of informativity or animacy (p's $> .3$). Thus, even when sources do tend to be encoded accurately, there is no statistically reliable difference between the Source Known versus Source Unknown conditions for either animate or inanimate events.

5 A previous version of this experiment did not show the motion events inside a television set. In that version, the interlocutor looked directly at the motion clip as it was shown to participants. This design led to concerns that the interlocutor was able to see more of the scene than participants could (e.g., that the interlocutor could see the baby behind the crib, even when the baby was not visible to participants). The "television screen" version described above rectified this issue. Both experiments yielded the same pattern of results but we report the results from the better-designed "television screen" version here. Results from the initial study (without the television) are given in Appendix A.

6 The final model for Experiment 2 is given here: glmer(data=AnimXPred_TV_Long, MentionCode ~ MentionType + MentionType:InformativityType + MentionType: AnimacyType +(1+MentionType|SubjID) + (1|ItemID), family=binomial, control=glmerControl(optimizer = "bobyqa", optCtrl = list(maxfun=2e5))).

7 The final model comparing the results from Experiment 1a and Experiment 2 is provided here: glmer(data=AllLang, MentionCode ~ MentionType + MentionType: InformativityType + MentionType:AnimacyType+ MentionType:Experiment + MentionType:InformativityType:Experiment + MentionType:AnimacyType:Experiment + (1+MentionType + MentionType:AnimacyType|SubjID) + (1+MentionType |ItemID), family=binomial, control=glmerControl(optimizer = "bobyqa", optCtrl = list(maxfun=2e5))).

Acknowledgments

A.P. acknowledges the support from NSF grant 1632849. We also thank members of the Language Learning Lab for their help in creating the stimuli and Jeremy Zehr, in particular, for his help in the implementation of Experiment 2. Finally, we are grateful to our reviewers for their thoughtful and helpful comments on this work.

Conflicts of interest

The authors have no conflicts to disclose.

Open Research Badges

This article has earned Open Data and Open Materials badges. Data are available at https://osf.io/f5bm3/?view_only=a9c506bd3294422b9bbf096cb64283ef and materials are available at https://osf.io/f5bm3/?view_only=a9c506bd3294422b9bbf096cb64283ef.

REFERENCES

Arnold, J. E., & Griffin, Z. M. (2007). The effect of additional characters on choice of referring expressions: Everyone competes. *Journal of Memory and Language*, *56*, 521–536.

Arts, A., Maes, A., Noordman, L., & Jansen, C. (2011). Overspecification in written instruction. *Linguistics*, *49*, 555–574. https://doi.org/10.1515/LING.2011.017

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Bates, D., Kliegl, R., Vasishth, S., & Baayen, R. H. (2015). Parsimonious mixed models. Available from arXiv:1506.04967 (stat.ME).

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *22*, 482–493.

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

Brown-Schimidt, S., & Tanenhaus, M. K. (2006). Watching the eyes when talking about size: An investigation of message formulation and utterance planning. *Journal of Memory and Language*, *54*, 592–609. https://doi.org/10.1016/j.jml.2005.12.008

Brown-Schmidt, S., & Duff, M. C. (2016). Memory and common ground processes in language use. *Topics in Cognitive Science*, *8*, 722–736.

Brown-Schmidt, S., & Heller, D. (2018). Perspective-taking during conversation. In S. Rueschemeyer & M. G. Gaskell (Eds.), *Oxford handbook of psycholinguistics* (pp. 551–574). New York: Oxford University Press.

Clark, H. H. (1992). *Arenas of language use*. Chicago, IL: University of Chicago Press.

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

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

Csibra, G., Bíró, S., Koós, O., & Gergely, G. (2003). One-year-old infants use teleological representations of actions productively. *Cognitive Science*, *27*, 111–133.

Deutsch, W., & Pechmann, T. (1982). Social interaction and the development of definite descriptions. *Cognition*, *11*, 159–184.

Do, M. L., Papafragou, A., & Trueswell, J. (2020). Cognitive and pragmatic factors in language production: Evidence from source-goal motion events. *Cognition*, *205*, e104447.

Engelhardt, P. E., Bailey, K. G. D., & Ferreira, F. (2006). Do speakers and listeners observe the Gricean Maxim of Quantity? *Journal of Memory and Language*, *54*, 554–573.

Fussell, S. R., & Krauss, R. M. (1989). Understanding friends and strangers: The effects of audience design on message comprehension. *European Journal of Social Psychology*, *19*, 509–525. https://doi.org/10.1002/ejsp.2420190603

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

Fussell, S. R., Kraut, R. E., & Siegel, J. (2000). Coordination of communication: Effects of shared visual context on collaborative work. Proceedings of the CSCW 2000 Conference on Computer-Support Cooperative Work (pp. 21–30). New York: ACM.

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

Galati, A., & Brennan, S. E. (2014). Speakers adapt gestures to addressees' knowledge: Implications for models of co-speech gesture. *Language, Cognition and Neuroscience*, *29*, 435–451.

Giles, H., & Powesland, P. F. (1975). *Speech styles and social evaluation*. London: Academic Press.

Grigoroglou, M., & Papafragou, A. (2019a). Interactive contexts increase informativeness in children's referential communication. *Developmental Psychology*, *55*, 951–966.

Grigoroglou, M., & Papafragou, A. (2019b). Children's (and adults') production adjustments to generic and particular listener needs. *Cognitive Science*, *43*, e12790. https://doi.org/10.1111/cogs.12790

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.

Heller, D., Gorman, K., & Tanenhaus, M. K. (2012). To name or to describe: Shared knowledge affects referential form. *Topics in Cognitive Science*, *4*, 290–305.

Horton, W. S. (2007). The influence of partner-specific memory associations on language production: Evidence from picture naming. *Language and Cognitive Processes*, *227*, 1114–1139.

Horton, W. S., & Gerrig, R. J. (2002). Speakers' 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. (2005). 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.

Ihara, H., & Fujita, I. (2000). A cognitive approach to errors in case marking in Japanese agrammatism: The priority of the goal -ni over the source -kara. In A. Foolen & F. Van der Leek (Eds.), *Constructions in cognitive linguistics: Selected papers from the fifth international cognitive linguistics conference* (pp. 123–140). Amsterdam: John Benjamins.

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

Johanson, M., Selimis, S., & Papafragou, A. (2019). The source-goal asymmetry in spatial language: Language-general vs. language-specific aspects. *Language, Cognition and Neuroscience*, *34*, 826–840.

Kabata, K. (2013). Goal-source asymmetry and crosslinguistic grammaticalization patterns: A cognitive-typological approach. *Language Sciences*, *36*, 78–89.

Krauss, R. M., & Wienheimer, S. (1964). Change in the length of reference phrases as a function of social interaction: A preliminary study. *Psychonomic Science*, *1*, 113–114.

Krauss, R. M., & Wienheimer, 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., Fussell, S. R., Brennan, S., & Siegel, J. (2002). A framework for understanding effects of proximity on collaboration: Implications for technologies to support remote collaborative work. In P. Hinds & S. Kiesler (Eds.), *Technology and distributed work* (pp. 137–162). Cambridge, MA: MIT Press.

Kraut, R. E., Miller, M. D., & Siegel, J. (1996). Collaboration in performance of physical tasks: Effects on outcomes and communication. Proceedings of the CSCW 96 (pp. 57–66). New York: ACM.

Kraut, R. E., Fussell, S. R., & Siegel, J. (2003). Visual information as a conversational resource in collaborative physical tasks. *Human–Computer Interaction*, *18*, 13–49.

Kuhlen, A. K., & Brennan, S. E. (2010). Anticipating distracted addressees: How speakers' expectations and addressees' feedback influence storytelling. *Discourse Processes*, *47*, 567–587.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*, 1–26. https://doi.org/10.18637/jss.v082.i13

Lakusta, L., & Carey, S. (2015). Twelve-month-old infants' encoding of goal and source paths in "agentive" and "non-agentive" motion events. *Language Learning and Development*, *11*, 152–175.

Lakusta, L., & DiFabrizio, S. (2017). And, the winner is…a visual preference for endpoints over starting points in infants' motion event representations. *Infancy*, *23*, 323–343.

Lakusta, L., & Landau, B. (2005). Starting at the end: The importance of goals in spatial language. *Cognition*, *96*, 1–33.

Lakusta, L., & Landau, B. (2012). Language and memory for motion events: Origins of the asymmetry between source and goal paths. *Cognitive Science*, *36*, 517–544.

Lakusta, L., Muentener, P., Petrillo, L., Mullanaphy, N., & Muniz, L. (2016). Does making something move matter? Representations of goals and sources in motion events with causal sources. *Cognitive Science*, *41*, 1–13.

Lakusta, L., Wagner, L., O'Hearn, K., & Landau, B. (2007). Conceptual foundations of spatial language: Evidence for a goal bias in infants. *Language Learning and Development*, *3*, 179–197.

Lakusta, L., Spinelli, D., & Garcia, K. (2017). The relationship between pre-verbal event representations and semantic structures: The case of goal and source paths. *Cognition*, *164*, 174–187. https://doi.org/10.1016/j.cognition.2017.04.003

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

Luo, Y., & Baillargeon, R. (2005). Can a self-propelled box have a goal? Psychological reasoning in 5-month-old infants. *Psychological Science*, *16*, 601–608.

Matuscheck, H., Reinhold, K., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305–315.

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. C. (2002). Evidence of perspective-taking constraints in children's on-line reference resolution. *Psychological Science*, *13*, 329–336.

Narasimhan, B., Kopecka, A., Bowerman, M., Gullberg, M., & Majid, A. (2012). Putting and taking events: A crosslinguistic perspective. In A. Kopecka & B. Narasimhan (Eds.), *Events of 'putting' and 'taking': A crosslinguistic perspective* (pp. 1–20). Amsterdam/Philadelphia: John Benjamins.

Papafragou, A. (2010). Source-goal asymmetries in motion representation: Implications for language production and comprehension. *Cognitive Science*, *34*, 1064–1092.

Paraboni, I., & van Deemter, K. (2014). Reference and the facilitation of search in spatial domains. *Language, Cognition and Neuroscience*, *29*, 1002–1017.

Pasupathi, M., Stallworth, L. M., & Murdoch, K. (1998). How what we tell becomes what we know: Listener effects on speakers' long-term memory for events. *Discourse Processes*, *26*, 1–15.

Pechmann, T. (1989). Incremental speech production and referential overspecification. *Linguistics*, *27*, 89–110.

R Core Team. (2017). *R: a language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Regier, T. (1996). *The human semantic potential: Spatial language and constrained connectionism*. Cambridge, MA: MIT Press.

Regier, T., & Zheng, M. (2007). Attention to endpoints: A cross-linguistic constraint on spatial meaning. *Cognitive Science*, *31*, 705–719.

Rissman, L., Woodward, A., & Goldin-Meadow, S. (2018). Occluding the face diminishes the conceptual accessibility of an animate agent. *Language, Cognition and Neuroscience*, *34*, 273–288. https://doi.org/10.1080/23273798.2018.1525495

Rubio-Fernández, P. (2016). How redundant are redundant color adjectives? An efficiency-based analysis of color overspecification. *Frontiers in Psychology*, *7*, 153.

Rubio-Fernández, P. (2019). Overinformative speakers are cooperative: Revisiting the Gricean Maxim of Quantity. *Cognitive Science*, *43*, e12797.

Shockley, K., Richardson, D., & Dale, R. (2009). Conversation and coordinative structures. *Topics in Cognitive Science*, *1*, 305–319.

Talmy, L.T. Shopen(1985). Lexicalization patterns: Semantic structure in lexical forms. *Language typology and syntactic description* (57–149). New York: Cambridge University Press.

Tatone, D., Geraci, A., & Csibra, G. (2015). Giving and taking: Representational building blocks of active resource-transfer events in human infants. *Cognition*, *137*, 47–62.

van Buren, B., Gao, T., & Scholl, B. J. (2017). What are the underlying units of perceived animacy? Chasing detection is intrinsically object-based. *Psychonomic Bulletin Review*, *24*, 1604–1610.

Wardlow Lane, L., & Ferreira, V. S. (2008). Speaker-external versus speaker-internal forces on utterance form: Do cognitive demands override threats to referential success? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 1466–1481.

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

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, *69*, 1–34.

Yoon, S. O., & Brown-Schmidt, S. (2018). Aim Low: Mechanisms of Audience Design in Multiparty Conversation. *Discourse Processes*, *55*, 566–592. https://doi.org/10.1080/0163853X.2017.1286225

Yoon, S. O., & Brown-Schmidt, S. (2019). Audience design in multiparty conversation. *Cognitive Science*, *48*, Article e12774. https://doi.org/10.1111/cogs.12774

Zehr, J., & Schwarz, F. (2018). PennController for Internet Based Experiments (IBEX). https://doi.org/10.17605/OSF.IO/MD832

Zheng, M., & Goldin-Meadow, S. (2002). Thought before language: How deaf and hearing children express motion events across cultures. *Cognition*, *85*, 145–175.

Zheng, M., & Goldin-Meadow, S. (2002). Thought before language: how deaf and hearing children express motion events across cultures. *Cognition*, *85*, 145–175. https://doi.org/10.1016/S0010-0277(02)00105-1

APPENDIX

Figure A1 reports the results of a separate online experiment that was run prior to Experiment 2. In this earlier version, motion events were not presented inside the television screen. Instead, the interlocutor appeared in either the upper right or upper left hand corner of the motion clip. We abandoned this version of the online experiment in favor of the design employed in Experiment 2 because one concern with the TV-less display was that the interlocutor might be able to "see" more of the event than the speaker could.

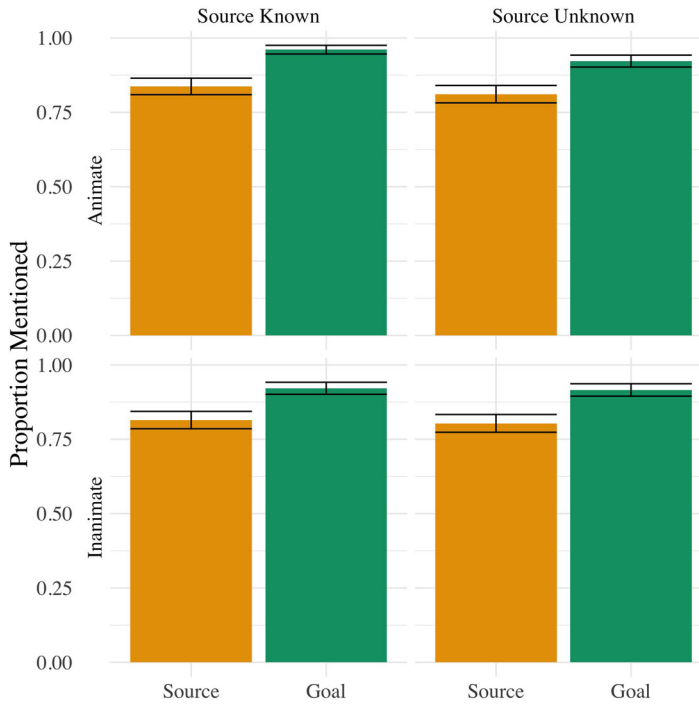

Fig A1. Results from an initial, web-based version of Experiment 2, in which motion events were not shown inside the television screen.

The results of this experiment were contrast coded and analyzed using logistic mixed effect regression models as in Experiment 1a and Experiment 2. As can be seen in Figure A1, we found the same pattern of results in this initial version as those reported in Experiment 2. Specifically, we detected only a main effect of Mention Type ($\beta = 1.76$, SE $= .37$, $|z| = 4.83$, $p < .01$) indicating that goals (93%) were mentioned significantly more frequently than sources (82%). As in Experiment 2, no other statistically significant effects were detected (p's $> .2$).