THE UNIVERSITY OF CHICAGO

Animacy and Audience Design: Representations of Core Components of Motion Events are Influenced by Conceptual and Pragmatic Factors

By James R. Kesan

August 2022

A paper submitted in partial fulfillment of the requirements for the Master of Arts degree in the Master of Arts Program in the Social Sciences.

Faculty Advisor: Monica Do

Preceptor: Ruben Rodriguez-Barron

Abstract

We examined the conceptual and pragmatic factors that influence memory (Experiment 1) and language (Experiment 2) for the conceptually core elements of simple source-goal motion events – the goal, or endpoint of motion. In Experiment 1, we presented animate or inanimate events to separate groups of participants and found that animate events were remembered more accurately than endpoints of inanimate events. In Experiment 2 we examined the influence of pragmatic factors related to audience design on participants' linguistic choices using an elicited production paradigm. We found that, unlike with sources or starting points of motion, the choice to mention or omit the goal was not affected by whether the goal was already known or unknown to the interlocutor. But the *way* that speakers talked about the goals did vary as a consequence of the pragmatic status of the goal. Our results shed light on the representation of inanimacy in the context of source-goal motion events and in conceptual representations of animacy, more generally. They also shed light on the way in which pragmatic factors related to audience design can affect language for conceptually core versus conceptually peripheral elements of events. We discuss the implications of our findings for theories of event cognition and language production.

1. Introduction

An important issue in the study of language is the seemingly simple question—How do people decide what they should talk about? As people's conceptual representations of the world are necessarily richer than their linguistic representations (i.e., we cannot possibly describe everything that we see in the world), speaking is a challenge that requires people to not only decide what to mention and what to omit, but to do so rapidly and appropriately. In describing an event in the world, how do speakers know, seemingly intuitively, what they should mention to be understood, and what they may safely disregard?

In theoretical and practical terms, language production requires speakers to translate their mental representations of the world into a linguistic form. Thus, before language can be produced, speakers have to build an internal representation of the events and concepts that are important to their message. The structures that people build in their minds to represent events are incredibly rich and detailed. Consequently, in an event many features are important to grasp what has actually happened: this includes the identity of participants or settings in the event; the causal relationship between elements; the temporal structure including whether the event is ongoing or completed (Rissman & Majid, 2019)—in basic terms the, "Who does what to whom" of an event (Papafragou & Grigoroglou, 2019). Critically, there is far more information within a speaker's conceptual representation of an event than can be effectively described or conveyed during brief bursts of speech. The richness of a speaker's conceptual representation of the world, compared to the relative poverty or 'sketchiness' of a linguistic utterance means that speakers have to constantly make choices about what to include and what not include in their utterances.

Consider, for example, a squid that swims from a coral reef to some seaweed, see Figure 1. This simple event may plausibly be described in a number of ways using varying levels of detail

(eg. “The squid is swimming.”, “The squid swims away from the coral.”, “The squid swims to the seaweed.”, “The colorful, elegant squid emerges from behind the big, green coral and swims over to the inviting leaves of the seaweed.”).

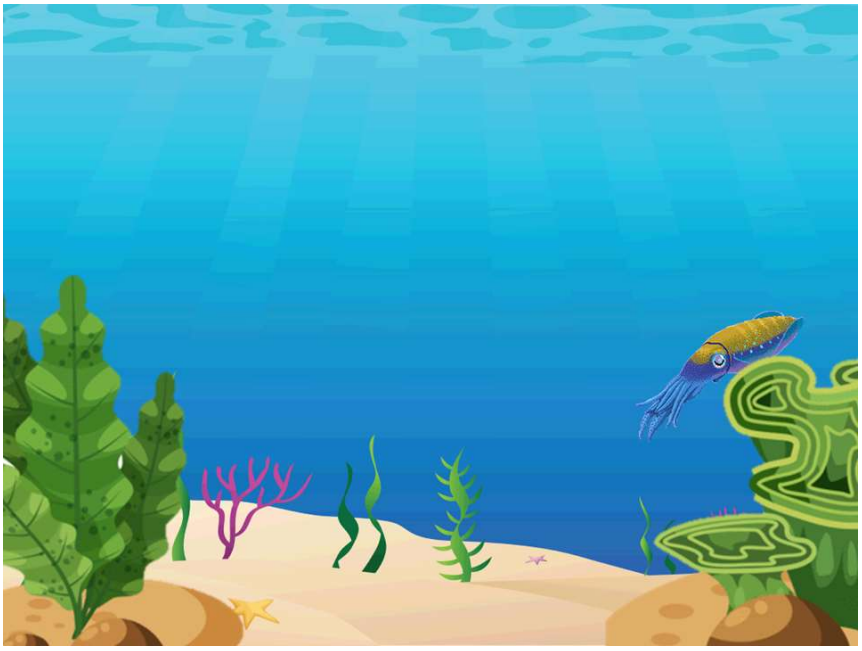

***Figure 1**. Sample event of a squid swimming from the coral to the seaweed.*

To make these choices, speakers often take into account the communicative context and the needs of listeners (Papafragou & Grigoroglou, 2019; Do et al., 2020). Prior work has shown, that one factor affecting the decision about what to mention and what to omit is audience design – the process of generating and continuously updating a model of the informational needs of an addressee. Although, prior work suggested that engaging processes associated with audience design may be computationally expensive (Pickering & Garrod, 2004; Brennan et al, 2010), it nevertheless allows speakers to manage the two competing pragmatic pressures that pervade everyday conversations: speakers must balance, on the one hand, the need to be informative by contributing new, relevant information with the benefits of minimizing the effort required to communicate (Grice, 2002). Audience design allows the speaker to identify and accommodate

information that is already in *common ground* – that is, information that constitutes shared knowledge between conversation partners typically stemming from prior discourse or from shared visual access during a conversation (Clark & Murphy, 1982; Lockridge & Brennan, 2002; Brennan et al., 2010).

A range of paradigms have been used to show that audience design can affect language production that refers to objects and events in the world. In the domain of events, prior studies typically using the referential communication paradigm, have investigated not only the ways in which speakers adjust to the knowledge state of the listener (e.g., by shortening words-Galati & Brennan, 2010 or calibrating the definiteness of their determiners-a/the/etc. Clark & Marshall, 1981) but also the limits of audience design for adults and children (Nadig & Sedivy, 2002; Fukumura, 2016; Grigoroglou & Papafragou, 2019a, 2019b).

In the domain of events, experiments have also used an elicited production task in which speakers describe an event to a listener or addressee who may or may not have visual access to the event. These studies have shown that informativity is influential in determining what speakers actually include in their utterances: initial work by Brown & Dell (1987), for instance, showed that speakers were more likely to mention atypical components of an event (e.g., using a spoon to hammer a nail into a piece of wood) than they were to mention typical components (e.g., using a hammer) because atypical event components were more challenging to infer from context (Brown & Dell, 1987). Building on this work, later research by Lockridge & Brennan (2002) found that this effect of atypicality was especially pronounced when speakers were recounting stories to listeners who also lacked visual access to the events. This demonstrates that speakers incorporate more general or generic information about what is readily inferable for their addressee with specific information derived from the visual perspective of the addressee.

1.1 Source-Goal Motion Events

Recently, though, investigations into the descriptions of source-goal motion events – events in which a figure in motion travels from a starting point (i.e., Source) along some path to an end point (i.e., Goal); for example, the event of a squid$_{\text{FIGURE}}$ swimming [from the coral$_{\text{SOURCE}}$] [to the seaweed$_{\text{GOAL}}$] – have shown that in addition to the pragmatic factors associated with audience design, there are also conceptual factors affecting the decision about what to mention or omit from an utterance.

Prior work, has shown that in the case of source-goal motion events, goals are linguistically privileged over sources. Across a number of typologically different languages, goals tend to be mentioned over sources (Ihara & Fujita, 2000; Regier & Zheng, 2007; Johanson et al., 2019) and carry finer-grained semantic distinctions (Regier & Zheng, 2007; Johanson et al., 2019). Importantly, though, other work has found a similar bias towards the goal completely independent from language, suggesting that the prominence of the goal is not just a feature of language, but instead a phenomenon traceable to privileged conceptual status of the goal itself. The bias for the goal, for instance, is present among homesigners and signers of an emerging language, both of whom lack prior exposure to conventional language (Zheng & Goldin-Meadow, 2002; Goldin-Meadow, 2003). Likewise, researchers in human development have independently noted a bias towards attending to the endpoints, but not the starting points, of actions and motion among pre-verbal infants (Gergely & Csibra, 2003; Biro et al., 2007; Woodward, 1998; Lakusta & Carey, 2015). In addition, during memory tasks that utilize verbal interference, goals are still remembered with greater frequency than sources (Papafragou, 2010; Lakusta & Landau, 2005, 2012; Do et al., 2020; 2022). Taken together, these studies have argued for a very tight relationship between the way that individual components of an event are conceptually represented and the way that they are

talked about: In particular, they argue that people tend to mention goals because they are a conceptually core element of the event; they tend to omit sources because sources are conceptually peripheral to the representation of the event.

In recent work, Do and colleagues (2020, 2022) have pointed to an interaction between the conceptual and the pragmatic factors influencing the decision about what to mention and what to omit when describing an event in the world. In particular, these studies have shown that unlike conceptually peripheral aspects of an event (e.g., sources), the more conceptually prominent event elements (e.g., goals), tend to be resistant to the usual consequences of audience design (Do et al., 2020, 2022). In their studies, participants provided descriptions of source-goal motion events to an addressee who either knew or did not know the source or goal of the motion event (Do et al., 2020; 2022). Do et al. found that sources tended to be mentioned when they were pragmatically important – i.e., when they were not already known to their addressee – and tended to be omitted when they were pragmatically unimportant to the addressee – i.e., already when the identity of the source object was already known. By contrast, changes to pragmatic status of the goal – i.e., whether it was or was not already known to the addressee – had little effect on language for the goal. This demonstrates that pragmatic factors do not affect sources and goals equivalently. Indeed, it appears that goals are largely impervious to the effects of their pragmatic status and thus do not respond to being made uninformative (Do et al., 2020; Papafragou & Grigoroglou, 2019).

1.2 The Current Study

The purpose of this the current study is to better understand both the conceptual and pragmatic factors that impact the decision to mention or omit the goal. In line with prior work, we focus our domain of inquiry on source-goal motion events because: (i) the representation of source-goal motion events has been well-studied in language and event conceptualization and (ii) because

these events offer speakers a wide range of choices in terms of how they may discuss the event. The origin, endpoint, or other elements could all be included or freely omitted.

In Experiment 1, we focus on one specific feature of source-goal motion events – namely, the animacy of the figure in motion – because the extent to which Figure animacy can modulate the conceptual (i.e., non-linguistic) prominence of the goal is unclear. Although initial work by Lakusta & Landau (2012) suggested that motion events with animate and inanimate figures are represented differently, more recent work investigating the role of animacy has yielded mixed results. Some studies (Lakusta & Landau, 2012; Lakusta & Carey, 2015) have shown that the goals of animate motion events are prominent in the non-linguistic representation of a motion event while goals of motion events containing inanimate figures are not (Lakusta & Landau, 2012). Other studies, using the same paradigm as Lakusta & Landau (2012), however, have found a similar bias for the goal in memory for both animate and inanimate figures (Do et al., 2022). Experiment 1 thus uses a memory task, specifically the change detection paradigm, to investigate the conceptual status of the goal when the figure is animate (Animate Figure Conditions) versus when it is inanimate (Inanimate Figure Conditions).

In Experiment 2, we use an elicited production paradigm to see whether it is possible for *pragmatic* factors related to audience design to exert an influence on speakers' decision to mention the goal of source-goal motion events. Given findings by Do et al. (2020), it is possible that the conceptual prominence of the goal will continue to affect its representation in language, resisting an influence from pragmatic cues: In addition to being the endpoint of motion, the Goal can be identified as the temporal endpoint of a motion event, meaning that the phrases used to describe the Goal often carry key information about when motion occurred or whether a motion event is ongoing or completed (Do et al., 2020). However, it is also possible that the manipulation used by

Do et al. (2020) was not quite strong enough to convey the knowledge state of the interlocutor. Prior work, for instance, has shown that adults can fail to make inferences using audience design if doing so is not obligatory to complete a task or they can make incomplete inferences that do not represent a full model of their addressee (Horton & Keysar, 1996; Nadig & Sedivy, 2002; Brennan et al., 2010). An alternative possibility, then, is that goals are sensitive to pragmatic cues, but more direct language for the status of information about whether the goal is in common ground (Goal Common Ground Conditions) or not (No Common Ground Conditions) is necessary to change the way in which the goal is discussed by speakers. In addition, given prior results showing that the conceptual status of goals depends on the animacy of the figure in motion, Experiment 2 also investigates the pragmatic status of the goal (previously known versus unknown) both when the goal is expected to be conceptually prominent (Animate Figure Conditions) and when it is not (Inanimate Figure conditions).

2. Experiment 1: Memory (Non-Linguistic Cognition)

An open question is the extent to which the conceptual status of the goal is impacted by the animacy of the figure in motion: Whereas some studies (Lakusta & Landau, 2012; Lakusta & Carey, 2015) have shown the goal to be conceptually prominent *only* when the figure is animate, other studies (Do et al., 2022) have shown the goal to be conceptually prominent both when the figure is animate and inanimate.

One explanation for these diverging sets of results may be related to differences in the sets of materials employed by these two groups of researchers. Where Lakusta & Landau (2012) chose to stage live-action stimuli that depicted rather different types of motion events in animate (e.g., a person walking from a chair to a ladder) versus inanimate conditions (e.g., a piece of paper falling off the edge of a container onto a candle), Do et al. (2022) used animated video clips that allowed

for far more similarities across animate (e.g., a baby crawling from a crib to a toy chest) versus inanimate (e.g., a toy ball rolling from a crib to a toy chest) conditions. Under this account, replicating a result where the goal is prominent in memory for animate and inanimate events is straightforward. If we present nearly identical animated stimuli to separate groups of participants, changing only the animacy/inanimacy of the figure, this account predicts that the goal should be equally prominent, or recalled with equal accuracy, for both groups of participants.

A second possibility, is that the distinct outcomes for animacy were a consequence of other differences in experiment design. While the studies by Lakusta and colleagues manipulated animacy between-subjects, the work by Do et al. (2022) manipulated animacy within-subjects. Importantly, as a consequence of allowing participants to view both animate and inanimate figures within the same study, it is possible that participants in the latter set of studies were presented with the opportunity to make direct comparisons between animate and inanimate figures while the participants in the Lakusta & Landau (2012) studies were not. In this case, the conceptual prominence of the goal found in the case of inanimates may have been due to high level inferencing effects in which participants reason that because the inanimate stimuli, like the animate stimuli, are the focus of attention in a scene, the two types of stimuli should be treated as equally prominent or newsworthy.

To tease apart these possibilities, the current study manipulates animacy between subjects. Crucially, though, all other aspects of the animate versus inanimate motion events – including the backdrop, the path of the figure in motion, the source landmark, the goal landmark, etc. – will be identical. We follow prior work in using a change detection task in which participants viewed pairs of animated motion events and were required to detect whether the second animate was "exactly the same as the first" (Papafragou, 2010). It follows that because an interference task was included

in the procedure to eliminate verbal encoding as an aid during memory encoding and recall, this paradigm allowed us to study the role of animacy in the conceptual representation of the goal itself.

If previously observed differences in the response to animate and inanimate figures are really an effect of the different stimuli used by Lakusta and Landau (2012) but not by Do et al. (2022), then highly similar stimuli in the present study should lead to very similar performance across participant groups, even when animacy is manipulated between participants. On the other hand, if the prior discrepancies in findings for animate and inanimate motion events can be explained by differences in between- versus within- subject experimental designs, then we might expect to find differences between animate and inanimate figures when they are being presented to entirely separate groups of participants – even though the backdrops, the paths of motion, and the speed/timing of the motion are otherwise identical. In particular, we would expect that the goal of an animate motion event should be more conceptually prominent than the goal of an inanimate motion event and that this should encourage a higher rate of accuracy in the memory task.

2.1 Methods

2.1.1 Participants

Participants were 96 adult native speakers of American English recruited using Prolific: 48 in the Animate condition and 48 in the Inanimate Condition. They were paid $3.50 for approximately 20 minutes of participation. Participants were excluded if they scored below chance on a set of filler items that did not depict source-goal motion events (n=3) or if they performed two standard deviations below the average on an unrelated interference task involving simple math questions (n=5). The data from the remaining 88 (42 Inanimate/46 Animate) participants was analyzed.

2.1.2 Materials & Design

The experiment was a 2x2 within-by-between design with the factors Change Type (No change/Goal change) and Animacy Type (Animate/Inanimate). These conditions were rotated across four lists with a Latin Square design that was used to rotate the Change type for the trials included within each list. This yields two lists with animate figures and two with inanimate figures that contain target items belonging to both change types.

Using Microsoft PowerPoint, we created a total of 40 items for each animacy condition: 16 target items that depicted source-goal motion events and 24 filler items that did not. Source-goal motion events depicted situations in which a single figure moved from a source location at the right or left side of the screen to a goal location grounded at the opposite side of the screen. The location of the source and goal were right/left counterbalanced across trials. Filler items depicted other motions like spinning, movement in circular patterns, or movement to an endpoint located offscreen. So that participants would not anticipate target items, some filler items were designed to look visually similar to target events.

For half of the target items (8 animations) and half of the filler items (12 animations) in each list, we also created a second, changed version of the video clips. For target items, this changed version always involved a Goal Change, in which the goal landmark object was replaced by a within semantic category variant, or similar class of item (eg. seaweed is replaced using an image of seaweed; Figure 2), occupying similar dimensions on the screen. For filler items, this changed version included substituting one background for another similar scene, exchanging an object located in the foreground of the scene, or changing the figure in motion see Figure 3.

A complete set of materials and demo version of the experiment, along with pre-registration information is available here: https://osf.io/ebvw7.

Figure 2. *Variations on a sample event- "The squid swam from the coral to the seaweed" tested in Experiment 1. The animate (squid) or inanimate (anchor) figure would proceed from the source (Coral) at the right of the screen to the goal (Seaweed) at the left. In the Goal Change Condition, the goal would be replaced by a within category variant, a different image of seaweed.*

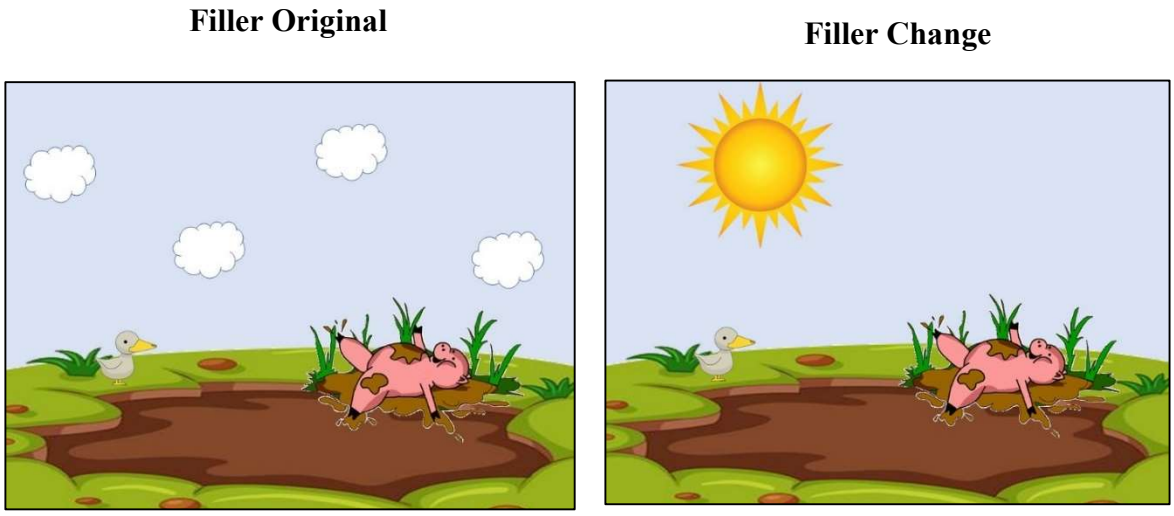

Figure 3. *Samples from filler animations used in Experiment 1. The image on the left is from the original filler item and the still on the right depicts the change to the background scene that was used in the alternate version of this filler item.*

2.1.3 Procedure

The experiment was scripted and implemented using PennController Ibex (Zehr & Schwarz, 2018). After consenting to participate in the experiment, participants completed 40 trials in either the Animate or Inanimate conditions. Within each trial, participants were first shown an animated video clip were asked to remember the clip to the best of their ability. Participants were only allowed to view the clip once and were only permitted to proceed to the next screen via button click after the animation had finished playing. After this, participants were asked to respond to two simple math questions containing simple arithmetic calculations like 45+4+5=54 or 75-25-3=47. This unrelated interference task was designed to interrupt attempts at verbally encoding information about the video. Participants then viewed either the exact same animation a second time (No Change Conditions) or the changed variant of that item (Goal Change Conditions). Participants clicked "Yes" if the video was exactly the same as the video they had just seen and "No" otherwise. Trials were pseudo-randomized such that participants never viewed more than two consecutive target items.

2.2 Results

2.2.1 Data Analysis

Accuracy in the memory task was analyzed using logistic mixed effects models in R (R Core Team, 2021) using the `glmer` function of the lme4 package (Bates et al., 2015). The outcomes for individual trials were binary coded (0 = Incorrect, 1 = Correct) according to whether an accurate answer was provided. The main variables of interest were the Animacy type of the figure (Animate/Inanimate) and the Change type (No change/Goal change). These variables were contrast coded using a dummy coding schema with Inanimate and Goal Change set as reference levels. An Animacy by Change type interaction term was also included in the model as a fixed effect.

We started with the maximal random effect structure which included Change Type and Animacy Type included as by-item effects and Change Type included as a factor in by-subject random effect structures. Following the procedure outlined by Matuschek et al. (2017), the random effect structure was simplified until (i) the model converged or was no longer overfitted and (ii) only those effects contributing significantly to model fit were included. The final model is listed below.[1]

2.2.2 Results

As can be seen in Figure 4, participants in our study performed at above chance levels across all conditions (mean 78.59% across all target items). In addition, our results showed that overall, the No Change conditions (teal bars; 85.52%) without an exchange of the goal item were significantly more accurate than Goal Change conditions (coral bars; 71.65%). This was corroborated statistically by the main effect of ChangeType (β=1.14, SE=.34, |z|=3.39, p<.001)

[1] Final model for Experiment 1: Sogomodel3=glmer(scoring ~ ChangeType*AnimacyType + (1+ChangeType|IP_address) + (1+ChangeType|ItemID) , family=binomial (link = "logit"), data=reg_list, control= glmerControl(optimizer = "bobyqa", optCtrl = list(maxfun=2e5)))

which suggests that the probability of observing an accurate response in the No Change condition was approximately 19.19% higher than in the Goal Change condition. Given that the procedure was structured such that participants provided a "yes" response to indicate the No Change condition, we interpret this as being consistent with the well-known "yes bias" (Hinz et al., 2007).

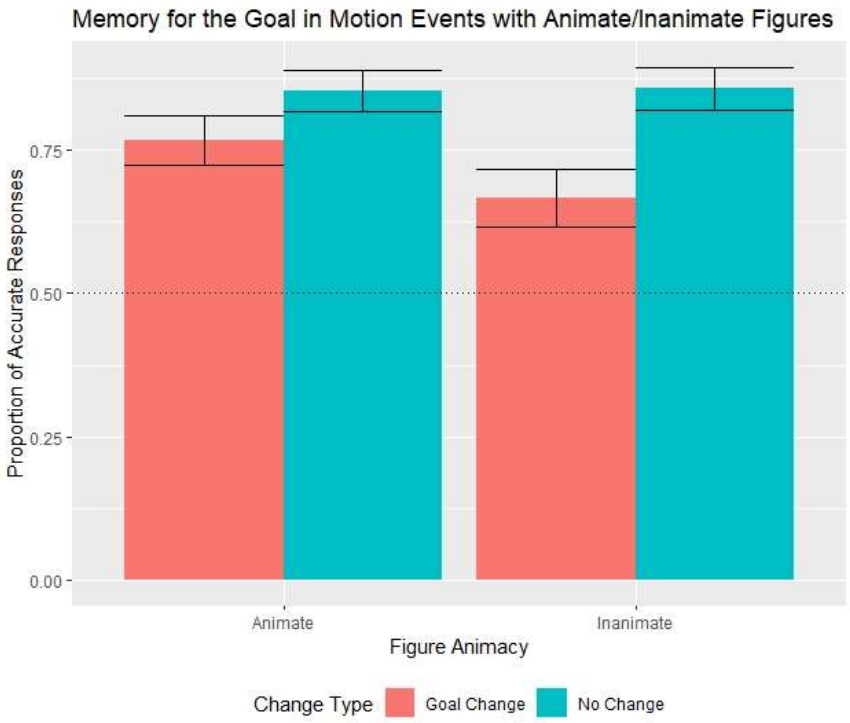

Figure 4. *Shows the mean proportion of correct responses in Experiment 1 with the dotted line representing chance performance and error bars indicating the 95% confidence interval.*

In addition, as demonstrated by a significant main effect of Animacy type (β=.56, SE=.25, |z|=2.21, p< .05), we found that responses in the Animate conditions (80.98%) were significantly more accurate than those in the Inanimate conditions (76.19 %). However, we did not find evidence for a significant interaction effect between Animacy and Change type (β=-.6, SE=.39, |z|=1.52, p=.13). Overall then, in line with our predictions, participants were more accurate at detecting changes to the goal when the figure in motion was Animate versus when it was Inanimate.

2.3 Discussion

The aim of Experiment 1a was to obtain more decisive evidence for whether differing animacy results in previous experiments may have been due to the qualities of the stimuli for animate and inanimate motion events or due to differences in the experiment design that permitted more higher-order reasoning about animates and inanimate figures. To do this, we used stimuli that were otherwise *identical* with the exception of the animate or inanimate figure and tested this variation between participants.

Consistent with the results reported by Lakusta and colleagues (2012) and unlike Do et al. (2022) we found that the animacy of the figure has a significant effect on participants' accuracy. The significantly higher accuracy rates in the Animate conditions showed a broad tendency to recall the elements of motion events, including endpoints of motion, more accurately when those events involve Animate versus Inanimate figures of motion. Given that the stimuli in the Animate and Inanimate conditions were otherwise identical, these results suggest that underlyingly the reason for the difference between Lakusta and colleagues (2012) and Do et al. (2022) was, indeed, attributable to the between-subjects versus within-subjects experimental design. In other words, whether participants did or did not have the opportunity to see animate and inanimate figures within the context of the same study played an important role in the way that inanimate motion events were construed. Importantly, when participants were exposed to an inanimate figure that had been foregrounded in the same way as an animate figure (Do et al., 2022), participants represented the motion event in the same way – that is, the goal of both inanimate and animate motion events were conceptually privileged.

In addition to clarifying an open question in the literature, our results also have broader implications for how we might best understand the relationships between animate versus

inanimate figures. In particular, our results point to an inequality in how animacy is construed in the mind of the viewer: If inanimate figures can be construed like animate figures as a consequence of a close temporal juxtaposition with animate figures, then this suggests that it may be relatively easy to promote a conceptually less prominent inanimate entity to a status comparable to the more prominent animate figure. Notably, it is possible for a motion event with an animate figure to be construed as inanimate, and consequently demoted or reduced in relative prominence as a consequence of being associated with inanimacy. But this reverse situation was not observed either in our data here or in any of the previous studies. This calls attention to a broader pattern that affects conceptual factors, including representations of sources and goals: it appears to be easier to promote conceptually less prominent event components like inanimate entities or sources of motion to greater prominence (Lakusta et al., 2016; Lakusta & DiFabrizio, 2017) than it is to reduce the prominence of already conceptually core elements of events like animate entities or goals of motion (Do et al., 2020). We return to this point in the General Discussion.

3. Experiment 2: Language Production

Given the results of Experiment 1 showing that the animacy of the figure in motion can directly affect conceptual representations of the goal, the aim of Experiment 2 is to investigate the extent to which pragmatic factors related to audience design can affect the decision to linguistically mention or omit the goal both when the goal is conceptually prominent (Animate Figure condition) and when it is not conceptually prominent (Inanimate Figure condition).

To accomplish this, we asked a new group of participants to describe a sequence of short videos depicting the same animate or inanimate source-goal motion events used in Experiment 1. We varied the knowledge state of the addressee by asking participants to describe these events to

a physically co-present confederate experimenter who they were told either: (i) had visual access to and thus already knew about the endpoint of the motion event (Goal Common Ground Condition) or (ii) had no information about the event at all (No Common Ground Condition).

Peripheral elements in motion events, like sources of motion, have been shown to be highly susceptible to pragmatic factors related to audience design and in particular to their informativity for the listener (Lockridge & Brennan, 2002; Brennan et al., 2010; Do et al., 2020, 2022): Speakers are overwhelmingly more likely to omit sources of motion from their description when the source is already known to the listener than when it is unknown. But this is not necessarily true for goals, which are conceptually prominent elements in motion events – at least when the figure in motion is Animate (Lakusta & Landau, 2012; Lakusta & Carey, 2015; Exp 1). If, as suggested by Do et al. (2020), conceptually prominent elements are resistant to pragmatic pressures related to audience design, then we might expect to find that goals, unlike sources, are not likely to be omitted even when they are pragmatically uninformative. If, however, Do et al. (2020) failed to find evidence of a pragmatic influence simply because the cue to the pragmatic status of the goal in their experiment was not strong enough we expect a different pattern. In particular, we should see a higher rate of goal omissions in the Goal Common Ground condition, when the goal is already known, and thus, uninformative to the listener as compared to the No Common Ground condition, where the goal is completely unknown and thus important to mention to the addressee.

An interesting question is how pragmatic factors related to audience design will affect goal mentions in the case of motion events containing inanimate figures, where goals are not expected to be conceptually prominent. One possibility is that we may see the goal being influenced by pragmatic factors even though, conceptually, it is not expected to be prominent. In this case, we might expect to see an effect driven by the pragmatic asymmetry in which the goal is a departure

from the initial state of a motion event, and is therefore in need of explication. If so, the goal would be mentioned frequently-similar to the goal of an animate motion event, despite the inanimate figure, because it is still pragmatically important to the individual engaged in packaging the essence of a source-goal motion event into language.

3.1 Methods

3.1.1 Participants

Participants for Experiment 2 were 64 students recruited from the University of Chicago. Determination for the number of participants was pre-registered (https://osf.io/eaj6z) and based on comparison to prior work. All participants were native speakers of American English and at least 18 years old at the time of the study. Participants who completed the study were compensated with either course credit for a psychology course or $13 for 45-60 minutes of participation. A small set of participants were excluded from the study and their data was not part of the final analysis (n=5). Three of these participants were excluded because they either asked explicit questions about whether the confederate experimenter was part of the study or expressed suspicions to this effect during debriefing. Two additional participants were excluded for failing to understand the instructions given for the main experiment task.

3.1.2 Materials & Design

Motion videos used in Experiment 2 were taken directly from the No Change condition of Experiment 1. The design of Experiment 2 was a 2 x 2 fully between-subjects design with the factors Animacy (Animate/Inanimate figure) and Ground Type (Goal Common Ground/No Common Ground). This design produced four different experimental lists, to which participants were randomly assigned.

3.1.3 Procedure

Upon entering the lab, participant and confederate were seated at the same table with two monitors facing in opposite directions, so that each individual could see the events unfolding on their own computer screen but were not able to see the monitor of their conversation partner.

Before the primary experiment procedure, the participant and confederate would collaborate to solve a puzzle known as the Tower of Hanoi. This provides the confederate experimenter with the opportunity to demonstrate that they are an active and engaged partner in the task and to convey the impression that they are a naïve participant in the procedure. This is significant as prior scholarship suggests that participants in elicited production procedures will provide more information to attentive and invested addressees (Pasupathi et al., 1998; Brennan et al., 2010).

In both Goal Common Ground and No Common Ground conditions, participants were told that they would be viewing short video clips and that they must describe them so that their listener, the confederate experimenter, could answer brief, innocuous questions about the events occurring in the video, such as, "How likely would you be to see an event like this in your own life?". In the No Common Ground condition, participants were told that their interlocutor (i.e., the confederate listener) would have no other information about the video and would need to rely exclusively on the participants' descriptions to answer the questions. In the Goal Common Ground condition, by contrast, participants were told that the listener would be able see the last frame of the video clip – namely, the frame depicting the figure at its final destination, the goal object/location – but nothing before that point. The instructions for the Goal Common Ground condition, therefore, position the goal as common knowledge between conversation partners. To ensure that participants fully understood what would/would not be visible to their conversational partners in both the Goal

Common Ground and No Common Ground conditions, participants were guided through one practice trial before continuing to the main experiment. To reinforce the practical implications of the Goal Common Ground condition for the participant, the confederate listener would point to their screen and ask the experimenter to confirm that the still image on their screen was indeed the end of the video clip. After the practice trial, the experimenter left the testing room.

Each trial in the experiment shared a common structure. First, the participant would press the spacebar to begin the videoclip. Next, they would describe the video clip to the confederate experimenter who would pretend to answer questions on their monitor. Once the confederate had successfully 'answered' the question, they indicated either verbally or non-verbally that they were ready for the next trial. Confederates were instructed to engage in consistent verbal (e.g., "ok", "uh-huh", "mhm") and non-verbal (e.g., nodding or maintaining eye-contact with the participant) back-channeling throughout the experiment (Clark & Krych, 2004; Ward, 2006). This sequence was repeated for all 40 trials. The audio for each experimental session was recorded using a microphone and the digital audio recording software Audacity. Materials for this study are also available on OSF (https://osf.io/eaj6z).

3.2 Results

3.2.1 Data Analysis

In Experiment 2, we were interested in three different variables: (i) the proportion of goal mentions (1 = Y/ 0 = N), (ii) the definiteness of the determiners immediately preceding the first mention of the source and/or goal landmarks (0=indefinite/1=definite); and (iii) the proportion of utterances in which the speakers made explicit reference to what they believed their interlocutor

could or could not see (1 = Y/ 0 = N) – what we call Explicit Reference to Addressee Visual Access.[2]

To analyze these variables, data were transcribed and coded by the first author and research assistants who were blind to the study conditions. Then, to check reliability between coders for each of these measures, the first author re-coded a randomly selected 30% of the data. For goal mentions, there was 97.1% agreement between coders and analyses using Cohen's kappa returned an inter-rater reliability score of $\kappa = .81$ for mentions of the goal. Source mentions were coded with 97.1 % agreement between coders and had an inter-rater reliability score of $\kappa = .89$. Interrater reliability scores calculated using Cohen's kappa returned inter-rater scores of $\kappa = .81$ for definiteness of goal determiners and $\kappa = .73$ for definiteness of source determiners. Our third variable, in which we coded whether participants made Explicit Reference to Shared Visual Access, showed 95.1% agreement with an interrater reliability score of $\kappa = .61$.

Analyses over all variables were performed using logistic mixed effects models in the lme4 package for R (Bates et al., 2015). In all cases, the independent variables Animacy Type and Ground Type were dummy coded with Inanimate and Goal Common Ground as the reference groups. Both of these factors, as well as their interaction, were included in each analysis as fixed effects. We always started with the maximal random effects structure, which included random intercept for each subject and each item, in addition, we also included random slopes for the Animacy and Ground Type factors in the by-item random effect terms. The model was simplified

[2]A number of other variables, which were not central to our research question/pre-registered as exploratory, were also coded – including variables such as the types of syntactic phrases used most frequently for goal/source mentions (Appendices A-C), etc. These are available in the appendix. We do not discuss these further since they do not relate to the question at hand.

following the procedure outlined by Matuschek et al. (2017) until it (i) converged or was no longer overfitted and (ii) only effects that contributed significantly to the model fit were maintained.

3.2.2 Goal and Source Mention Results

Given that our main question of interest was whether participants would be more likely to omit the goal when it was already known to an interlocutor, our primary variable of interest was the overall proportion of utterances that featured a mention of the goal. If goals or conceptually core elements of an event *are* sensitive to pragmatic factors related to informativity, then we might expect to see a higher proportion of goal mentions in the No Common Ground condition and a lower proportion of goal mentions in the Goal in Common Ground condition.

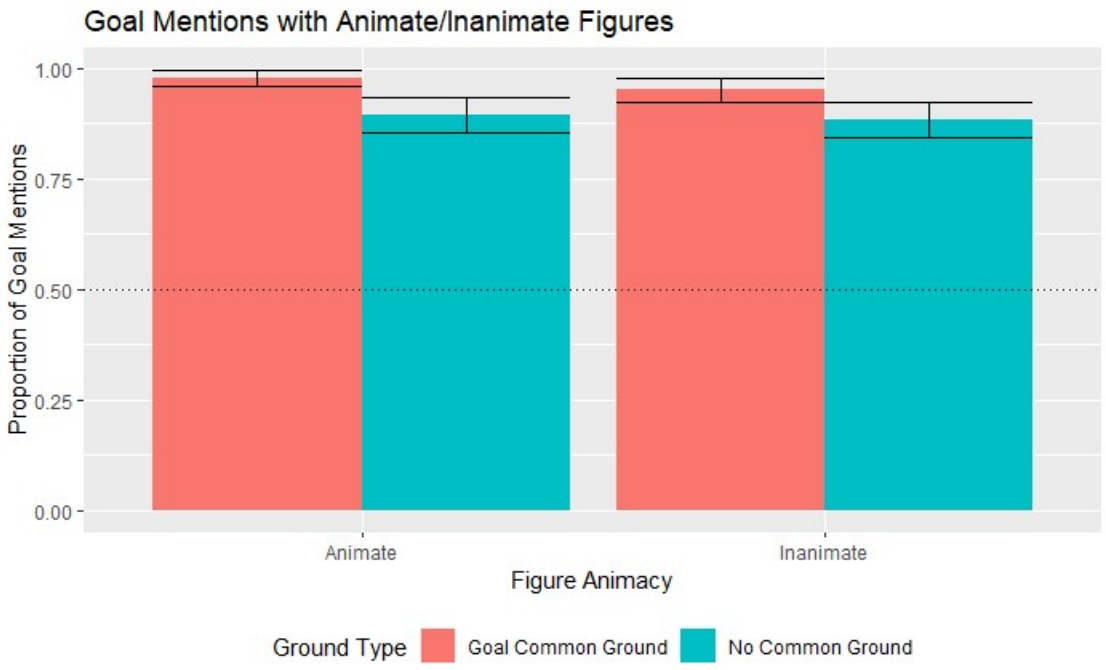

Figure 5. *Proportion of in-lab trials featuring Goal Mentions in language production. Error bars indicate the 95% confidence interval*

As seen in Figure 5, this prediction was not borne out. In particular, we did not observe any statistically significant differences in the rate of Goal mentions as a result of Animacy or

Common Ground Condition.[3] The mean rate of Goal mentions in the Animate condition (93.55%) was quite similar to the rate in the Inanimate condition (91.99%). This negative result is illustrated by a non-significant main effect of Animacy type (β=1.24, SE=1.05, |z|=1.18, p=.24). Nor did the model provide evidence for an effect of Common Ground. The Goal in Common Ground condition had a numerically higher rate of goal mentions (96.48%), as compared to the No Common Ground condition (89.06%); but this difference was not found to be statistically reliable (β=-1.19, SE=.76, |z|=1.57, p=.18). The interaction also did not reach significance, suggesting that there is not a reliable difference in the effect of animacy or inanimacy on goal mentions across the two Common Ground conditions (β=-.99, SE=1.13, |z|=.88 p=.38).

Although the rate of source mentions was not central to our hypotheses, we nevertheless submitted these to analysis because they, being susceptible to pragmatic factors related to audience design (Do et al., 2020, 2022), offered a useful counterpoint to goals. [4] Interestingly, we found that sources were mentioned significantly more frequently (β=-3.95, SE=1.53, |z|=2.58, p<.01) in the Goal in Common Ground Condition (94.53%) compared to the No Common Ground condition (77.54%). Although it is possible to imagine a possible explanation for this pattern of results, the effect is neither predicted by extant pragmatic accounts nor pertinent to the present study, and so we do not interpret it further as further replication is needed[5].

[3] The final model after model comparison is given here: glmer(GoalMention2 ~ AnimacyType2*GroundType2 + (1|SubjID) + (1+AnimacyType2 |ItemID), family=binomial(link = "logit"), data=gdata2, control=glmerControl(optimizer = "bobyqa",optCtrl = list(maxfun=2e5)))

[4] Final model for Source mentions- glmer(SourceMention2 ~ AnimacyType2*GroundType2 + (1|SubjID) + (1+AnimacyType2|ItemID), family=binomial(link = "logit"), data=gdata2, control=glmerControl(optimizer = "bobyqa",optCtrl = list(maxfun=2e5)))

[5] One possible account of this is that because goals are evidently known to the interlocutor, speakers in the Goal in Common Ground condition construe sources as being especially useful for an interlocutor. In the No Common Ground condition, by comparison, both goals and sources are presented as unknown, therefore this condition would not allow speakers to infer that the source is especially important to their listener.

As in the case of goals, this main effect of Ground Type was not modulated by effects of animacy ($p > .2$); nor, did we detect any significant main effect of animacy on source mentions ($p > .6$).

3.2.3 Goal Definiteness Results

In addition to the proportion of goal mentions, we also categorized the determiners immediately preceding the first mention of both source and goal as being either a definite or indefinite determiners (0=indefinite/1=definite) following the classification system outlined by Abbott (2006).[6] We were interested in determiners because they can provide a subtle, but useful proxy of what the participant believes the confederate knows. Prior work has shown, for instance, that indefinite determiners, like "a/an" or "some" are typically used to introduce information that is discourse-new, and thus previously unknown to an interlocutor, while definite determiners like "the" or "that" are typically known mark 'given' information – that is, information that is already known to both parties in a discourse. In addition to the mention or omission of the goal, then, the definiteness of the determiner can also provide insight into whether the speaker views the goal as already in common ground (e.g., via definite determiners) or not yet mutually known (e.g., via indefinite determiners) (Clark & Marshall, 1981; Clark & Murphy, 1982).

For goals, we found that participants were significantly more likely to use definite determiners (e.g., "the", "that", "those") in the Goal Common Ground condition (94.74%), where participants were explicitly told that the goal would be visible to interlocutors on the final frame of the video, than in the No Common Ground Condition (23.53%), where participants were told

[6] The final model for Goal Determiners is given here: glmer(GoalDeterminer2 ~ AnimacyType2*GroundType2 + (1|SubjID) + (1|ItemID), family=binomial(link = "logit"), data=gdata2, control=glmerControl(optimizer = "bobyqa",optCtrl = list(maxfun=2e5))) and the final model for Source Determiners is here: Final Model for Source Determiners-glmer(SourceDeterminer2 ~ AnimacyType2*GroundType2 + (1|SubjID) + (1|ItemID), family=binomial(link = "logit"), data=gdata2, control=glmerControl(optimizer = "bobyqa",optCtrl = list(maxfun=2e5)))

that their interlocutor would not be able to see any part of the video. This was indexed by a significant main effect of Common Ground condition (β=-6.15, SE=.92, |z|=6.7, p<.001) indicating that the odds of observing a definite goal determiner were significantly lower in the No Common Ground condition than in the Goal Common Ground condition. Participants' determiners when talking about the goals of motion events thus suggested that they were not only attending to, but also adjusting their utterances to accommodate the knowledge state of their addressee: They used definite determiners when talking about goals which they knew were visually accessible, and thus, already known to their interlocutor; but used indefinite determiners when introducing goals that were not visually accessible, and thus, completely new to their addressee.

A look at source determiners showed the same general pattern of results. In the Goal Common Ground condition approximately 94.14% of the determiners preceding sources were definite, while in the No Common Ground condition only 26.3% of the observed determiners were definite. Statistical analyses showed that the odds of a speaker using a definite source determiner were significantly lower in the No Common Ground condition (β=-5.12, SE=.74, |z|=6.94, p<.001), confirming that participants were aware of the fact that the ground items located at the sources were visible to their addressee in the final frame of the Goal Common Ground, but not in the No Common Ground conditions.

As expected, for both goal and source determiners, there were no main or interaction effects involving Animacy (all p's > .2).

3.2.4 Explicit Reference to Interlocutor's Visual Access

A final variable of interest, and perhaps the most course-grained, was whether participants made any explicit reference to privileged information that their interlocutor could or could not

attain by viewing their computer monitor. [7] These included utterances such as "The squid swam from the coral to where it is now.", "The squid you now see at the seaweed, it swam from the coral." or "You should see some seaweed on your left, that is where the squid is right now."

Overall, the use of these explicit references to the interlocutor's visual access was not particularly common in our data: they only occurred in approximately 8.4% of trials. Although, a significant main effect of Ground Type showed that utterances containing these explicit visual access references were overwhelmingly more likely to occur in Goal in Common Ground conditions (95.35%) than in the No Common Ground conditions (4.65%; β=-3.83, SE=1.32, $|z|$=2.9, $p<.01$). The explicit references to visual access were also well represented in terms of the number of participants in the Goal Common Ground condition that chose to include an explicit reference to the interlocutor's visual access at some point during the procedure (50%), by comparison only (9.4%) included references to visual access in the No Common Ground condition.

We did not expect Animacy to play a role in the decision to make an explicit reference to the interlocutor's visual state. Indeed, it did not: we found no significant main or interaction effects involving Figure Animacy (p's > .4).

3.3 Discussion

Prior work (Do et al., 2020, 2022) has shown goals of motion, unlike sources of motion, to be resilient to pragmatic factors associated with audience design. To see whether this apparent resilience is more attributable to the conceptual status of the goal or to the efficacy of the pragmatic cues used in prior work, we replicated the design of their experiment using a much stronger

[7] Final Model for Shared Knowledge References- glmer(SharedKnowledgeReferences2 ~ AnimacyType2*GroundType2 + (1|SubjID) + (1|ItemID), family=binomial(link = "logit"), data=gdata2,control=glmerControl(optimizer = "bobyqa",optCtrl = list(maxfun=2e5)))

pragmatic cue and found that even in spite of this much stronger pragmatic cue, participants were still very unwilling to omit the goal from their utterances.

Given what is known of audience design – namely, that speakers tend to maximize efficiency by omitting information that is already known to their addressee – our results are surprising. If speakers were making their choices based on (pragmatic) informativity, goals should in principle have been mentioned less frequently in the Goal in Common Ground conditions, where it was clear that the goal was already known to the interlocutor. Theoretically, there should also be an impulse to include the goal more in the No Common Ground condition, where the goal constituted completely new information for the addressee.

At first glance, one way to interpret the results here may be to say that participants in our study, as in Do et al. (2020), were simply insensitive to the knowledge state of their interlocutor, because the cue used here was still too weak a signal of the interlocutor's knowledge state. We believe this account, though, is unlikely for three reasons. First, half of our participants in the Goal in Common Ground condition made explicit references to what was visible/known to their addressee and this occurred almost exclusively in the Goal in Common Ground conditions. Thus, it could not have been the case that participants were completely unaware of their interlocutor's knowledge state or more specifically, that the goal of the motion event was already known to their interlocutor.

A second piece of evidence against this account comes from the relatively high rate of source mentions in the present study compared to rates reported in prior work. In particular, prior research has shown that because sources of motion are conceptually peripheral to the representation of the event, they tend to be omitted from utterances (Papafragou, 2010; Lakusta & Landau, 2012) unless

they are especially conceptually/perceptually salient (Lakusta et al., 2016; Lakusta & DiFabrizio, 2017) or, relevant for our purposes, pragmatically informative (Do et al., 2020). In the present experiment, sources were mentioned in the vast majority of utterances (94.5% in Goal Common Ground, 77.5% in No Common Ground) which represents a very high rate compared to prior studies. Papafragou (2010) found sources being mentioned in just 56% of utterances, and Do et al. (2020) found that sources were mentioned around 85% of the time in a No Common Ground condition that was similar to the No Common Ground condition in the current study. One way to interpret these results is that the higher rates of source mentions in the present work were the consequence of pragmatic factors related to audience design. Speakers in our study were especially likely to mention sources because information about the source – which was always unknown to the listener – was what they determined to be most useful to their addressee.

Most compelling, though, is data from the distribution of definite versus indefinite determiners. In particular, our results show that when the goal is in common ground, speakers are more likely to refer to it using definite determiners that typically correspond to mutually-known, information that has already been introduced either by the discourse or visual context. By contrast, in No Common Ground conditions, speakers were more likely to refer to the goal using indefinite determiners that typically mark unknown, new, previously un-introduced information (Clark & Murphy, 1982; Lockridge & Brennan, 2002). This choice indicates that speakers are, indeed, aware of whether the Goal is or is not in common ground between speaker and addressee. Interestingly though, unlike with sources (Do et al., 2020, 2022), this awareness simply does not impinge on speakers' decisions to mention or omit the goal.

Cumulatively, it does not appear that our results are consistent with the possibility that speakers neglected to engage in audience design or that they engaged incompletely because the

pragmatic cue to the knowledge state of the interlocutor was too weak. Nor are they fully consistent with the possibility (Do et al., 2020) that goals are completely resistant to pragmatic factors related to audience design. Indeed, evidence from the types of determiners used to talk about goals (i.e., definite versus indefinite) shows that participants were, in fact, very aware of whether their addressee could identify the goal. This awareness simply does not prompt speakers to assess the goal in terms of uninformativity and omit it. Instead, our data suggest that the pragmatic factors associated with audience design appear to affect sources and goals in different ways. Informativity may affect decisions about whether to mention the peripheral components of events, like sources, but these same pragmatic factors affect decisions about – not whether, but *how* conceptually prominent elements of an event are discussed. In the General Discussion, we explore candidate explanations for why goals are stubbornly maintained by speakers.

4. General Discussion

Conceptual representations of events are far more detailed than the corresponding linguistic utterances used by speakers to characterize these events. This motivates compromises during language production as speakers determine the information that is most important to convey about an event and the most efficient means of making this understood. These concerns are balanced with the need of listeners or addressees, using a process called audience design.

Recent research in the area of source-goal motion events has shown, though, that the core elements of an event, such as the goal or endpoint of motion, and the peripheral elements of an event, such as the source or starting point of motion, are influenced differently by the pragmatic pressures associated with the needs of listeners. The decision to mention peripheral event components do appear to be susceptible to factors relating to audience design (Do et al., 2020,

2022). But conceptually core elements do not appear to be – at least, not in motion events involving Animate figures in motion.

The aim of this work was to further investigate whether the decision to mention conceptually core elements in a motion event can ultimately be influenced by the needs of addressees or whether – as previously advanced by Do et al. (2020) – they are "largely impervious to communicative status" (Do et al., 2020, p. 9). Moreover, given that prior experiments have diverged in terms of their findings for the significance of the animacy of the figure for simple motion events, a second aim of this work was to see how conceptual factors such as the animacy of the figure within a motion event influences how the goal is represented in memory (Experiment 1) and language (Experiment 2).

4.1 Pragmatics and Goals

The present set of results paint an interesting picture for the role of audience design in source-goal motion events. Prior results have shown that a specific aspect of audience design, informativity, has distinct effects on how the source and the goal are represented in language production (Do et al, 2020, 2022). In order to better understand the effect of informativity, a stronger set of cues to the informativity of the goal were included in the present procedure. In spite of this, we found that decisions about whether to mention or omit the goal were not affected at all by informativity. Nonetheless, our results strongly imply that our participants were still responding to the pragmatic status of the goal. We observed speakers making explicit references to shared visual information, and, more importantly, participants reliably used definite/indefinite determiners with goals in a manner consistent with the communicative status or givenness of the goal. It appears that goals are not completely impervious to pragmatic influences, instead pragmatic cues may change how the goal is articulated or used by speakers.

Why might participants choose to elaborate on the goal despite the fact that it is uninformative? Speakers may continue to mention the goal even when the endpoint of the motion is already known to their interlocutor (Goal in Common Ground Condition), because doing so serves a purpose beyond informing the addressee about the physical or spatial endpoint reached by the figure in a motion event. While it is not yet clear what that purpose may be, our data point to several possible reasons why a speaker may continue to mention even pragmatically "uninformative" goals.

One possibility suggested by Do et al. (2020) is that the goal is the endpoint of a movement in space and it conveys the telicity or boundedness of an event (Do et al., 2020). An event with no goal (eg. "The squid swims away."), for instance, does not necessarily have an end-boundary. However, an event with a goal communicates an event that ends once the figure has reached the goal (eg. "The squid swims to the coral."). If the goal is serving this dual communicative role, then, speakers may have been unwilling to drop goals from their speech because they communicate important temporal information about the event.

A related possibility, highlighted by the speaker's explicit references to their addressee's visual access, is that mentioning the goal serves a synchronization function in discourse that may benefit the conversation in two ways. Some support for this possibility comes from the types of sentences produced when speakers explicitly referenced information that was visually accessible to their interlocutor – sentences like "The squid went from there to where it is now." Although these types of sentences did not constitute a majority of our data, the presence of phrases like "to where it is now" nevertheless suggest that speakers may choose to mention the goal as a means of establishing a commonly understood timeline between speaker and addressee. If so, then goal

mention may serve a turn-taking function – a conversational "baton pass" indicating not the end of the event *per se*, but the end of what speakers have to say about the event.

Likewise, utterances in which the endpoint of the motion is mentioned prior to the starting point of the motion ("The squid that you see by that seaweed$_{GOAL}$, it started by the coral$_{SOURCE}$."), suggest that a separate benefit of using the goal to establish a commonly understood endpoint in time may be to allow the speaker to work backwards from that endpoint in their description of the event. Notably, speakers in prior work (Do et al., 2020) rarely ever described motion events using the goal-source or goal-only word orders (approximately 1% of their utterances were these goal-first orders), goal-first utterances comprised roughly 12% of all utterances in the present study. Importantly, the vast majority of these goal-source utterances, as in the example above, co-occurred with explicit references to the listener's visual access: utterances that explicitly referenced addressee knowledge showed a much higher rate of Goal-Source word order (35.71%) as compared to the rate of this order in the data overall (4.59%).

While additional work is necessary to further investigate these latter accounts of goal mention, they do point to an aspect of audience design that goal representations participate in more actively. Goal mentions in these cases do not necessarily fit under the umbrella of informativity, but do fit within the broader concept of pragmatic adjustments to that speakers make for the benefit of their listeners.

4.2 The Role of Animacy in Memory & Language for Source Goal Motion Events

Although some studies (Lakusta & Landau, 2012; Lakusta & Carey, 2015) have found goals to be conceptually prominent only in motion events involving animate but not inanimate figures in motion, other studies (Do et al., 2022; Papafragou, 2010) have found evidence for conceptual prominence in both Animate and Inanimate figures. Consequently, the influence of

animacy has been an open question in the literature for source goal motion events and has complicated attempts to understand the relationship between conceptual and linguistic representations of events.

In the present study, we tested an important design difference between works by Lakusta and colleagues and research from Do and colleagues that, among other things, may have contributed to these differing sets of results – namely, the manipulation of Animacy Type as part of a within-subject or a between-subject design. In contrast with Do et al. (2022), which used a within-subjects procedure, we used a between-subjects design in Experiment 1 and found, in line with Lakusta and colleagues, that participants were significantly more likely to accurately recall goals of animate motion events than they were to accurately recall goals of inanimate motion events. These results suggest that when participants have the opportunity to view both animate and inanimate motion events within the same series of trials this changes how they reason about the prominence relationships between inanimate objects and animate entities.

Importantly, these findings allow us to resolve many of the apparent contradictions in prior work. In work by Lakusta and colleagues (Lakusta & Landau, 2012; Lakusta & Carey, 2015), animacy was manipulated between-subjects, so participants would not have been presented with the opportunity to make comparisons between animate and inanimate motion events. As a result, the goals of motion within inanimate events were treated as fully distinct from animate ones. The opposite is true for the Do et al. (2022) study. There, the within-subjects design allowed participants to directly compare events involving animate and inanimate figures. At a methodological level, then, one implication of our work may be that future studies examining effects of animacy, particularly in the context of source-goal motion events, may wish to take careful consideration in the design and/or blocking of experimental trials.

At a theoretical level, our results also bear on open questions concerning the relationship between a speaker's conceptual and linguistic representations of an event. Prior work by Lakusta and Landau (2012) noted an interesting asymmetry between the goals of animate versus inanimate events: while it was possible to motivate the bias to mention goals over sources in animate events via the relative prominence of the goal versus the source in event cognition (i.e., in memory), it was not possible to do so in the context of inanimate events, where there was no evidence that goals were conceptually more prominent than sources. Although we did not directly compare memory for goals versus sources, the results presented here essentially replicate the pattern observed by Lakusta and Landau (2012). As indicated by the lower rates of accuracy in the memory task, goals of inanimate motion events were treated as conceptually less prominent than the goals of animate events. Yet, this difference in conceptual prominence did not appear to feed forward robustly into people's language for motion events: in Experiment 2, participants still chose to mention the goal at approximately the same rate even though Experiment 1 showed that the presence of an inanimate figure should have rendered the goal less prominent conceptually and by extension, less important to talk about.

If there is a relative absence of conceptual weight behind the goals of inanimate motion events, as found in the present study and by Lakusta & Landau (2012), why are the goals of inanimate motion events still mentioned at such a high rate? Our results provide some tentative support for the pragmatic account of this asymmetry initially posited by Do et al. (2020) – namely, that goals of inanimate events continue to be mentioned at high rates, not because they are conceptually important, but because they are important for pragmatic reasons. Indeed, in line what may be expected by this pragmatic account, we were able to change the way that speakers talked about the goal in inanimate events by manipulating the pragmatic status of the goal. Moreover, our

results showed that these pragmatic factors could affect language for the goal in inanimate events in exactly the same way as it affected animates.

At the same time, though, there is some tentative evidence that animacy did influence language for other aspects of the event – beyond the motion event, itself. In particular, we found that participants tended to describe the backdrop and 'set the scene' more frequently in the Inanimate conditions than in the Animate conditions (Appendix D) both in the Goal in Common Ground (four times in Animate and 21 times in Inanimate conditions) and the No Common Ground conditions (111 observations in the Animate conditions, 170 observations in the Inanimate conditions). These results suggest that speakers may have reasoned differently about what else in the scene was important to mention in animate versus inanimate conditions. Moreover, this choice to describe the scene occurring behind the motion event reflects the participant's judgement about what is newsworthy or requiring elaboration in the trial under discussion. In inanimate conditions, the inanimate figure seems to be treated as less distinct from the rest of the scene and participants appear to spend more time describing the entire scene rather than fixating on the motion event. As this was not the central aim of the present work, though, it is not possible to comment extensively on the qualitative differences between the elicited descriptions of the motion events. An area for future research may be trying to assess, in a more controlled fashion, whether having an animate or inanimate participate in an event will change how speakers prioritize and construe elements beyond the foregrounded aspects of event.

4.3 The Conceptual & Linguistic Representation of (In)Animates

More broadly, our findings have implications for the study of judgements around animacy and inanimacy even beyond the domain of source-goal motion events. Importantly, our findings showed that it was easier to promote inanimate figures to a status comparable to animate figures

than it is to demote an animate figure. Indeed, it is not difficult to imagine the evolutionary or adaptive justifications for why this is the case. For example, the ability to quickly reassess the status of an entity that initially appears inanimate, but is in fact dangerous, moving, or living, could be very beneficial (Nairne et al., 2017). Further, the costs of being unable to rapidly shift and incorporate new information about animate or moving entities in one's immediate environment could be very high.

At the same time, while the animate/inanimate distinction has typically been treated as categorical, a growing number of experiments have not only begun identifying the ways in which animacy/inanimacy binary is incomplete, but have also suggested that animacy may best be viewed as a dimensional hierarchy in which features of animacy and prominence can be accrued. An easy way to see this middle ground is to consider a nearly philosophical question: Is fire animate or inanimate? An open flame has features of both inanimacy (eg. non-volitional) and animacy (e.g. it moves under its own power and consumes material). In the lab, experiments using cartoon-like contexts to impart animate or anthropomorphic traits to inanimate objects (eg. a peanut that sings and falls in love), for instance, have shown that participants very quickly adjust their expectations to accommodate this situation (Nieuwland & Van Berkum, 2006; Rich et al., 2022): basically, in the presence of cues that an inanimate object should be treated more like an animate, it often will be.

These findings accord with similar results, including ours, from the domain of source-goal motion events. Our findings – in particular, the fact that construals of (in)animacy can depend simply on contextual factors like viewing animates and inanimate events in close proximity – offer additional support for a model of animacy in which animacy is not a strict binary defined by non-salient inanimate objects and salient animate agents. Likewise, work by Papafragou (2010) using

inanimate, but self-propelled figures in motion (e.g., a soccer ball moving with a self-determined trajectory) showed that proximity may not be the only substrate by which the goal of an inanimate motion event can increase in prominence. In their study, goals of inanimate motion events were construed as conceptually salient because self-propulsion was a very strong cue to the intentionality and by extension, animacy of the figure in motion (Gergely & Csibra, 2003; Biro et al., 2007; Gao et al., 2012).

Taken together, then, these studies highlight the ways in which the conceptual representations used for inanimacy are highly contextually sensitive in a way that our conceptualizations of animacy are not. Essentially, inanimates appear to be more easily shifted upward or downward on the animacy hierarchy, whereas the position of animate entities is more stable.

Given this, an important question, then, is why the animate figure has such a striking effect on the way in which inanimate figures are generally construed when the two are in proximity. On this question, there are at least three logical origins or etiologies that may contribute to this finding. This could be a process that points to the contagiousness associated with animate entities (Nairne et al., 2017, Pandeirada et al., 2022), the degree of intentionality conveyed by the inanimate and animate figures relative to one another (Heider & Simmel, 1944; Gao et al., 2012), or to a process in which animate representations can serve in a templating role for the representations used for inanimate figures (Lakusta & Carey, 2015). Theorists have suggested that animate entities may be inherently more memorable. By most measures living things cluster high on the animacy hierarchy, and this pattern (or bias) in prior findings is sometimes labeled the Animacy effect (Pandeirada et al., 2022). Human cognitive systems seem to be tuned to notice and maintain information about animate entities over inanimate objects (Pandeirada et al., 2022). But experiments from outside

the literature on motion events lend support to the perhaps stranger notion that this sense of priority or prominence associated with animacy is a little contagious. The conceptual importance granted by animacy can extend to objects that are associated with an animate entity through abstract relationships, like ownership, or – as in the case of source-goal motion events investigated here – through more concrete channels like physical contact (Nairne et al., 2017). This idea has two related applications within simple motion events. If the prominence associated with animate figures is contagious, then this could have an effect on how the highly context sensitive inanimate representations are construed within motion events. This is particularly true when animate entities and inanimate objects are in close physical or temporal proximity. Second, the goal of an animate motion event may in some cases be viewed as more prominent because the "contagious" animate figure has made physical contact with the inanimate ground object located at the endpoint of motion.

It is also possible that the issue is not entirely about the contagiousness of animacy and may relate to other features that usually differ between animate and inanimate figures. One of the best illustrations of the human tendency to grant a wide range of animate traits to inanimate objects based on less than extensive evidence comes from work by Heider & Simmel (1944) in which participants viewed 2-D geometric shapes in motion. In animated displays in which a triangle moves along a simple path, appearing to "follow" two other geometric shapes, participants will readily describe the triangle mentalistically, as chasing the other shapes. This involves granting a very abstract and minimal type of object not only animacy but a kind of intentionality, as only an intentional actor can "chase" rather than just being moved (Gao et al., 2012; Scholl & Gao, 2013). The feature of intentionality may also be a viable avenue for exploring the differences observed between events with animate and inanimate figures. Events involving animate agents may be

viewed as conceptually important, in part, because animate agents are intentional. In the real world, animacy and intentionality tend to co-occur, it is very often animate entities that move in a way that conveys intentionality (Tremoulet & Feldman, 2000; Scholl & Gao, 2013). Yet, when this is not the case, as with the triangles, humans are rather susceptible to perceiving intentionality in motion. In a cartoon or animated setting, if participants view animate agents and inanimate objects that move in essentially similar fashion across the screen, the figures may convey a similar degree of intentionality. This may lead participants to grant both groups a similar status within a motion event. An area for future research may be to attempt to systematically change the conveyed intentionality of figures in motion events.

A final candidate explanation for the subordinate relationship in which animate representations influence the construal of inanimate source-goal motion events concerns a theory about the structural relationships between animate and inanimate conceptual representations beginning in development. Lakusta & Carey (2015) proposed that during childhood the conceptual representations for animate or agentive motion events serve as templates for the linguistic structure used for motion events. This more core conceptual structure is also used during development to build the linguistic structure used for motion events with inanimate figures and underlies the conceptual structure associated with inanimate motion events. This reasoning implies that because the same patterns of processing activated for animate motion events underlie the processing of inanimate events, once they are activated inanimate events will be treated more similarly to animate motion events. This templatic relationship also speaks to why it is that when animate and inanimate motion events are viewed close to one another in time, it is more likely to be the inanimate motion events that conform to the animate template. As a consequence, the within-subjects procedure for memory may have permitted participants to use the animate stimuli as a

template that was subsequently applied to the inanimate stimuli. This opportunity would not have been presented in the current study. This may be another possible source of the differing rates of accuracy in memory for endpoints of animate and inanimate motion events.

Our data do not allow us to distinguish between these possibilities. Nor do they allow us to understand which aspect of proximity ultimately encourages people to treat inanimates like animates. It is not clear, for example, whether this is because proximity invites participants to make direct comparisons between animate agents and objects in motion or if this is really a result of temporal proximity. In other words, seeing animate and inanimate stimuli within a short timespan could lead the viewer to represent animate agents and inanimate objects more similarly. An area for future research may be to see if these two possibilities can be properly partitioned. It may be possible to present animate and inanimate stimuli close together temporally in a way that does not invite comparison. However, this may be challenging with source-goal motion events that, at some level, have a common, simple underlying form. Another strategy might be to use stimuli that can be readily compared but to manipulate the time elapsed between viewing blocks of animate and inanimate stimuli to see what the temporal limitations of this effect might be. We leave these issues to be further explored in further studies.

5. Conclusion

We focused on a single prominent element (the goal) of simple source-goal motion events in order to determine how conceptual and pragmatic factors can influence the status of the goal in memory (Experiment 1) and language (Experiment 2), respectively. In Experiment 1, we manipulated the Animacy of the Figure in motion and found that animate figures contribute to the conceptual prominence of the goal, whereas inanimate figures did not. Taken with prior work, our results point to a model of animacy that is not categorical, but may instead be multi-dimensional and heavily dependent on contextual factors. In Experiment 2, we additionally manipulated the pragmatic status of the goal by varying whether the goal was inside or outside common ground. We found that unlike for conceptually peripheral event components, such as sources, factors relating to audience design do not fundamentally determine whether conceptually core event elements such as goals are included during production– they almost always are. Instead, pragmatic factors related to audience design influence how speakers strategize and use the goal to convey relevant information. We discuss the implications of our work for theories of event cognition and language production.

Acknowledgments

I would like to thank Dr. Monica Do for her guidance throughout the experiment design, editing, and revision process. I would like to thank the undergraduate research assistants affiliated with the Language, Events, and Mind (LEM) lab—Sophia Rodriguez-Bell, Ryan Hathaway, Amelia Orwant, Charlotte Gilmore, and Lourdes Benecia-Courreges who were adept confederates and provided valuable assistance in the transcribing and coding process. I am also grateful to, Ruben Rodriguez-Barron, who provide useful feedback during the revision process.

References

Abbott, B. (2006). Definiteness and Indefiniteness. In L. R. Horn & G. Ward (Eds.), *The Handbook of Pragmatics* (pp. 122–149). Blackwell Publishing Ltd. https://doi.org/10.1002/9780470756959.ch6

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using **lme4**. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Biro, S., Csibra, G., & Gergely, G. (2007). The role of behavioral cues in understanding goal-directed actions in infancy. In *Progress in Brain Research* (Vol. 164, pp. 303–322). Elsevier. https://doi.org/10.1016/S0079-6123(07)64017-5

Brennan, S. E., Galati, A., & Kuhlen, A. K. (2010). Two Minds, One Dialog. In *Psychology of Learning and Motivation* (Vol. 53, pp. 301–344). Elsevier. https://doi.org/10.1016/S0079-7421(10)53008-1

Brown, P. M., & Dell, G. S. (1987). Adapting production to comprehension: The explicit mention of instruments. *Cognitive Psychology*, *19*(4), 441–472. https://doi.org/10.1016/0010-0285(87)90015-6

Clark, H. H., & Krych, M. A. (2004). Speaking while monitoring addressees for understanding. *Journal of Memory and Language*, *50*(1), 62–81. https://doi.org/10.1016/j.jml.2003.08.004

Clark, H. H., & Marshall, C. R. (1981). *Definite Reference and Mutual Knowledge*. Cambridge University Press.

Clark, H. H., & Murphy, G. L. (1982). Audience Design in Meaning and Reference. In *Advances in Psychology* (Vol. 9, pp. 287–299). Elsevier. https://doi.org/10.1016/S0166-4115(09)60059-5

Do, M. L., & Kaiser, E. (2021). Sentence formulation is easier when thematic and syntactic prominence align: Evidence from psych verbs. *Language, Cognition and Neuroscience*, 1–23. https://doi.org/10.1080/23273798.2021.2008458

Do, M. L., Papafragou, A., & Trueswell, J. (2020). Cognitive and pragmatic factors in language production: Evidence from source-goal motion events. *Cognition*, *205*, 104447. https://doi.org/10.1016/j.cognition.2020.104447

Do, M. L., Papafragou, A., & Trueswell, J. (2022). Encoding Motion Events During Language Production: Effects of Audience Design and Conceptual Salience. *Cognitive Science*, *46*(1). https://doi.org/10.1111/cogs.13077

Fukumura, K. (2016). Development of audience design in children with and without ASD. *Developmental Psychology*, *52*(1), 71–87. https://doi.org/10.1037/dev0000064

Galati, A., & Brennan, S. E. (2010). Attenuating information in spoken communication: For the speaker, or for the addressee? *Journal of Memory and Language*, *62*(1), 35–51. https://doi.org/10.1016/j.jml.2009.09.002

Gao, T., Scholl, B. J., & McCarthy, G. (2012). Dissociating the Detection of Intentionality from Animacy in the Right Posterior Superior Temporal Sulcus. *Journal of Neuroscience*, *32*(41), 14276–14280. https://doi.org/10.1523/JNEUROSCI.0562-12.2012

Gee, J. P. (1974). Jackendoff's Thematic Hierarchy Condition and the Passive Construction. *Linguistic Inquiry*, *5*(2), 304–308.

Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naïve theory of rational action. *Trends in Cognitive Sciences*, *7*(7), 287–292. https://doi.org/10.1016/S1364-6613(03)00128-1

Goldin-Meadow, S. (2003). *the resilience of language: What gesture creation in deaf children can tell us about how all children learn language*. Psychology Press.

Grice, H. P. (2002). *Logic and conversation.*

Grigoroglou, M., & Papafragou, A. (2019a). Interactive contexts increase informativeness in children's referential communication. *Developmental Psychology*, *55*(5), 951–966. https://doi.org/10.1037/dev0000693

Grigoroglou, M., & Papafragou, A. (2019b). Children's (and Adults') Production Adjustments to Generic and Particular Listener Needs. *Cognitive Science*, *43*(10). https://doi.org/10.1111/cogs.12790

Hafri, A., Papafragou, A., & Trueswell, J. C. (2013). Getting the gist of events: Recognition of two-participant actions from brief displays. *Journal of Experimental Psychology: General*, *142*(3), 880–905. https://doi.org/10.1037/a0030045

Heider, F., & Simmel, M. (1944). An Experimental Study of Apparent Behavior. *The American Journal of Psychology*, *57*(2), 243–259. https://doi.org/10.2307/1416950

Hinz, A., Michalski, D., Schwarz, R., & Herzberg, P. Y. (2007). The acquiescence effect in responding to a questionnaire. *GMS Psycho-Social Medicine*, *4*, Doc07.

Horton, W. S., & Keysar, B. (1996). When do speakers take into account common ground? *Cognition*, *59*(1), 91–117. https://doi.org/10.1016/0010-0277(96)81418-1

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.), *Current Issues in Linguistic Theory* (Vol. 178, p. 123). John Benjamins Publishing Company. https://doi.org/10.1075/cilt.178.09iha

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*(7), 826–840. https://doi.org/10.1080/23273798.2019.1584323

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*(2), 152–175. https://doi.org/10.1080/15475441.2014.896168

Lakusta, L., & DiFabrizio, S. (2017). And, the Winner Is…A Visual Preference for Endpoints over Starting Points in Infants' Motion Event Representations. *Infancy*, *22*(3), 323–343. https://doi.org/10.1111/infa.12153

Lakusta, L., & Landau, B. (2005). Starting at the end: The importance of goals in spatial language. *Cognition*, *96*(1), 1–33. https://doi.org/10.1016/j.cognition.2004.03.009

Lakusta, L., & Landau, B. (2012). Language and Memory for Motion Events: Origins of the Asymmetry Between Source and Goal Paths. *Cognitive Science*, *36*(3), 517–544. https://doi.org/10.1111/j.1551-6709.2011.01220.x

Lakusta, L., Muentener, P., Petrillo, L., Mullanaphy, N., & Muniz, L. (2017). Does Making Something Move Matter? Representations of Goals and Sources in Motion Events With Causal Sources. *Cognitive Science*, *41*(3), 814–826. https://doi.org/10.1111/cogs.12376

Levine, D., Hirsh-Pasek, K., Pace, A., & Michnick Golinkoff, R. (2017). A goal bias in action: The boundaries adults perceive in events align with sites of actor intent. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(6), 916–927. https://doi.org/10.1037/xlm0000364

Lockridge, C. B., & Brennan, S. E. (2002). Addressees' needs influence speakers' early syntactic choices. *Psychonomic Bulletin & Review*, *9*(3), 550–557. https://doi.org/10.3758/BF03196312

Maguire, M. J., Brumberg, J., Ennis, M., & Shipley, T. F. (2011). Similarities in Object and Event Segmentation: A Geometric Approach to Event Path Segmentation. *Spatial Cognition & Computation*, *11*(3), 254–279. https://doi.org/10.1080/13875868.2011.566955

Matuschek, H., Kliegl, R., 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. https://doi.org/10.1016/j.jml.2017.01.001

Nadig, A. S., & Sedivy, J. C. (2002). Evidence of Perspective-Taking Constraints in Children's On-Line Reference Resolution. *Psychological Science*, *13*(4), 329–336. https://doi.org/10.1111/j.0956-7976.2002.00460.x

Nairne, J. S., VanArsdall, J. E., & Cogdill, M. (2017). Remembering the Living: Episodic Memory Is Tuned to Animacy. *Current Directions in Psychological Science*, *26*(1), 22–27. https://doi.org/10.1177/0963721416667711

Nieuwland, M. S., & Van Berkum, J. J. A. (2006). When Peanuts Fall in Love: N400 Evidence for the Power of Discourse. *Journal of Cognitive Neuroscience*, *18*(7), 1098–1111. https://doi.org/10.1162/jocn.2006.18.7.1098

Pandeirada, J. N. S., Félix, S. B., & Nairne, J. S. (2022). Animacy. *Encyclopedia of Animal Cognition and Behavior*, 268–272. https://doi.org/10.1007/978-3-319-55065-7_2069

Papafragou, A. (2010). Source-Goal Asymmetries in Motion Representation: Implications for Language Production and Comprehension. *Cognitive Science*, *34*(6), 1064–1092. https://doi.org/10.1111/j.1551-6709.2010.01107.x

Papafragou, A., & Grigoroglou, M. (2019). The role of conceptualization during language production: Evidence from event encoding. *Language, Cognition and Neuroscience*, *34*(9), 1117–1128. https://doi.org/10.1080/23273798.2019.1589540

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), 1–25. https://doi.org/10.1080/01638539809545035

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, *27*(02). https://doi.org/10.1017/S0140525X04000056

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, *36*(4), 329–347. https://doi.org/10.1017/S0140525X12001495

Regier, T., & Zheng, M. (2007). Attention to Endpoints: A Cross-Linguistic Constraint on Spatial Meaning. *Cognitive Science*, *31*(4), 705–719. https://doi.org/10.1080/15326900701399954

Rich, S., Balachandran, L., Duff, J., Kogan, M., Cavallaro, M., Handel, N., & Wagers, M. (2022). *"It's alive!": Animacy-based structural expectations rapidly update with context*. https://virtual.oxfordabstracts.com/#/event/public/2562/submission/278

Rissman, L., & Majid, A. (2019). Thematic roles: Core knowledge or linguistic construct? *Psychonomic Bulletin & Review*, *26*(6), 1850–1869. https://doi.org/10.3758/s13423-019-01634-5

Scholl, B. J., & Gao, T. (2013). Perceiving Animacy and Intentionality. In M. D. Rutherford & V. A. Kuhlmeier (Eds.), *Social Perception* (pp. 197–230). The MIT Press. https://doi.org/10.7551/mitpress/9780262019279.003.0009

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

Talmy, L. (n.d.). *Lexicalization patterns: Semantic structure in lexical forms*. 47.

Tremoulet, P. D., & Feldman, J. (2000). Perception of Animacy from the Motion of a Single Object. *Perception*, *29*(8), 943–951. https://doi.org/10.1068/p3101

Ward, N. (2006). Non-lexical conversational sounds in American English. *Pragmatics & Cognition*, *14*(1), 129–182. https://doi.org/10.1075/pc.14.1.08war

Wilson, D., & Sperber, D. (1981). *On Grice's theory of conversation* (pp. 155–178).

Woodward, A. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, *69*(1), 1–34. https://doi.org/10.1016/S0010-0277(98)00058-4

Zheng, M., & Goldin-Meadow, S. (2002). Thought before language: How deaf and hearing children express motion events across cultures. *Cognition*, *85*(2), 145–175. https://doi.org/10.1016/S0010-0277(02)00105-1

Appendix A

Goal phrases	Types	n	Proportion
Particle+Prepositional	away-part	2	0.0018416
Particle+Prepositional	off-part	3	0.0027624
Particle+Prepositional	on-part	1	0.0009208
Particle+Prepositional	other-part	3	0.0027624
Particle+Prepositional	over-part	84	0.0773481
Particle+Prepositional	to-part	19	0.0174954
Prepositional Phrase	at-pp	12	0.0110497
Prepositional Phrase	behind-pp	39	0.0359116
Prepositional Phrase	by-pp	18	0.0165746
Prepositional Phrase	in-front-of-pp	38	0.0349908
Prepositional Phrase	in-pp	139	0.1279926
Prepositional Phrase	near-pp	1	0.0009208
Prepositional Phrase	of-pp	44	0.0405157
Prepositional Phrase	on-pp	157	0.1445672
Prepositional Phrase	other-pp	23	0.0211786
Prepositional Phrase	over-part	1	0.0009208
Prepositional Phrase	to-pp	449	0.4134438
Prepositional Phrase	towards-pp	43	0.0395948
Verb Phrase	v-np	10	0.0092081

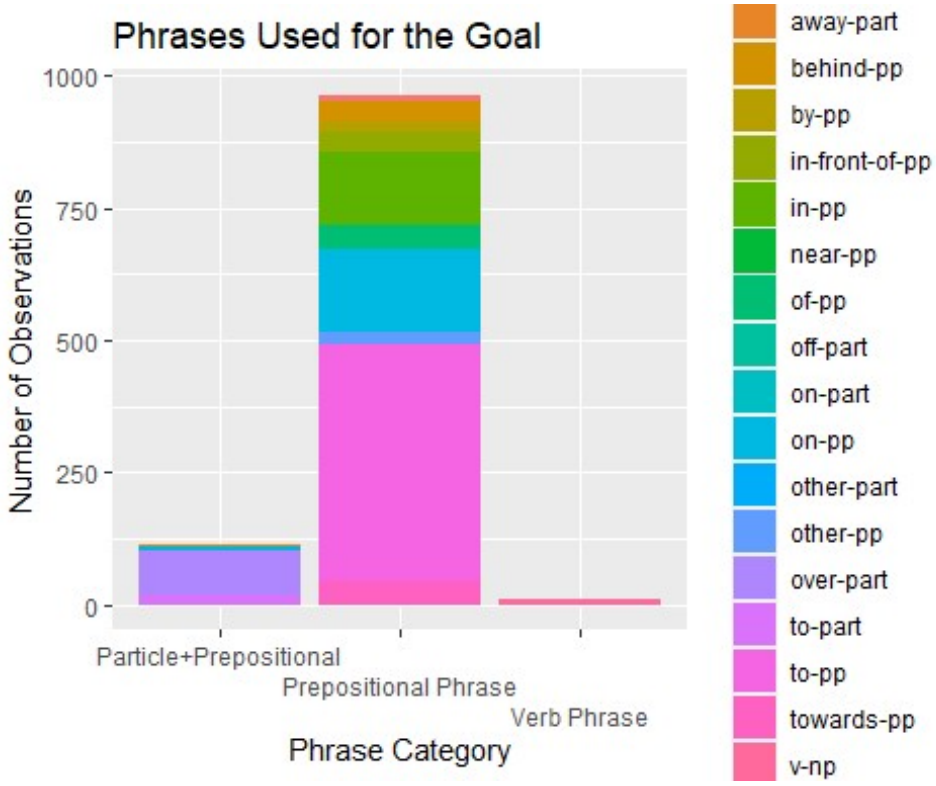

Figure A1. *Frequency of use for different syntactic frames for Goals in Experiment 1b. Detailed proportions for each type of phrase are in the table above.*

For both sources and goals, mentions were most frequently observed within prepositional phrases (eg. "The squid swam from the coral$_{SOURCE}$ to the seaweed$_{GOAL}$"). Mentions were also observed in numerically fewer cases as part of a particle and prepositional phrase structure (e.g., "The squid swam from behind the coral$_{SOURCE}$ over to the seaweed$_{GOAL}$") or in a few cases a verb and noun phrase structure (e.g., "The squid found the seaweed$_{GOAL}$").

Appendix B

Source Phrase	Types	n	Proportion
Particle+Prepositional	away-part	1	0.0010604
Particle+Prepositional	from-part	51	0.0540827
Particle+Prepositional	off-part	25	0.0265111
Particle+Prepositional	other-part	3	0.0031813
Particle+Prepositional	out-part	62	0.0657476
Particle+Prepositional	over-part	9	0.0095440
Prepositional Phrase	at-pp	25	0.0265111
Prepositional Phrase	behind-pp	42	0.0445387
Prepositional Phrase	by-pp	21	0.0222694
Prepositional Phrase	from-pp	434	0.4602333
Prepositional Phrase	in-front-of-pp	11	0.0116649
Prepositional Phrase	in-pp	36	0.0381760
Prepositional Phrase	near-pp	14	0.0148462
Prepositional Phrase	of-pp	35	0.0371156
Prepositional Phrase	off-pp	11	0.0116649
Prepositional Phrase	on-pp	127	0.1346766
Prepositional Phrase	other-pp	19	0.0201485
Prepositional Phrase	out-pp	5	0.0053022
Prepositional Phrase	over-part	1	0.0010604
Prepositional Phrase	to-pp	3	0.0031813
Prepositional Phrase	under-pp	1	0.0010604
Verb Phrase	v-np	7	0.0074231

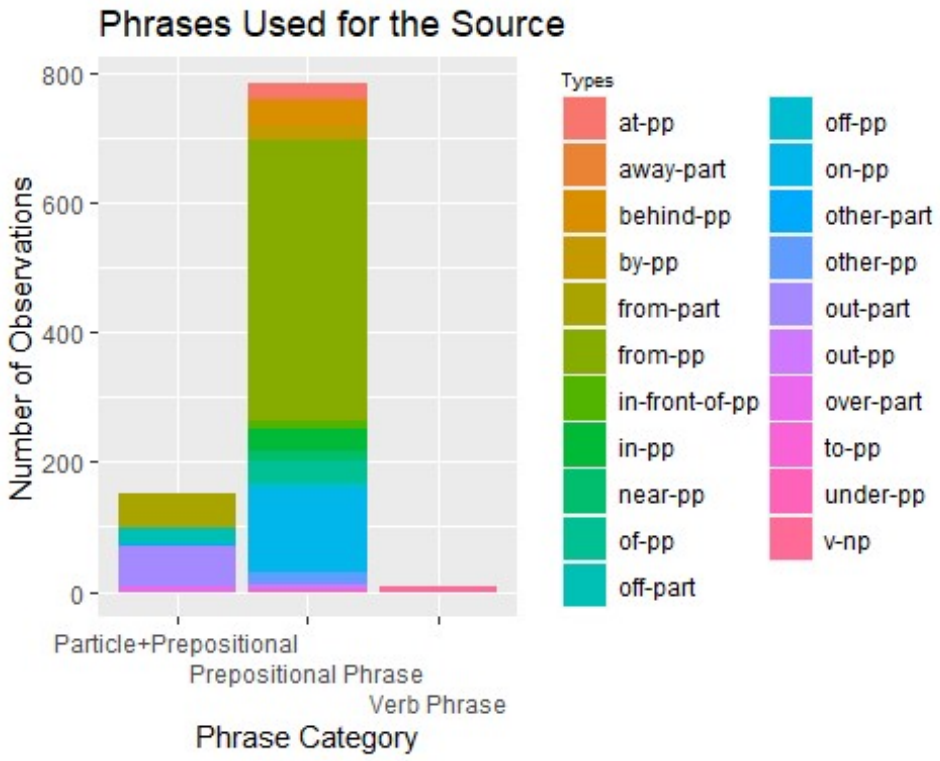

Figure B1. *Frequency of use for different syntactic frames for Sources in Experiment 1b. Detailed proportions for each type of phrase are in the table above.*

For both sources and goals, mentions were most frequently observed within prepositional phrases (eg. "The squid swam from the coral$_{SOURCE}$ to the seaweed$_{GOAL}$"). Mentions were also observed in numerically fewer cases as part of a particle and prepositional phrase structure (e.g., "The squid swam from behind the coral$_{SOURCE}$ over to the seaweed$_{GOAL}$") or in a few cases a verb and noun phrase structure (e.g., "The squid left the coral$_{SOURCE}$").

Appendix C

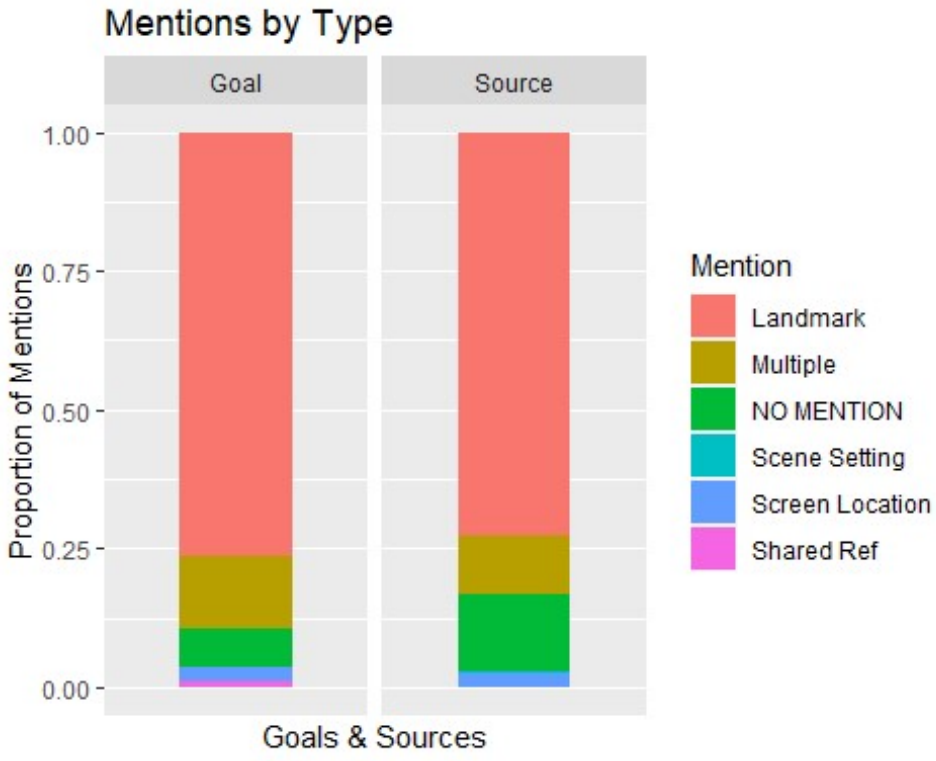

Figure C1. *Proportions of Source and Goal Mentions by categories used in coding for the mention types. Landmarks were physical locations. Screen Locations were uses like "the right of the screen" or "the bottom left". Shared Ref included cases where the primary mention of the Goal was an Explicit Reference to Shared Visual Access. In Multiple more than one type of Goal Mention was included.*

Goal Mention	n	Proportion	Example
Landmark	783	0.7646484	"To the seaweed"
Multiple	133	0.1298828	"To the seaweed that is located in the lower left"
NO MENTION	73	0.0712891	"A squid swam away"
Screen Location	25	0.0244141	"To the lower left"
Shared Ref	10	0.0097656	"To where it is now"

Source Mention	n	Proportion	Example
Landmark	745	0.7275391	"From the coral"
Multiple	108	0.1054688	"From the coral that is located at the upper right"
NO MENTION	143	0.1396484	"A squid swam away"
Scene Setting	2	0.0019531	"We're under the sea, there's some coral"
Screen Location	26	0.0253906	"From the upper right"

Appendix D

Scene Setting

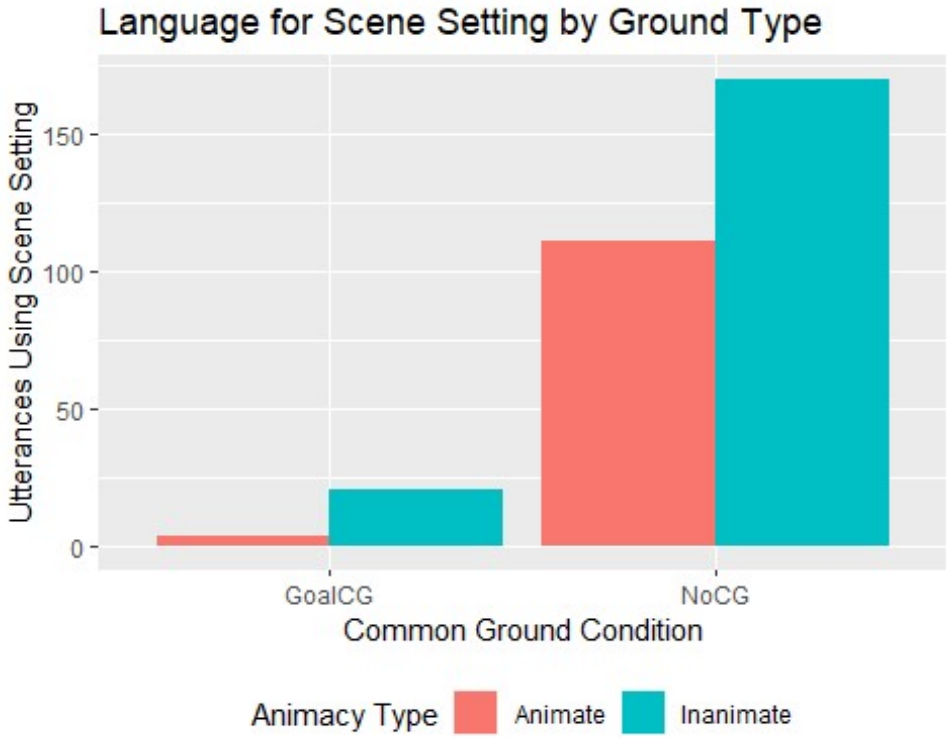

Figure D1. *The number of utterances that used Scene Setting divided by Common Ground condition and Animacy Type. Scene setting was observed in approximately 30% of utterances overall.*