HHS Public Access

Author manuscript

Lang Cogn Neurosci. Author manuscript; available in PMC 2020 October 02.

Published in final edited form as:

Lang Cogn Neurosci. 2019 ; 34(3): 273–288. doi:10.1080/23273798.2018.1525495.

Occluding the face diminishes the conceptual accessibility of an animate agent

Lilia Rissman[a], **Amanda Woodward**[b], **Susan Goldin-Meadow**[b,c]

[a]Center for Language Studies, Radboud University, Erasmusplein 1, Nijmegen, the Netherlands
[b]Department of Psychology, University of Chicago, 5848 South University Ave., Chicago, IL 60637
[c]Department of Psychology, University of Chicago, Center for Gesture, Sign, and Language, University of Chicago

Abstract

The language that people use to describe events reflects their perspective on the event. This linguistic encoding is influenced by *conceptual accessibility*, particularly whether individuals in the event are animate or agentive—animates are more likely than inanimates to appear as Subject of a sentence, and agents are more likely than patients to appear as Subject. We tested whether *perceptual* aspects of a scene can override these two conceptual biases when they are aligned: whether a visually prominent *inanimate patient* will be selected as Subject when pitted against a visually backgrounded *animate agent*. We manipulated visual prominence by contrasting scenes in which the face/torso/hand of the agent were visible vs. scenes in which only the hand was visible. Events with only a hand were more often associated with passive descriptions, in both production and comprehension tasks. These results highlight the power of visual prominence to guide how people conceptualize events.

Keywords

event representation; animacy; sentence production; passive voice

1 Introduction

When people view an event, different perspectives on the event are often available, and these perspectives are reflected in the linguistic descriptions that people choose. For example, an event may be described as *the sleek gray cat tipped over the plastic cup of water*, or simply *she tipped the cup over,* or even *the cup was tipped over.* These choices are influenced by a variety of conceptual, linguistic, and contextual factors, such as the animacy of the individuals in the event, or whether the person describing the event heard a linguistic prime beforehand (see Bock, Irwin & Davidson, 2004; Bock & Ferreira, 2014; Ferreira & Slevc, 2007, for review). The Subject of a sentence is a crucial indicator of the perspective taken on

Correspondence concerning this article should be directed to Lilia Rissman, Center for Language Studies, Radboud University, Erasmusplein 1, Nijmegen, the Netherlands. l.rissman@let.ru.nl. +1 773-206-3954. Other authors' contact information: Amanda Woodward, woodward@uchicago.edu, (773) 702-8829. Susan Goldin-Meadow, sgsg@uchicago.edu. (773) 702-2585.

an event. For example, when describing a transitive event with an agent and a patient, participants can either produce an active sentence, with the agent as Subject (e.g., *the cat tipped over the water*), or a passive sentence, with the patient as Subject (e.g., *the water was tipped over by the cat*).[1] Extensive research has confirmed that conceptual properties of an event (e.g., whether the agent is animate) can influence choice of Subject. However, we know less about the impact that lower-level visual properties of events (e.g., whether an animate agent is visible) have on event perspective and Subject choice. Here we ask whether lower-level visual properties of how an event is displayed can override conceptual properties of the event, thus shaping the dominant perspective taken on the event.

1.1 Conceptual accessibility

Sentence production is an incremental process that proceeds through multiple phases, including message formation, grammatical function assignment, lexical retrieval, structural assembly, and phonological spell-out (Bock, 1982; Bock & Ferreira, 2014; Levelt, 1989). In English and many other languages, the Subject of a sentence is usually also the topic (Givón, 1983; Lambrecht, 1994). Thus active and passive sentences differ not only in their Subject, but also in their topic; that is, the "emphasis" of the sentence (Bernolet, Hartsuiker & Pickering, 2009), or what the sentence is primarily "about" (Brown & Yule, 1993; Lambrecht, 1994). Bock and Warren (1985) characterize this difference in terms of *conceptual accessibility:* "conceptual accessibility is the ease with which the mental representation of some potential referent can be activated in or retrieved from memory.… accessible concepts being those that are in some sense most 'thinkable'" (Bock & Warren, 1985: 50). Assuming an incremental model of language production, entities that are more conceptually accessible are easier to mentally activate and are therefore more likely to surface as Subject.

A variety of factors have been shown to influence conceptual accessibility. One of the most robust and well documented of these is *animacy* (Bock, 1986; Bock, Loebell & Morey, 1992; Branigan, Pickering & Tanaka, 2008; Clark & Begun, 1971; Corrigan, 1986; Itagaki & Prideaux, 1985; McDonald, Bock & Kelly, 1993; Prat-Sala & Branigan, 2000; see also Christianson & Ferreira, 2005). In McDonald et al. (1993), for example, participants had more success recalling active sentences with animate agents (e.g., *a farmer purchased a refrigerator*) than active sentences with inanimate agents (e.g., *the sound frightened the students*). And in Prat-Sala & Branigan (2000), speakers were more likely to produce passive descriptions when the patient of an event was animate than when the patient was inanimate (i.e., *the man was hit by the swing* was more common than *the scooter was hit by the swing*). Animacy also has a strong influence on grammatical structure, affecting a range of cross-linguistic patterns, including word order, differential object marking, and person systems (Aissen, 2003; Comrie, 1979; Croft, 1988; Yamamoto, 1999).

Agency is another factor that has been shown to influence conceptual accessibility.[2] Agents are more conceptually accessible than patients (Gleitman, January, Nappa & Trueswell, 2007; Prat-Sala & Branigan, 2000; Van Nice & Dietrich, 2003; Vogels, Krahmer & Maes,

[1]We take the term "agent" to roughly indicate an entity who initiates an event, playing a causal role, and "patient" to indicate an entity who is affected by an agent (see Jackendoff, 1990; Dowty, 1991).

2013a). For example, in a norming study conducted by Gleitman et al. (2007), when speakers described scenes in which animacy was controlled (i.e., both the agent and patient were animate), they overwhelmingly produced active sentences, placing the agent in Subject position on 86% of trials. The influence of agency is also well documented with respect to other aspects of grammatical structure, with widespread cross-linguistic effects of agency on word order, case-marking, and agreement (Comrie, 1989; Dowty, 1991; Primus, 1999).

Animacy and agency are both conceptual properties of events: of the individual entities and their roles in the event, respectively. Conceptual properties are not alone, however, in influencing conceptual accessibility: the linguistic and non-linguistic context of an event, as well as perceptual features of the event, can also play a role. For example, when participants describe pictures with two entities, if they have seen one of the entities beforehand, they are more likely to mention this entity as Subject (Antón-Méndez, 2017; Prentice, 1967; Osgood, 1971; Sridhar, 1988), perhaps because it has become familiar or Given in the discourse. Prat-Sala and Branigan (2000) find a similar effect by manipulating linguistic discourse: hearing a narrative that puts linguistic focus on one of the entities in a picture increases the likelihood that the focused entity will be mentioned as Subject in a subsequent description task. Crucially, Prat-Sala and Branigan found that this discourse effect overrode the influence of animacy: when an inanimate agent was highlighted by previous linguistic discourse, it was mentioned as Subject more often than an animate patient. Prat-Sala and Branigan thus argue that “inherent” influences, such as animacy, and “derived” influences, such as linguistic context, have an additive effect on conceptual accessibility: “context can make an inherently accessible entity still more accessible; but given a sufficiently strong context, an inherently inaccessible entity may be temporarily more accessible than in inherently accessible entity” (179).

In this theory of conceptual accessibility, none of the influences on accessibility (inherent properties like animacy, role properties like agency, discourse factors like whether an entity is Given, or perceptual properties like the size of an entity) has special priority. If the purely additive proposal of Prat-Sala and Branigan (2000) is correct, then conceptual factors such as animacy and agency should not play a privileged role in shaping conceptual accessibility. This framework thus predicts that factors such as the linguistic context and perceptual features of an event should, in principle, be able to override *both* the animacy and agency of the event in determining which entity is most conceptually accessible.[3]

Such a result, however, would come as a surprise given previous theoretical and empirical research. An event in which an animate agent causes a change in an inanimate patient constitutes a prototypical transitive event, a pattern with high regularity across the languages of the world (Dowty, 1991; Hopper & Thompson, 1980; Tsunoda, 1985). As such, animacy and agency, when combined, are assumed to have a super-additive effect on conceptual

[2]According to Bock & Warren (1985), “conceptual role differences may themselves be ordered in terms of accessibility, with roles such as that of the agent being more accessible than that of the patient” (64).

[3]Prat-Sala & Branigan (2000) explicitly discuss inherent properties of referents, such as animacy and concreteness, as well as derived properties, such as linguistic and non-linguistic context. We additionally consider in this paper the effects of thematic role on conceptual accessibility (what Van Nice & Dietrich, 2003, call “role accessibility”), as well as effects of visual prominence (see Vogels, Krahmer & Maes, 2013b).

accessibility, overriding more transient influences, such as whether a picture of the inanimate patient was previously viewed. In Bock et al. (1992), participants viewing events with an animate agent and an inanimate patient (e.g., a person carrying a boat) *never* produced passive descriptions such as *the boat was carried by the person*. These data suggest that when conceptual properties such as animacy and agency are aligned, perceptual and discourse factors are not likely to have an effect on conceptual accessibility. Indeed, we are aware of few studies within this psycholinguistic literature that explore the impact of animacy *and* agency on conceptual accessibility; that is, few studies that pit animate agents against inanimate patients (but see Altmann & Kemper, 2006). Given the increasing number of studies over the past 10 years demonstrating effects of visual perception on sentence production, we ask here whether visual prominence can override animacy and agency when they are aligned.

1.2 Visual prominence and language production

A wealth of research demonstrates that visual perceptual factors affect Subject selection (Antón-Méndez, 2017; Baltaretu, Krahmer, van Wijk & Maes, 2016; Coco & Keller, 2009; Coco & Keller, 2015; Flores d'Arcais, 1975; Forrest, 1996; Gleitman et al., 2007; Hwang & Kaiser, 2015; Kuchinsky, 2009; Myachykov, Thompson, Garrod & Scheepers, 2012; Myachykov, Garrod & Scheepers, 2012; Myachykov, Garrod & Scheepers, 2018; Osgood, 1971; Osgood & Bock, 1977; Sridhar, 1988; Tomlin, 1997; Vogels et al., 2013a). For example, in Gleitman et al. (2007), participants viewed illustrations with two animate entities, such as a cat licking a dog. Just prior to viewing the illustration, a square flashed briefly over the space where one of the entities would ultimately appear. When the square had flashed over the patient, participants were more likely to produce a passive description (e.g., *the dog was licked by the cat*) than when the square had flashed over the agent.

What is the mechanism by which more visually prominent referents are more likely to surface as Subject? Gleitman et al. (2007) argue that the dominant mechanism is lexical access: that guiding the eyes to one of the entities, in this case, the dog, facilitates access of the word *dog*, which has the effect of promoting *dog* to the beginning of the sentence. Myachykov et al. (2012a) propose a direct mapping from attentional focus to the Subject position, based on the finding that viewing a preview picture has the same effect as the attention capture manipulation used by Gleitman et al. (2007).

An alternate explanation is that visual prominence affects Subject selection by influencing conceptual accessibility, specifically, how an event is construed (Antón-Méndez, 2017; Bock & Ferreira, 2014; Kuchinsky, 2009; Vogels et al., 2013a). Any single event in the world can be construed in multiple ways; for example, an event of one army running away from another can be described as either chasing or fleeing (DeLancey, 1991; Fisher, Hall, Rakowitz & Gleitman, 1994). This notion of construal is similar to the notion of topic introduced in Section 1.1. Kuchinsky (2009) showed events that had a straightforward construal, as well as more semantically ambiguous events. She found that the attention capture manipulation used by Gleitman et al. (2007) only affected Subject selection for the ambiguous events (i.e., events with more than one likely construal), suggesting that the visual manipulation of cuing eye gaze directed participants toward a particular construal. If

cuing eye gaze only affected lexical access, then this facilitation should have been the same for both straightforward and ambiguous events (see Bock & Ferreira, 2014, for discussion). Gleitman et al. themselves acknowledge the possible role of event construal, which they describe in terms of figure-ground assignment: “in our dog-chasing-man scene, either the dog or the man can be viewed as the Figure, and the surrounding information will thus serve as the background” (565). The current study contributes to the debate on whether visual prominence influences conceptual accessibility, specifically, whether visual prominence influences the perspective taken on an event.

1.3 Research approach

When people view events, a variety of factors determine which referent is most conceptually accessible. We ask here whether information about visual prominence can override animacy and agency when they are aligned, i.e. whether making an animate agent less visually prominent decreases its conceptual accessibility relative to an inanimate patient. We use Subject choice as a measure of conceptual accessibility; that is, which entity is mentioned as Subject in a language production task, and how sentences with different Subjects are interpreted in a language comprehension task. The majority of studies demonstrating perceptual effects on sentence formulation involve two entities that are matched for animacy (Antón-Méndez, 2017; Hwang & Kaiser, 2015; Gleitman et al., 2007; Myachykov et al., 2012a; Myachykov et al., 2012b; Myachykov et al., 2018; Tomlin, 1997; Vogels et al., 2013a; see also Griffin & Bock, 2000), perhaps reflecting a tacit assumption that perceptual influences will only reveal themselves when animacy is controlled and thus not a factor. To address our question, we selected an overt manipulation of visual prominence: partial occlusion of one of the referents. Specifically, we contrasted events where the face, torso and hands of an animate agent were visible (Body-Agent events, thus focusing attention on the agent) against events where only the hand and forearm of the agent were visible (Hand-Agent events, thus taking attention away from the agent and focusing more on the patient).

We tested whether occluding the face and torso of an animate agent leads this agent to be less conceptually accessible, using both a sentence production task (Experiments 1–2) and a sentence comprehension task (Experiment 3). In Experiments 1–2, participants described videos of events and we measured whether the animate agent or the inanimate patient was mentioned as Subject. In Experiment 3, participants read a sentence and had to choose which video (Body-Agent or Hand-Agent) best matched the sentence; we measured whether their choice of video differed depending on whether the animate agent or the inanimate patient was the Subject. The majority of previous studies investigating conceptual accessibility have focused on language production data. Nonetheless, data about how people comprehend and interpret sentences are also relevant to this question, and can provide converging evidence for an effect of visual prominence on conceptual accessibility.

A variety of evidence predicts that this perceptual manipulation (Body-Agent vs. Hand-Agent) should have *no* effect. Animacy and agency are two of the strongest, if not the strongest, predictors of Subject choice that have been observed. A person’s hand is a strong cue to both animacy and agency: infants as young as 9 months understand that hands are connected to animate agents (Saxe, Tzelnic & Carey, 2007; Slaughter & Heron-Delaney,

2010; Woodward, 1998), and we use our hands to explore and manipulate the world, as well as to communicate through gesture and sign.

However, there is also empirical reason to believe that visual prominence could override animacy. Altmann and Kemper (2006) asked adults to formulate sentences after viewing displays with three written words—one animate noun, like *butler*; one inanimate noun, like *juice*; and one verb, like *stirred*. The three words were shown in a vertical list, and the authors manipulated whether the inanimate noun was on the top or the bottom of the list. When the inanimate noun was on the top of the list, participants were 13% more likely to produce a passive sentence than when it was on the bottom of the list. A perceptual manipulation can thus mitigate the strong effects of animacy and agency. But this study was restricted to written words—participants' behavior might reasonably be different when viewing actual events, particularly given the visual salience of people (Fletcher-Watson, Findlay, Leekam, & Benson, 2008; Kirchner & Thorpe, 2006; New, Cosmides, & Tooby, 2007).

In Experiments 1–3, we manipulated the visual prominence of an animate agent (Body-Agent vs. Hand-Agent). In Experiment 4, we conducted a validation task to assess whether participants rate the agent as having the same level of animacy in the Body-Agent vs. Hand-Agent conditions; participants rated the animacy of entities in events and we measured whether animacy ratings differed depending on whether an agent's face was visible or not.

2 Experiment 1

In Experiment 1, we assessed how speakers describe Body-Agent events relative to Hand-Agent events, using three different designs (Experiment 1A, Experiment 1B, Experiment 1C).

2.1 Participants

English-speaking adults were tested on Amazon Mechanical Turk (Experiment 1A: N = 52, F = 17, age range = 21 – 57, mean age = 33; Experiment 1B: N = 16, F = 8, age range = 20 – 55, mean age = 33; Experiment 1C: N = 16, F = 5, age range = 20 – 50, mean age = 33). Participants self-reported being native speakers of American English, where native speaker was defined as someone who lived in the United States prior to age 13, and whose parents spoke English to them during that time. Participants were also located in the United States. An additional two participants were tested but excluded for not being native English speakers. For completing the study, participants received $1 (1A), $2.50 (1B), and $2.50 (1C).

2.2 Design & Materials

Experiment 1 used three types of videos: (1) In Body-Agent videos, a person acted on an object (e.g., a woman tipped over a book), and the person's face, torso and hands were visible throughout the video. (2) In Hand-Agent videos, a person's hand entered the video frame from off screen, acted on the object, and then exited the frame. (3) As a control condition, participants also described No Agent videos in which the object acted on its own (e.g., a book fell over), with no person visible in the video. We elicited descriptions of these

No Agent videos as a baseline measure of how people describe events that have no animate agent at all, and we expected intransitive descriptions, e.g. *the book fell over*. Figure 1 shows still images from each of these three conditions.

Experiment 1A had a between-subjects design: 18 people described Body-Agent videos, 18 people described Hand-Agent videos, and 16 people described No Agent videos. Experiment 1B had a within-subject design in which participants described videos from each of the Body-Agent, Hand-Agent, and No Agent conditions in a randomized order. Experiment 1C had a blocked within-subjects design in which participants viewed Hand-Agent videos, then Body-Agent videos, then No Agent videos. These three designs serve as an opportunity for replication and also test whether participants only distinguish Body-Agent events from Hand-Agent events when they are observed contrastively, as in a within-subjects design.

There were five different items (featuring three different agents) across the three conditions: (1) a woman tipping over a book/a book falling, (2) a man shutting a jewelry box/a jewelry box shutting, (3) a woman opening a dresser door/a dresser door swinging open, (4) a woman pushing a tube down into water/a tube sinking in water, and (5) a man pushing a marker down a ramp/a marker rolling down a ramp.

In addition to the test items, each participant also described filler videos. In half of the fillers, a person performed a repeated activity, such as jumping or spinning in place (Activity fillers); in the other half, an inanimate object was undergoing an action with no obvious start or endpoint, such as a flag waving or water running in a faucet (Object fillers). The purpose of these fillers was to increase the variety in the stimuli, and to encourage participants to produce sentences with an animate object as Subject (Activity fillers, e.g. *a man was jumping*) or with an inanimate object as Subject (Object fillers, e.g. *the flag was waving*). The fillers were presented randomly throughout the test items; there were roughly equal numbers of fillers and test items across Experiments 1A-C.

2.3 Procedure

For each video, participants typed a written description of what they saw happening. Participants saw each video once, and were told that their descriptions could be as long or short as they liked. No time limit was imposed on the descriptions, nor any restrictions on the form of the description (e.g., a full sentence, a sentence with a verb, etc.). Participants could make changes to their descriptions within a single trial as they were typing, but could not revisit their responses from previous trials.

2.4 Coding

Two research assistants coded each description as either “active,” “passive,” “intransitive,” or “other.” Active descriptions included an agent as Subject and a patient as Object, as in *a man pushed a pen down a slope*, or *a hand spinning a tire on a rope*. In passive descriptions, the patient is the Subject, the verb is in past participle form, and there is a passive auxiliary—BE, as in *a small tire swing was spun around*; GET, as in *some kind of tube getting put in water*; or HAVE, as in *a makeshift ramp has a marker rolled down it*.[4] We also coded whether passive descriptions included an agentive *by*-phrase, as in *a small chest was shut by a man*. In intransitive descriptions, the patient is the Subject and there is no passive

morphology, as in *the marker rolled down the cardboard* or *book falls down.* Descriptions coded as Other include statives such as *rotating tire on string* and reflexives such as *a box closed itself*. To establish reliability, the first author coded a random 15% of the descriptions. Agreement was 94% (1A), 97% (1B), and 100% (1C).

2.5 Results

The proportion of trials in which participants produced active, passive, and intransitive descriptions for each of the Body-Agent, Hand-Agent, and No Agent videos is shown in Figure 2.

As expected, participants almost always produced transitive descriptions (active or passive) in the Body-Agent and Hand-Agent conditions, and intransitive descriptions in the No Agent condition. The Body-Agent and Hand-Agent conditions were not identical, however. In all three experiments, participants produced more passive descriptions in Hand-Agent than in Body-Agent conditions—in Experiment 1A, Hand-Agent = 25% (SE = 5%) and Body-Agent = 4% (SE = 2%); in Experiment 1B, Hand-Agent = 18% (SE = 4%) and Body-Agent = 1% (SE = 1%), and in Experiment 1C, Hand-Agent = 29% (SE = 5%) and Body-Agent = 0%. To assess whether these differences were statistically significant, we modeled the data using logistic regression, where the dependent variable was whether participants produced a passive description. In Experiments 1A-B, we fit regression models using *glm* for R (R Core Team, 2017), where the possible predictor was Agent Type (Body-Agent vs. Hand-Agent). There were significantly more passive descriptions for Hand-Agent than Body-Agent scenes in both Experiment 1A ($\beta = 2.00$, SE = .57, $p < .001$) and in Experiment 1B ($\beta = 2.82$, SE = 1.05, $p < .01$). In Experiment 1C, participants never produced passive descriptions in the Body-Agent condition. Because *glm* models do not provide reliable estimates when the values of one of the levels of the dependent variable have no variance, we modeled these data using the *logistf* package for R, which uses Firth's penalized likelihood logistic regression (Firth, 1993). Again, Agent Type was a significant predictor ($\beta = 4.19$, SE = 1.45, $p < .001$). 86% of the passive sentences produced in Experiment 1 were short passives without a *by*-phrase, e.g., *the box was opened*. This strong preference for short passives is consistent with previous corpus studies on English passive (Jespersen, 1992 [1924]; Stein, 1979; Thompson, 1987). Taken together, these results demonstrate that visually deemphasizing the agent can override the bias for animate agents to be mentioned in Subject position.

3 Experiment 2

Experiment 1 demonstrated that English speakers are more likely to use passive voice syntax when describing Hand-Agent scenes than when describing Body-Agent scenes. There were multiple perceptual differences between the Body-Agent and Hand-Agent scenes—in Hand-Agent scenes, the agent's face/torso were absent and the frame was more "zoomed in" on the patient. In addition, in Body-Agent scenes, the agent was visible for the entire duration

[4]Sentences lacking a passive auxiliary were still coded as Passive if the verb was incompatible with a non-agentive event, e.g. *book knocked over* or *pipe dunked under water*. Of the 101 Passive descriptions across Experiments 1–2, these were the only two exemplars lacking a passive auxiliary.

of the event, whereas in Hand-Agent scenes, the hand entered from off screen, performed the action, and then exited. In Experiment 2, we asked whether the absence of the face/torso would alone override the conceptual bias to mention animate agents in Subject position. We asked speakers to describe a new set of stimuli in which the sole difference between Body-Agent and Hand-Agent scenes was the presence or absence of the face/torso.

3.1 Participants

32 native English-speaking adults participated in Experiment 2 via Amazon Mechanical Turk, receiving $1.75 (F = 14, age range = 19 – 54, mean age = 29). These participants had not completed Experiments 1A-1C, as determined by Worker ID.

3.2 Design, Materials & Procedure

In a between-subjects design, participants viewed videos from one of two conditions: (1) Body-Agent or (2) Hand-Agent. In the Body-Agent videos, a person rests his or her hand on a table, then acts on an inanimate object, then returns the hand to the resting position. Hand-Agent videos are identical, except that a screen occludes the person's body, revealing only the hand and forearm. There were seven items (performed by two different women actors): tipping over a book, pushing a ball down a ramp into a cup, closing a box, pushing a paint roller down a ramp, rolling a cup across a table, grabbing a hacky sack off of a box, and opening a box. Figure 3 includes still images from one of these items. In both conditions, participants described seven Activity and seven Object filler videos (as described in Experiment 1), which were randomly interspersed among the Body-Agent and Hand-Agent videos. We chose a between-subjects design given that we observed an effect for all three designs in Experiment 1, and a between-subject design avoids possible concerns about interactions between the conditions. Procedure and data coding were the same as for Experiment 1. For Experiment 2, intercoder reliability was 98%.

3.3 Results

The results of Experiment 2 are shown in Figure 4. We again observed a contrast between the two conditions. Participants produced passive descriptions on 20% of Hand-Agent trials (SE = 4%) but only 5% of Body-Agent trials (SE = 2%). We modeled these data as in Experiments 1A-B. Participants were significantly more likely to produce passive descriptions in the Hand-Agent condition than in the Body-Agent condition ($\beta = 1.48$, SE = .48, $p < .01$). Here, 100% of the passive descriptions were short passives.

3.4 Discussion

Experiments 1–2 show that a perceptual manipulation of occluding most of the agent's body can override the strong bias for animate agents to be mentioned in Subject position. Does this result show that visually occluding part of the agent leads the agent to have lower conceptual accessibility? We argue that the effects in Experiments 1–2 reflect conceptual differences in event construal, i.e., figure-ground assignment. If Hand-Agent events are construed as more patient-oriented, that is, as having a more topical patient, this focus would also lead to more passive descriptions.

However, an alternate (non-mutually exclusive) explanation is that the increased use of passive voice in the Hand-Agent condition reflects lexical processing. As described in Section 1.2, Gleitman et al. (2007) argue that covertly directing participants' eye gaze to the patient leads participants to retrieve the lexical item for that patient more quickly than they would have without the covert direction, resulting in more descriptions with the patient in Subject position; that is, more passive descriptions. And in fact, in an eyetracking study, Rissman, Goldin-Meadow and Woodward (2018) found that participants viewing Hand-Agent scenes directed their initial fixation to the patient more often than participants viewing Body-Agent scenes did. Thus for Hand-Agent scenes in Experiments 1–2, the lexical item for the patient might be retrieved more rapidly than the lexical item for the agent, leading to more passive voice descriptions. Similarly, occluding the face in the Hand-Agent condition could have reduced the number of lexical options available for describing the agent. Although participants in the Hand-Agent condition produced a variety of labels for the agent in their active descriptions, including *someone*, *a person* and *a hand*, they infrequently produced gendered labels such as *man* and *woman*. Thus the fact that the gender of the agent was more ambiguous in the Hand-Agent condition could have increased the time needed to access a lexical item for the agent.

To understand whether participants conceptualize Body-Agent videos differently from Hand-Agent videos, and whether these differences are associated with the linguistically prominent Subject position, we sought converging evidence through a language comprehension task. In Experiment 3, we conducted a matching task in which participants viewed Body-Agent and Hand-Agent scenes side-by-side, accompanied by a single sentence description. This sentence was either active or passive (agent is Subject vs. patient is Subject), and participants' task was to select which video best matched the sentence. If Body-Agent and Hand-Agent videos differ in the conceptual prominence of the agent, then participants should match the passive sentence to a Hand-Agent video and the active sentence to a Body-Agent video. If, alternatively, the agent in Hand-Agent events is as conceptually prominent as the agent in Body-Agent events, participants' selection of one video over the other should not be predicted by the type of sentence they are given (active vs. passive). Lexical retrieval is not a factor in this task, as participants were provided the sentence descriptions.

4 Experiment 3

In Experiment 3, we assessed whether participants associate Body-Agent events with active sentences and Hand-Agent events with passive sentences. We used two different designs to test the robustness of any observed effect (Experiment 3A, Experiment 3B).

4.1 Participants

Native English-speaking adults participated in Experiment 3 via Amazon Mechanical Turk (Experiment 3A: N = 22, F = 12, age range = 26 – 59, mean age = 37; Experiment 3B: N = 63, F = 29, age range = 24 – 59, mean age = 37). These participants had not completed Experiments 1–2, as determined by Worker ID. An additional eight participants were tested, but were excluded for several reasons: inaccuracy on control trials (N = 3), being a non-

native speaker of English (N = 4), or having completed a previous experiment (N = 1). Participants received $2.75 (3A) and $3 (3B).

4.2 Design, Materials & Procedure

Participants viewed two videos positioned on the left- and right-hand sides of the screen. Beneath the videos appeared a single sentence. Participants' task was to choose which video best matched the sentence. In the crucial trials, a Body-Agent video was pitted against a Hand-Agent video, and we asked whether participants made a different selection when they were given an active sentence rather than a passive sentence.

Experiment 3A featured a within-subjects design, where participants viewed trials from each of the eight conditions in Table 1. Conditions #1–2 are the crucial test conditions and conditions #3–8 serve as controls—will participants correctly select Body-Agent and Hand-Agent videos over No Agent videos when the sentence describes a transitive event (either active or passive syntax), and will they correctly select No Agent videos when the sentence is intransitive? No Agent videos were perceptually matched to the Agent videos: in No Agent [Body] videos, a person sat without moving while an object underwent a change (e.g., a book falling over). In No Agent [Hand] videos, the person was occluded behind a screen but the hand was visible throughout the duration of the non-agentive event.

In Experiment 3A, each participant saw four trials in each condition, featuring four different items. Each participant thus saw 32 trials, which were presented in a pseudorandom order. In Experiment 3B, Sentence Type (active vs. passive) was a between-subjects variable and intransitive sentences were not tested. Given the absence of intransitive sentences, participants only saw trials from Conditions #1–6. Each participant saw eight Body-Agent vs. Hand-Agent trials, eight Body-Agent vs. No Agent [Body] trials, and eight Hand-Agent vs. No Agent [Hand] trials. These 24 trials were presented in random order. In both studies, the appearance of Video 1 and Video 2 on the left vs. right-hand sides of the screen was counterbalanced. The particular stimulus items were drawn from Experiment 2.

We selected short passive sentences for this task, as the vast majority of the passive descriptions produced in Experiments 1–2 were short passives. All sentences featured gerund syntax, e.g., *a box being closed,* because sentences with a past participle (e.g., *a box was closed*) are ambiguous between a passive reading, in which someone closed the box, and a stative reading, in which the box was in a closed-state.[5] Active sentences contained an indefinite subject, e.g., *someone rolling a ball into a cup,* to avoid the issue that the gender of the agent is relatively ambiguous in Hand-Agent videos. See Appendix for full list of sentences.

Participants were told that, in each trial, they would see an English sentence accompanied by two videos, and that their task was "to choose the video that best matches the English sentence."

[5]We used gerunds (e.g., *a box being closed*) rather than past tense sentences (e.g., *a box was closed*) because, on their stative reading, past tense sentences are equally felicitous at the end of Body-Agent events as at the end of Hand-Agent events, as both of these events end with a box in a closed-state. Although clausal gerunds lack tense and agreement, they do license overt subjects, which get assigned case, the syntactic property most relevant to the active vs. passive distinction (Pires, 2006).

4.3 Results

Participants were excluded from analysis if they scored less than 87% correct on the control conditions (Experiment 3A ~ 21/24; Experiment 3B ~ 14/16 trials). Three participants failed to meet this criterion (3A: N = 2; 3B: N = 1). For the remaining participants, the mean scores on the control conditions were 96% (3A) and 98% (3B). Turning to the crucial test conditions, Figure 5 shows the proportion of Body-Agent vs. Hand-Agent trials where participants chose the Body-Agent video, for both active and passive sentences.

In Body-Agent vs. Hand-Agent trials, participants may have had a baseline bias to choose one type of video over the other (e.g., participants may prefer a video with a person's face, all else being equal). For this reason, we tested whether rates of selecting the Body-Agent video were different for active vs. passive sentences, rather than testing whether selection of the Body-Agent video differed from chance in each sentence condition. The data in Figure 5 show that participants were more likely to choose the Body-Agent video when they saw an active sentence than when they saw a passive sentence in both Experiments 3A and 3B. We modeled the likelihood of choosing the Body-Agent video using mixed-effects logistic regression and the *lme4* package for R (Bates & Maechler, 2009). For Experiment 3A, the best-fitting model included Subject as a random effect and Sentence Type (active vs. passive) as a fixed effect, but not Item as a random effect. In this model, participants chose the Body-Agent video more often when the sentence was active than when it was passive ($\beta = 2.09$, SE = .42, $p < .001$). For Experiment 3B, the best-fitting model also included Subject as a random effect and Sentence Type (active vs. passive) as a fixed effect, but not Item as a random effect. Again, participants chose the Body-Agent video more often when the sentence was active than when it was passive ($\beta = 3.15$, SE = .92, $p < .001$).

4.4 Discussion

We found in Experiment 3 that participants were more likely to select a Body-Agent video given an active sentence than a passive sentence. These results suggest that participants conceptually distinguish Body- and Hand-Agent events: that the latter are more patient-oriented than the former. This interpretation supports the findings of Kuchinsky (2009), Vogels et al. (2013a), Bock and Ferreira (2014) and Antón-Méndez (2017), who argue that differences in visual prominence affect not only lexical retrieval but also event construal (i.e., figure-ground assignment), and thereby conceptual accessibility. We suggest that in Experiment 3, when participants were given an agent-oriented syntactic form (an active sentence), they were more likely to choose a Body-Agent video because these videos highlight the role of the agent more than Hand-Agent videos do (i.e., the agent is more conceptually prominent).

An alternate explanation for our findings in Experiment 3 is that the indefinite label we used for the agent, *someone,* biased participants to choose the Body-Agent videos (in the same manner that a label such as *the woman* would introduce a bias for Body-Agent videos). Indefinite descriptions, however, are used precisely when the identity of a referent is under-specified: for the Hand-Agent descriptions in Experiments 1–2, *someone* was the Subject in 33% of active descriptions. For Body-Agent videos, by contrast, *someone* was never the

Subject of an active description. It is therefore unlikely that the lexical content of *someone* biased participants in Experiment 3 to choose Body-Agent videos.

Taken together, Experiments 1–3 indicate that partially occluding an animate agent decreases its conceptual accessibility. However, we cannot rule out the possibility that the sentence production results from Experiments 1–2 reflect differences in lexical accessibility, and it is of course possible that the differing descriptions for Body-Agent vs. Hand-Agent scenes were driven by differences in *both* lexical and conceptual accessibility. But the fact that we found similar results in Study 3, a comprehension task where lexical access was not a factor, as the sentences were given to the participants, supports the interpretation that people construe Body-Agent and Hand-Agent events in different ways. Although the relation between representations underlying production vs. comprehension is a topic of long-standing debate, few researchers argue that these representations are strictly separate (see Kitterage & Dell, 2016; Meyer, Huettig & Levelt, 2016, for review). More importantly for the current study, conceptual representations of events serve as input to both language production and language comprehension, and production findings (here, the likelihood of mentioning the agent vs. patient as Subject) and comprehension findings (here, different interpretations for agent vs. patient Subjects) both reflect how people conceptualize events.

5 Experiment 4

This final experiment serves to validate the methodology used in Experiments 1–3. Our goal in this paper is to understand whether visual prominence can have such a strong effect on conceptual accessibility that it can override animacy and agency when they are aligned. Importantly, we do not know whether the Body-Agent videos and Hand-Agent videos used in our studies are, in fact, matched on the two relevant conceptual dimensions, animacy and agency. Participants describing Hand-Agent videos almost never produced intransitive descriptions such as *the book fell over*, indicating that participants do represent the hand as an agent. But the face of a person is a strong cue to animacy (Johnson, Slaughter & Carey, 1998; Kuhlmeier, Wynn & Bloom, 2003), raising the possibility that participants conceptualize the agent in Body-Agent videos as *more animate* than the agent in Hand-Agent videos. We tested this possible explanation in Experiment 4 using an animacy rating task.

5.1 Participants

36 Native English-speaking adults participated in Experiment 4 via Amazon Mechanical Turk (F = 14, age range = 21 – 58, mean age = 35). These participants had not completed Experiments 1–3. An additional twelve participants were tested, but were excluded either due to inaccuracy on control trials (N = 10) or for being a non-native speaker of English (N = 2). Participants received $2.

5.2 Design, Materials & Procedure

Participants viewed live-action videos and rated the animacy of one of the entities in each video. Videos were drawn from eight different conditions, shown in Table 2.

The Body and Hand trials were the crucial test conditions; the stimuli for these conditions were the same as in the Body-Agent and Hand-Agent conditions in Experiments 2 & 3, with the person occluded behind a screen in the Hand condition. Activity and Object trials served as control conditions, as these featured unambiguously animate and unambiguously inanimate entities, respectively. The remaining conditions served to vary the stimuli and to encourage participants to use the full range of the scale. The Activity, Object, and No Agent videos were the same as in Experiment 1. In all videos, the target entity was moving, so that participants would not conflate being static with being inanimate.

We tested eight videos per condition, with the exception that the No Agent condition contained only four videos. For the seven conditions where eight videos were tested, each participant viewed four videos selected at random from the group of eight. Each participant viewed all four of the No Agent videos. Thus, each participant viewed 32 videos (four videos in each of eight conditions), presented in a random order.

In each trial, the video began with a one-second still frame, with a colored dot placed over one of the entities. In the Body trials, for example, the colored dot was placed over the person's chest. The video then began playing and the colored dot disappeared. Each video lasted about four seconds and was followed by another one-second still frame, with the colored dot appearing again over the same entity. Participants were instructed to rate the animacy of the entity picked out by the colored dot on a scale from 1 to 7, where 1 corresponded to "not animate at all" and 7 corresponded to "highly animate." Following work by Dahl (2008) and Dahl and Fraurud (1996), we defined animacy for participants using the following text: "There is no single definition of animacy, but animacy is associated with: entities that are **alive**, entities that have **minds**, and entities that can **act** (i.e., perform actions in the world)." Previous studies have elicited judgments about animacy by asking participants to rate the extent to which an entity is "alive" (Tremoulet & Feldman, 2000), or is "animate" (Radanović, Westbury & Milin, 2016). We selected the instructions above in order to elicit a notion of animacy that distinguishes humans from plants (which are both alive), and because we were not confident that participants would have strong notions about the meaning of "animate" without any additional instruction.

5.3 Results

Participants needed to rate the Activity trials at least 2 points higher than Object trials to be included in the analysis. 10 participants failed to meet this criterion. The remaining 36 participants ranked Activity trials 4.7 points higher on average than Object trials. Figure 6 shows the average animacy rating for each of the eight conditions in Table 2.

We tested for a main effect of condition using Friedman's ANOVA, as the rating data was ordinal rather than interval. Animacy ratings were not the same across the eight conditions (Friedman's $\chi^2(8) = 255.7$, $p < .001$). We conducted post hoc tests to assess differences between conditions using the *friedmanmc* function and *pgirmess* package for R. Table 3 shows the difference value for each pairwise comparison, where 64.9 was the critical difference value for $\alpha = .05$, with Bonferroni correction for multiple comparisons. Significant comparisons are indicated by a shaded cell.

Table 3 shows that the animacy ratings for the Body and Hand videos were not statistically different from the animacy ratings for the Activity and Animal videos, and were not statistically different from each other. All four of these event types elicited high animacy judgments (6.2 – 6.6 on a scale from 1 to 7).

5.4 Discussion

We found no evidence in Experiment 4 that participants conceptualize a human agent as being more animate when the face of the human is visible than when the face is hidden behind a screen. It is possible that our measure was not sensitive enough to capture a true difference in representation of animacy between the Body and Hand conditions, or that our instructions biased participants to report that Body and Hand scenes are more similar than they actually are. However, we did find that participants were willing to rate other entities as having intermediary animacy status, namely plants and machines. In addition, the animacy ratings for the Body and Hand conditions were nearly identical and did not differ statistically from the ratings for Activity videos, which was our high-animacy control condition. These findings therefore moderate concerns that the agents in the Body-Agent and Hand-Agent videos from Experiments 1–3 do not share the same animacy status.

Humans are highly prone to interpret the behavior of a seemingly inanimate entity as animate given sufficient context, for example, by simply having eyes (Hamlin, Wynn & Bloom, 2007; Lowder & Gordon, 2015; Nieuwland & Van Berkum, 2006). The reverse, conceptualizing an animate entity as inanimate, that is, as not having the potential to act or think, may be more difficult (Vogels et al., 2013b). Experiment 4 indicates that, for simple causative actions, occluding the face of the agent does not diminish the perception of animacy.

6 General discussion

This paper explored perceptual influences on how we think about events, measured through the lens of linguistic description and interpretation. In Experiments 1–2, we found that English-speakers were more likely to produce passive descriptions when only the hand of an agent was visible, that is, when the animate agent was perceptually minimized and the inanimate patient was perceptually maximized. In Experiment 3, we found that speakers were more likely to choose a video with a perceptually minimized agent when comprehending a passive description than when comprehending an active description. Previous researchers have argued that the conceptual accessibility of an entity is the sum of various conceptual, perceptual, and discourse properties of that entity and its role in the event (Prat-Sala & Branigan, 2000). Consistent with this theory, we found, in a series of experiments, that visual prominence can override animacy and agency even when they are aligned to determine which referent in an event is most conceptually accessible. Conceptual properties of events do not appear to have a privileged role in shaping conceptual accessibility, highlighting the power of visual prominence in guiding event conceptualization and subsequent linguistic description and interpretation.

An important debate in the literature on how visual perceptual cues affect language production concerns the mechanism of these effects. Gleitman et al. (2007) and Myachykov

et al. (2012), among others, argue that visual perceptual effects are not modulated by structured conceptual representations of events. In Experiments 1–2, for example, it is likely that speakers were faster to fixate on the patient in the Hand-Agent condition (where the patient was perceptually prominent) than in the Body-Agent condition (see Rissman et al., 2018). It could therefore be true that participants in the Hand-Agent condition were faster to access a lexical item for the patient, e.g. "book," which would then be placed in Subject position, resulting in a passive sentence. Participants in the Hand-Agent condition also did not have sufficient information to access lexical labels such as *man* and *woman*.

Further research is needed to test the influence of lexical accessibility in driving the production results in Experiments 1–2. If the Hand-Agent events were modified such that they included stronger cues to gender (e.g., long, painted fingernails on the hand) and these events elicited more active descriptions, this would suggest that greater access to the lexical item *woman* can influence whether the agent is mentioned as the Subject. It is theoretically possible that differences in lexical access are the *only* factor driving our production findings. However, the results from the comprehension task in Experiment 3 (where lexical access is not an issue since the sentences are presented to the participants) suggest that Body-Agent scenes and Hand-Agent scenes are construed as having different levels of agent/patient prominence, and that speakers associate these different levels of prominence with different syntactic structures. In turn, this finding supports the interpretation that manipulations of visual prominence can affect conceptual accessibility, specifically, figure-ground assignment. As hypothesized by Vogels et al. (2013a), Subject choice "may be affected by protagonist or figure-ground assignment, which may be separate from the accessibility of the mental representations associated with individual entities" (1345).

The difference that we observed between Body-Agent and Hand-Agent events has methodological implications—psycholinguistic researchers often show experimental stimuli featuring just the hand of an agent, rather than the entire face and torso (e.g., Goldin-Meadow, Brentari, Coppola, Horton, & Senghas, 2015; Novack, Wakefield & Goldin-Meadow, 2016; Woodward, 1998; among others). Since speakers describe and interpret Hand-Agent scenes differently from Body-Agent scenes, experimenters need to consider whether relevant agentive information is being lost from the stimulus by *not* showing the face/torso. In other words, Hand-Agent scenes are a more ambiguous representation of an agentive event than Body-Agent scenes, which means that they might not be the cleanest way to elicit knowledge of, and language about, agentivity.

Our findings raise several unanswered questions. First, is the absence of the face *per se* driving our results? In Experiments 1–3, the face was absent from the Hand-Agent condition, but there were also fewer agent-pixels in the Hand-Agent condition than the Body-Agent condition. One possibility is that we would find the same increase in passive descriptions simply by making the agent in the videos visually smaller, but still showing the face/torso. Alternatively, removing the face may have an important impact simply because face processing has a privileged status in human cognition (Kanwisher, 2000), although we did find in Experiment 4 that occluding the face led to no decrease in perceived animacy. If the presence/absence of the face is the decisive factor, then showing the agent's torso and

head, but occluding the face (e.g. by blurring the face digitally) should lead to more passive descriptions than an unoccluded face.

A second unanswered question pertains to the role of modality effects in accounting for the findings in Experiments 1–2. As described in Section 1, Bock et al. (1992) reported that participants in their priming studies *never* produced passive descriptions of events with an animate agent and an inanimate patient, in stark contrast to the results of our experiments. Note, however, that passive is more common in written language than spoken language, at least in English (Biber, 1988), and that participants in our experiments (unlike Bock's) were always typing. We may have been able to observe a perceptual effect on animacy/agency simply because we elicited descriptions in a modality that promotes passive use.

Experiment 3 suggested that the contrast between Body-Agent and Hand-Agent scenes influences event construal, that is, figure-ground assignment. A third unanswered question then is—at what stage of event processing is this difference in construal represented? One possibility is that these different construals emerge rapidly in event processing, corresponding to different eye gaze patterns, with speakers visually inspecting Body-Agent scenes differently from Hand-Agent scenes. Alternatively, these different construals might reflect differences in higher-level implicit reasoning, rather than differences in patterns of visual inspection. For example, speakers may make the pragmatic inference that the agent must be less important in the Hand-Agent scenes, otherwise the experimenters would have shown the entire agent. Or, participants may have a general conceptual bias against representing only part of an object as the Figure, a bias which should be manifest for other types of referents, e.g., part of a boat. Tracking participants' eye gaze during the task in Experiments 1–2, and comparing visual inspection of Body-agent scenes vs. Hand-Agent scenes, would allow researchersto explore these alternatives.

7 Conclusion

The ability to construe events from multiple perspectives is pervasive in language and cognition: the same event in the world may be described alternatively as *kicking a ball* or *playing soccer*, and languages differ in terms of how events are commonly construed linguistically. In this study, we investigated the interplay between conceptual and perceptual influences on sentence formulation and sentence interpretation. Although the influence of animacy and agency has been extensively documented, the *interaction* of these factors with perceptual influences has received less attention. We find that decreasing the perceptual prominence of an animate agent, and thereby increasing the perceptual prominence of an inanimate patient, affects the conceptual accessibility of the agent. This manipulation affects production, decreasing the likelihood that the agent will be produced as the Subject of the sentence, as well as comprehension, influencing how participants interpret active vs. passive sentences. Thus, perceptual information can, at times, override conceptual cues over the course of event representation, shedding light on event construal at the linguistic/non-linguistic interface.

Acknowledgments

This research was supported by NIH RO1 DC000491 awarded to Goldin-Meadow, as well as a postdoctoral diversity supplement to Rissman through this parent grant. Thank you to Chiara Pandolfi for her hard work coding, and to all study participants.

Appendix.: Stimuli sentences in Experiments 3A-B by Item and Sentence type

Item	Sentence type	Sentence
ball	Active	Someone rolling a ball into a cup.
ball	Passive	A ball being rolled into a cup.
ball	Intransitive	A ball rolling into a cup.
book	Active	Someone tipping over a book.
book	Passive	A book being tipped over.
book	Intransitive	A book tipping over.
box	Active	Someone closing a wooden box.
box	Passive	A wooden box being closed.
box	Intransitive	A wooden box closing.
cover	Active	Someone opening the cover of a book.
cover	Passive	The cover of a book being opened.
cover	Intransitive	The cover of a book opening.
cup	Active	Someone rolling a cup across a table.
cup	Passive	A cup being rolled across a table.
cup	Intransitive	A cup rolling across a table.
hackey	Active	Someone grabbing a hackeysack off the top of a box.
hackey	Passive	A hackeysack being grabbed off the top of a box.
hackey	Intransitive	A hackeysack falling off the top of a box.
roller	Active	Someone pushing a roller down a ramp.
roller	Passive	A roller being pushed down a ramp.
roller	Intransitive	A roller rolling down a ramp.
soft-box	Active	Someone opening a striped box.
soft-box	Passive	A striped box being opened.
soft-box	Intransitive	A striped box opening.

References

Aissen J (2003). Differential Object Marking: Iconicity vs. Economy. Natural Language & Linguistic Theory, 21(3), 435–483. doi:10.1023/a:1024109008573

Altmann LJP, & Kemper S (2006). Effects of age, animacy and activation order on sentence production. Language and Cognitive Processes, 21(1–3), 322–354. doi:10.1080/01690960544000006

Antón-Méndez I (2017). Visual salience effects on speaker choices: Direct or indirect influences on linguistic processing? Applied Psycholinguistics, 38(3), 601–631. doi:10.1017/S0142716416000345

Baltaretu A, Krahmer EJ, van Wijk C, & Maes A (2016). Talking about Relations: Factors Influencing the Production of Relational Descriptions. Frontiers in Psychology, 7. doi:10.3359/fpsyg.2016.00103

Bates D, & Maechler M (2009). lme4: linear mixed effects models using S4 classes (Vol. R package version 0.999375–31)

Bernolet S, Hartsuiker RJ, & Pickering MJ (2009). Persistence of emphasis in language production: A cross-linguistic approach. Cognition, 112(2), 300–317. doi:10.1016/j.cognition.2009.05.013 [PubMed: 19539276]

Biber D (1988). Variation across speech and writing Cambridge; New York, NY: Cambridge University Press.

Bock JK (1982). Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. Psychological review, 89(1), 1–47. doi:10.1037/0033-295X.89.1.1

Bock JK (1986). Syntactic persistence in language production. Cognitive Psychology, 1986, 18, 3, July, 18(3), 355–387.

Bock JK, & Warren RK (1985). Conceptual accessibility and syntactic structure in sentence formulation. Cognition, 21(1), 47–67. doi:10.1016/0010-0277(85)90023-X [PubMed: 4075761]

Bock K, & Ferreira VS (2014). Syntactically speaking. In Ferreira VS, Goldrick M, & Miozzo M (Eds.), The Oxford handbook of language production Oxford: Oxford University Press.

Bock K, Irwin DE, & Davidson DJ (2004). Putting first things first. In Henderson JM & Ferreira F (Eds.), The interface of language, vision, and action: eye movements and visual world New York: Psychology Press.

Bock K, Loebell H, & Morey R (1992). From conceptual roles to structural relations: bridging the syntactic cleft. Psychological review, 99(1), 150–171. doi:10.1037/0033-295x.99.1.150 [PubMed: 1546115]

Branigan HP, Pickering MJ, & Tanaka M (2008). Contributions of animacy to grammatical function assignment and word order during production. Lingua, 118(2), 172–189. doi:10.1016/j.lingua.2007.02.003

Brown G, & Yule G (1983). Discourse analysis Cambridge: Cambridge University Press.

Christianson K, & Ferreira F (2005). Conceptual accessibility and sentence production in a free word order language (Odawa). Cognition, 98(2), 105–135. doi:10.1016/j.cognition.2004.10.006 [PubMed: 16307955]

Clark HH, & Begun JS (1971). The Semantics of Sentence Subjects. Language and Speech, 14(1), 34–46. doi:10.1177/002383097101400105 [PubMed: 5580469]

Coco M, & Keller F (2009). The impact of visual information on reference assignment in sentence production. Proceedings of the Annual Meeting of the Cognitive Science Society, 31(31), 274–279.

Coco MI, & Keller F (2015). Integrating mechanisms of visual guidance in naturalistic language production. Cognitive Processing, 16(2), 131–150. doi:10.1007/s10339-014-0642-0

Comrie B (1979). Definite and Animate Direct Objects: A Natural Class. Linguistica Silesiana, 3, 13–21.

Comrie B (1989). Language universals and linguistic typology: Syntax and morphology Chicago, IL: University of Chicago Press.

Corrigan R (1986). The internal structure of English transitive sentences. Memory & cognition, 14(5), 420–431. doi:10.3758/bf03197017 [PubMed: 3773708]

Croft W (1988). Agreement vs. case marking and direct objects. In Barlow M & Ferguson CA (Eds.), Agreement in natural language: Approaches, theories, descriptions (pp. 159–179). Stanford, CA: CSLI Publications.

Dahl Ö (2008). Animacy and egophoricity: Grammar, ontology and phylogeny. Lingua, 118(2), 141–150. doi: 10.1016/j.lingua.2007.02.008

Dahl Ö, & Fraurud K (1996). Animacy in grammar and discourse. In Fretheim T & Gundel J (Eds.), Reference and Referent Accessibility (pp. 47–64). Amsterdam; Philadelphia: John Benjamins.

DeLancey S (1991). Event Construal and Case Role Assignment. In Sutton L, Johnson C, & Shields R (Eds.), Proceedings of the 17th Annual Meeting of the Berkeley Linguistics Society (pp. 338–353). Berkeley, CA: Berkeley Linguistics Society.

Dowty D (1991). Thematic proto-roles and argument selection. Language, 67(3), 547–619.

Ferreira VS, & Slevc LR (2007). Grammatical encoding. The Oxford handbook of psycholinguistics (pp. 453–470). Oxford: Oxford University Press.

Fisher C, Hall DG, Rakowitz S, & Gleitman L (1994). When it is better to receive than to give: Syntactic and conceptual constraints on vocabulary growth. Lingua, 92(0), 333–375. doi:10.1016/0024-3841(94)90346-8

Fletcher-Watson S, Findlay JM, Leekam SR, & Benson V (2008). Rapid Detection of Person Information in a Naturalistic Scene. Perception, 37(4), 571–583. doi:10.1068/p5705 [PubMed: 18546664]

Flores d'Arcais GB (1975). Some perceptual determinants of sentence construction. In Flores d'Arcais GB (Ed.), Studies in perception: Festschrift for Fabio Metelli (pp. 344–373). Milan: Martello-Guinti.

Forrest LB (1996). Discourse goals and attentional processes in sentence production: the dynamic construal of events. In Goldberg A (Ed.), Conceptual structure, discourse and language Stanford, CA: CSLI Publications.

Givon T (Ed.) (1983). Topic continuity in discourse: a quantitative cross-language study Amsterdam; Philadelphia: John Benjamins.

Gleitman LR, January D, Nappa R, & Trueswell JC (2007). On the give and take between event apprehension and utterance formulation. Journal of Memory and Language, 57(4), 544–569. [PubMed: 18978929]

Goldin-Meadow S, Brentari D, Coppola M, Horton L, & Senghas A (2015). Watching language grow in the manual modality: Nominals, predicates, and handshapes. Cognition, 136(0), 381–395. doi:10.1016/j.cognition.2014.11.029 [PubMed: 25546342]

Griffin ZM, & Bock K (2000). What the Eyes Say About Speaking. Psychological Science, 11(4), 274–279. doi:10.1111/1467-9280.00255 [PubMed: 11273384]

Hamlin JK, Wynn K, & Bloom P (2007). Social evaluation by preverbal infants. Nature, 450(7169), 557–559. doi:http://www.nature.com/nature/journal/v450/n7169/suppinfo/nature06288_S1.html [PubMed: 18033298]

Hopper PJ, & Thompson SA (1980). Transitivity in Grammar and Discourse. Language, 56(2), 251–299.

Hwang H, & Kaiser E (2015). Accessibility effects on production vary cross-linguistically: Evidence from English and Korean. Journal of Memory and Language, 84(Supplement C), 190–204. doi:10.1016/j.jml.2015.06.004

Itagaki N, & Prideaux GD (1985). Nominal properties as determinants of subject selection. Lingua, 66(2), 135–149. doi:10.1016/S0024-3841(85)90292-X

Jackendoff R (1990). Semantic structures Cambridge, MA: MIT Press.

Jespersen O (1992 [1924]). The philosophy of grammar Chicago: University of Chicago Press.

Johnson S, Slaughter V, & Carey S (1998). Whose gaze will infants follow? The elicitation of gaze-following in 12-month-olds. Developmental Science, 1(2), 233–238.

Kanwisher N (2000). Domain specificity in face perception. Nature neuroscience, 3(8), 759. [PubMed: 10903567]

Kirchner H, & Thorpe SJ (2006). Ultra-rapid object detection with saccadic eye movements: Visual processing speed revisited. Vision Research, 46(11), 1762–1776. doi:10.1016/j.visres.2005.10.002 [PubMed: 16289663]

Kittredge AK, & Dell GS (2016). Learning to speak by listening: Transfer of phonotactics from perception to production. Journal of Memory and Language, 89, 8–22. doi:10.1016/j.jml.2015.08.001 [PubMed: 27840556]

Kuchinsky SE (2009). From seeing to saying: Perceiving, planning, producing: Doctoral dissertation, University of Illinois at Urbana-Champaign.

Kuhlmeier V, Wynn K, & Bloom P (2003). Attribution of Dispositional States by 12-Month-Olds. Psychological Science, 14(5), 402–408. doi:10.1111/1467-9280.01454 [PubMed: 12930468]

Lambrecht K (1994). Information structure and sentence form: topic, focus, and the mental representations of discourse referents Cambridge; New York, NY: Cambridge University Press.

Levelt WJ (1989). Speaking: From intention to articulation Cambridge: MA: MIT Press.

Lowder MW, & Gordon PC (2015). Natural forces as agents: Reconceptualizing the animate–inanimate distinction. Cognition, 136(0), 85–90. doi:10.1016/j.cognition.2014.11.021 [PubMed: 25497518]

McDonald JL, Bock K, & Kelly MH (1993). Word and World Order: Semantic, Phonological, and Metrical Determinants of Serial Position. Cognitive psychology, 25(2), 188–230. doi:10.1006/cogp.1993.1005 [PubMed: 8482072]

Meyer AS, Huettig F, & Levelt WJM (2016). Same, different, or closely related: What is the relationship between language production and comprehension? Journal of Memory and Language, 89, 1–7. doi:10.1016/j.jml.2016.03.002

Myachykov A, Garrod S, & Scheepers C (2012). Determinants of structural choice in visually situated sentence production. Acta Psychologica, 141(3), 304–315. doi:10.1016/j.actpsy.2012.09.006 [PubMed: 23085142]

Myachykov A, Garrod S, & Scheepers C (2018). Attention and Memory Play Different Roles in Syntactic Choice During Sentence Production. Discourse Processes, 55(2), 218–229. doi:10.1080/0163853X.2017.1330044

Myachykov A, Thompson D, Garrod S, & Scheepers C (2012). Referential and visual cues to structural choice in visually situated sentence production. Frontiers in Psychology, 3. doi:10.3389/fpsyg.2011.00396

New J, Cosmides L, & Tooby J (2007). Category-specific attention for animals reflects ancestral priorities, not expertise. Proceedings of the National Academy of Sciences, 104(42), 16598–16603. doi:10.1073/pnas.0703913104

Nieuwland MS, & Van Berkum JJA (2006). When Peanuts Fall in Love: N400 Evidence for the Power of Discourse. Journal of Cognitive Neuroscience, 18(7), 1098–1111. doi:10.1162/jocn.2006.18.7.1098 [PubMed: 16839284]

Novack MA, Wakefield EM, & Goldin-Meadow S (2016). What makes a movement a gesture? Cognition, 146, 339–348. doi:10.1016/j.cognition.2015.10.014 [PubMed: 26513354]

Osgood CE (1971). Where do sentences come from? In Steinberg DD & Jakobovitz LA (Eds.), Semantics: An interdisciplinary reader in philosophy, linguistics and psychology (pp. 497–529). London: Cambridge University Press.

Osgood CE, & Bock JK (1977). Salience and sentencing: Some production principles. In Rosenberg S (Ed.), Sentence production: developments in research and theory (pp. 89–140). Hillsdale, NJ: Lawrence Erlbaum.

Pires A (2006). The minimalist syntax of defective domains: gerunds and infinitives Amsterdam; Philadelphia: John Benjamins.

Prat-Sala M, & Branigan HP (2000). Discourse Constraints on Syntactic Processing in Language Production: A Cross-Linguistic Study in English and Spanish. Journal of Memory and Language, 42(2), 168–182. doi:10.1006/jmla.1999.2668

Prentice JL (1967). Effects of cuing actor vs cuing object on word order in sentence production. Psychonomic Science, 8(4), 163–164. doi:10.3758/bf03331600

Primus B (1999). Cases and thematic roles: Ergative, accusative and active (Vol. 393): Tübingen: Niemeyer.

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

Radanović J, Westbury C, & Milin P (2016). Quantifying semantic animacy: How much are words alive? Applied Psycholinguistics, 37(6), 1477–1499. doi:10.1017/S0142716416000096

Rissman L, Goldin-Meadow S & Woodward A (2018). From visual prominence to event construal: influences (and non-influences) of eyegaze. Proceedings of the Annual Meeting of the Cognitive Science Society.

Saxe R, Tzelnic T, & Carey S (2007). Knowing who dunnit: Infants identify the causal agent in an unseen causal interaction. Developmental psychology, 43(1), 149–158. doi:10.1037/0012-1649.43.1.149 [PubMed: 17201515]

Slaughter V, & Heron-Delaney M (2010). When do infants expect hands to be connected to a person? Journal of Experimental Child Psychology, 108(1), 220–227. doi:10.1016/j.jecp.2010.08.005 [PubMed: 20864118]

Sridhar SN (1988). Cognition and sentence production: A cross-linguistic study New York: Springer-Verlag.

Stein G (1979). Studies in the function of the passive Tübingen: Narr.

Thompson SA (1987). The passive in English: A discourse perspective. In Channon R & Shockey L (Eds.), In Honor of Ilse Lehiste (pp. 497–511). Dordrecht; Providence: Foris publications.

Tomlin RS (1997). Mapping conceptual representations into linguistic representations: The role of attention in grammar. In Nuyts J & Pederson E (Eds.), Language and conceptualization (pp. 162–189). Cambridge; New York, NY: Cambridge University Press.

Tremoulet PD, & Feldman J (2000). Perception of Animacy from the Motion of a Single Object. Perception, 29(8), 943–951. doi:10.1068/p3101 [PubMed: 11145086]

Tsunoda T (1985). Remarks on transitivity. Journal of Linguistics, 21, 385–396.

van Nice KY, & Dietrich R (2003). Task sensitivity of animacy effects: evidence from German picture descriptions. Linguistics, 41(5), 825. doi:10.1515/ling.2003.027

Vogels J, Krahmer E, & Maes A (2013a). Who is where referred to how, and why? The influence of visual saliency on referent accessibility in spoken language production. Language and Cognitive Processes, 28(9), 1323–1349. doi:10.1080/01690965.2012.682072

Vogels J, Krahmer E, & Maes A (2013b). When a stone tries to climb up a slope: the interplay between lexical and perceptual animacy in referential choices. Frontiers in Psychology, 4, 1–15. [PubMed: 23382719]

Woodward AL (1998). Infants selectively encode the goal object of an actor's reach. Cognition, 69(1), 1–34. doi:10.1016/S0010-0277(98)00058-4 [PubMed: 9871370]

Yamamoto M (1999). Animacy and reference: a cognitive approach to corpus linguistics Amsterdam; Philadelphia: John Benjamins

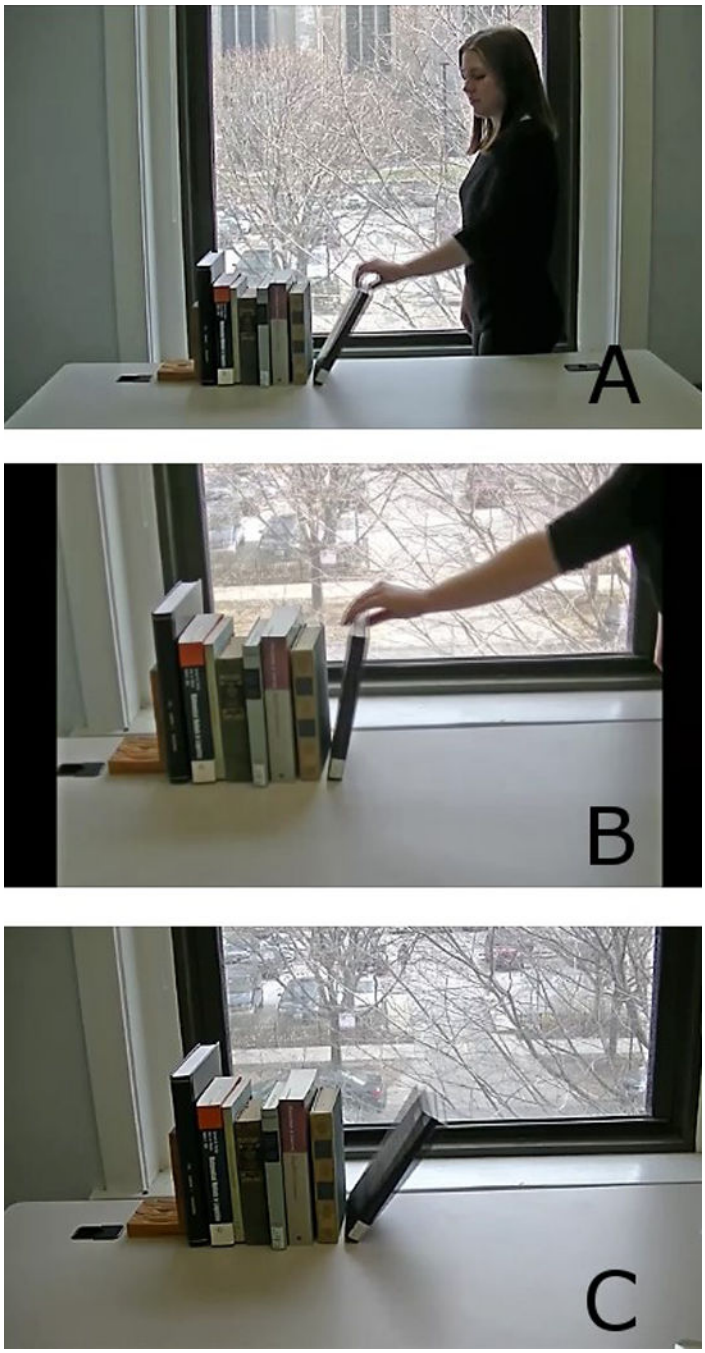

Figure 1.
Examples of Body-Agent (A), Hand-Agent (B), and No-Agent (C) events in Experiment 1. In A and B a woman is tipping over a book; in C a book is falling over.

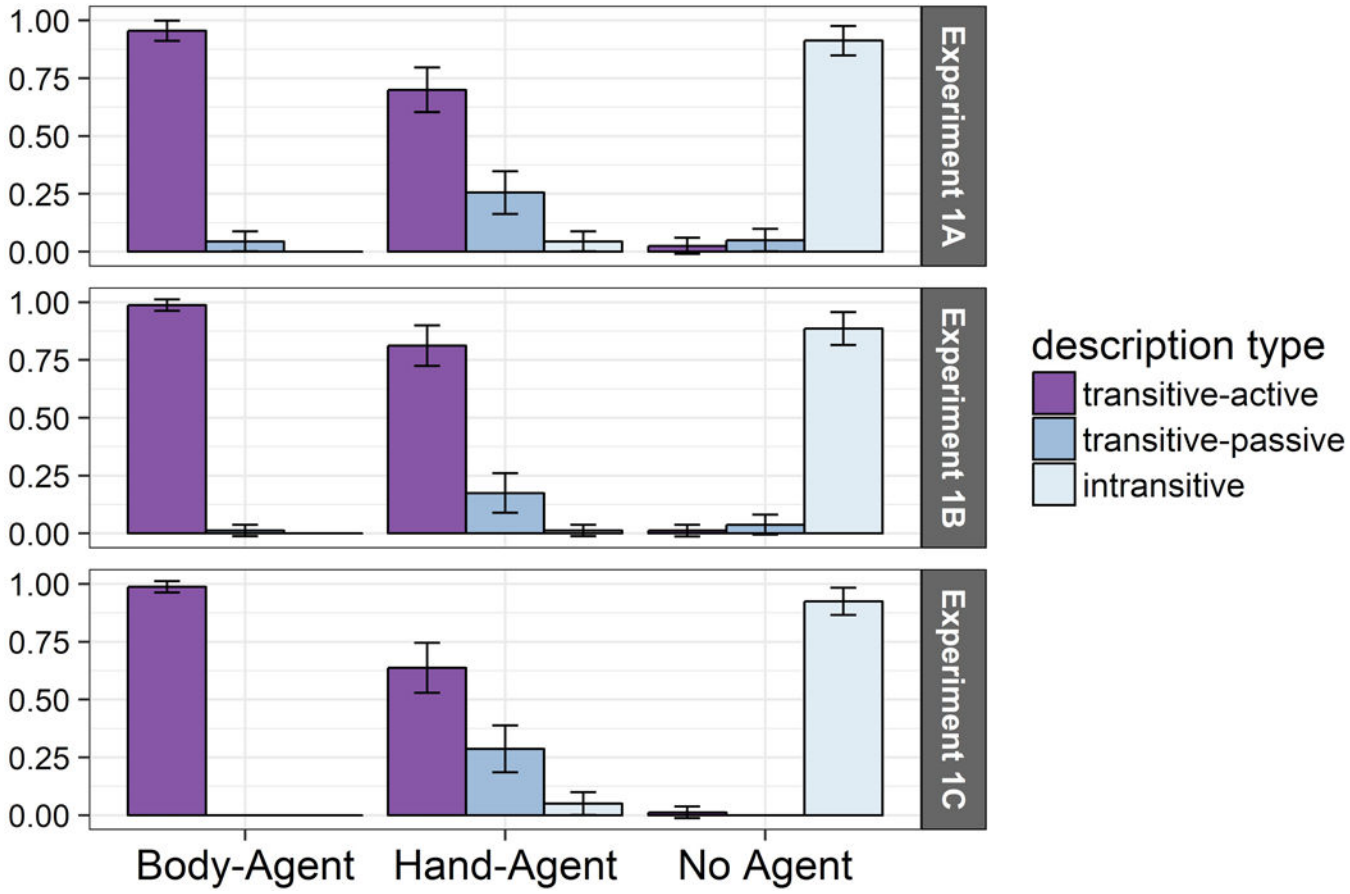

Figure 2.

Proportion of responses coded as active, passive or intransitive in the Body-Agent, Hand-Agent and No Agent conditions, in Experiments 1A-1C (1A: between subjects; 1B: within subjects randomized; 1C: within subjects blocked). Error bars show 95% confidence intervals.

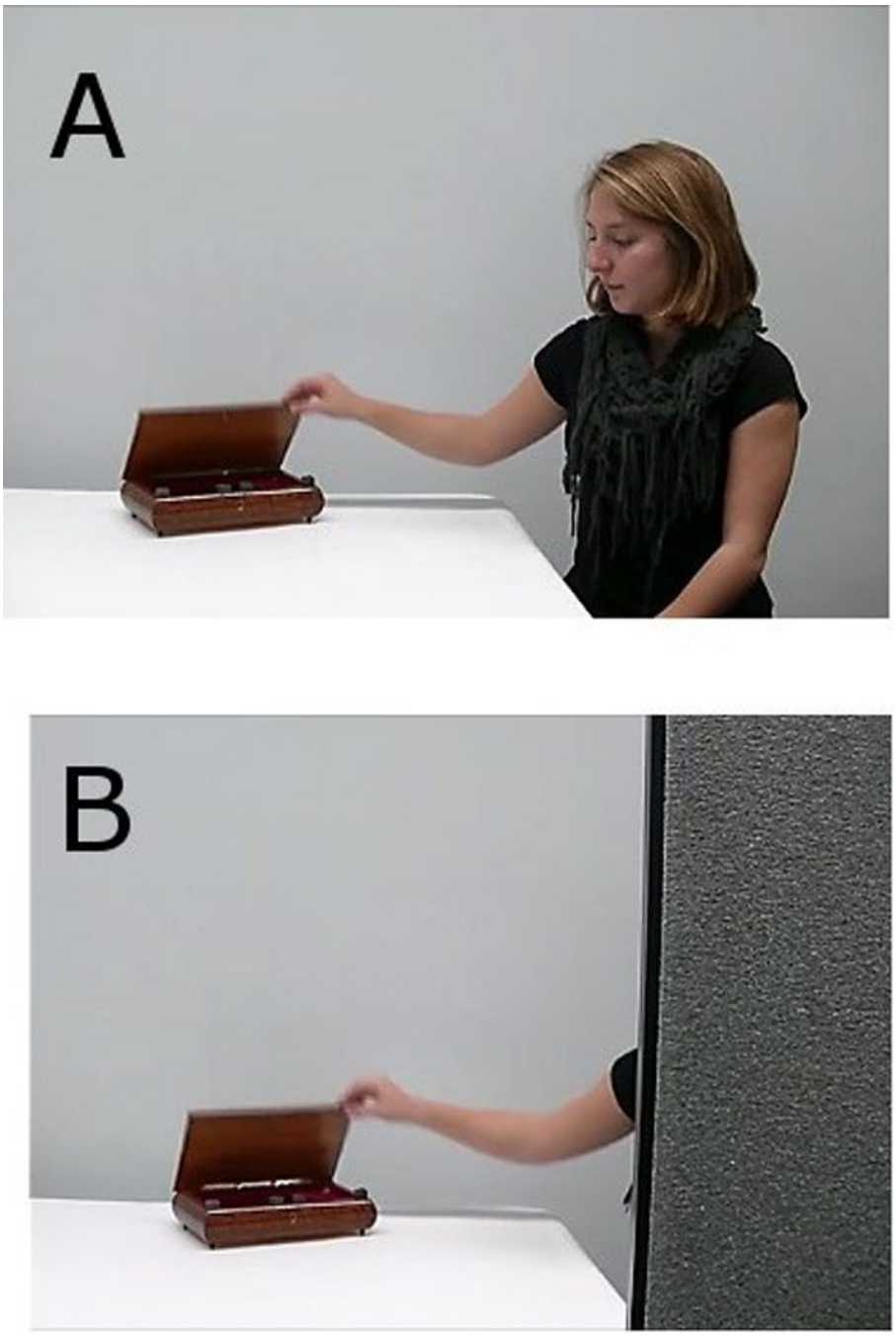

Figure 3.
Examples of Body-Agent (A) and Hand-Agent (B) events from Experiment 2. In the pictured event the woman is shutting a box.

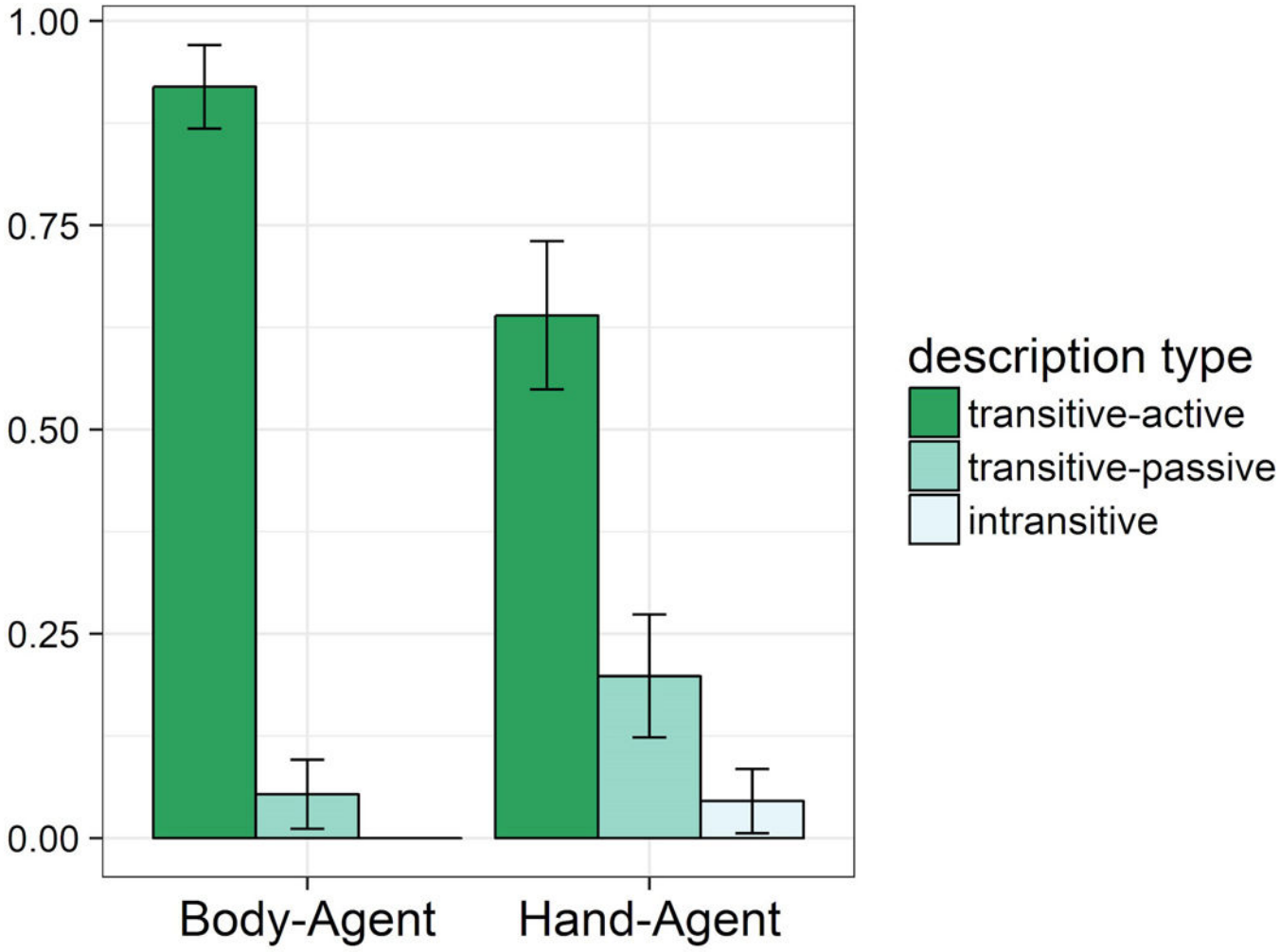

Figure 4.
Proportion of responses coded as active, passive or intransitive in the Body-Agent and Hand-Agent conditions in Experiment 2. Error bars show 95% confidence intervals.

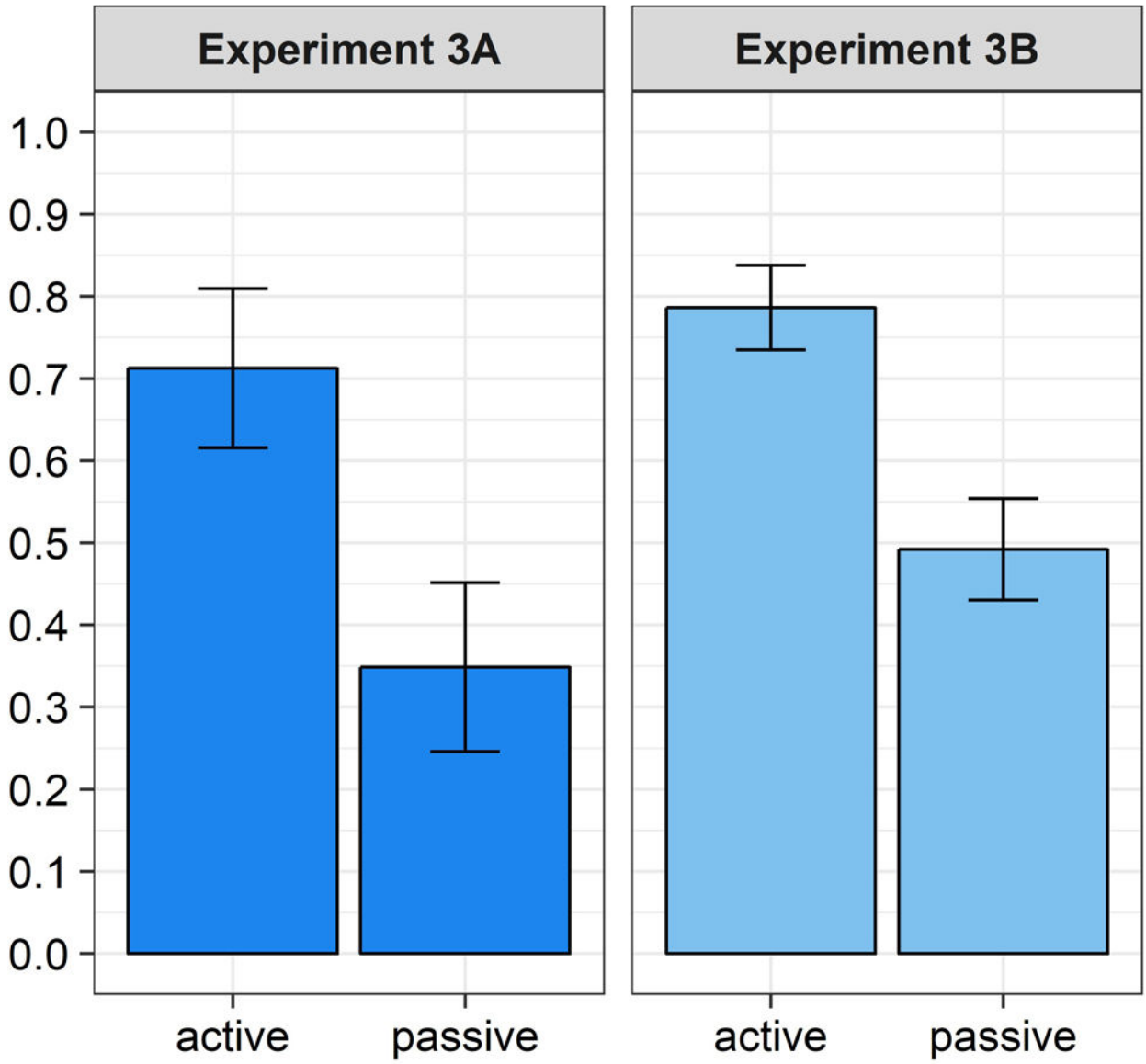

Figure 5.
Proportion of Body-Agent vs. Hand-Agent trials where the Body-Agent video was chosen, for each sentence type in Experiment 3A and 3B. Error bars show 95% confidence intervals.

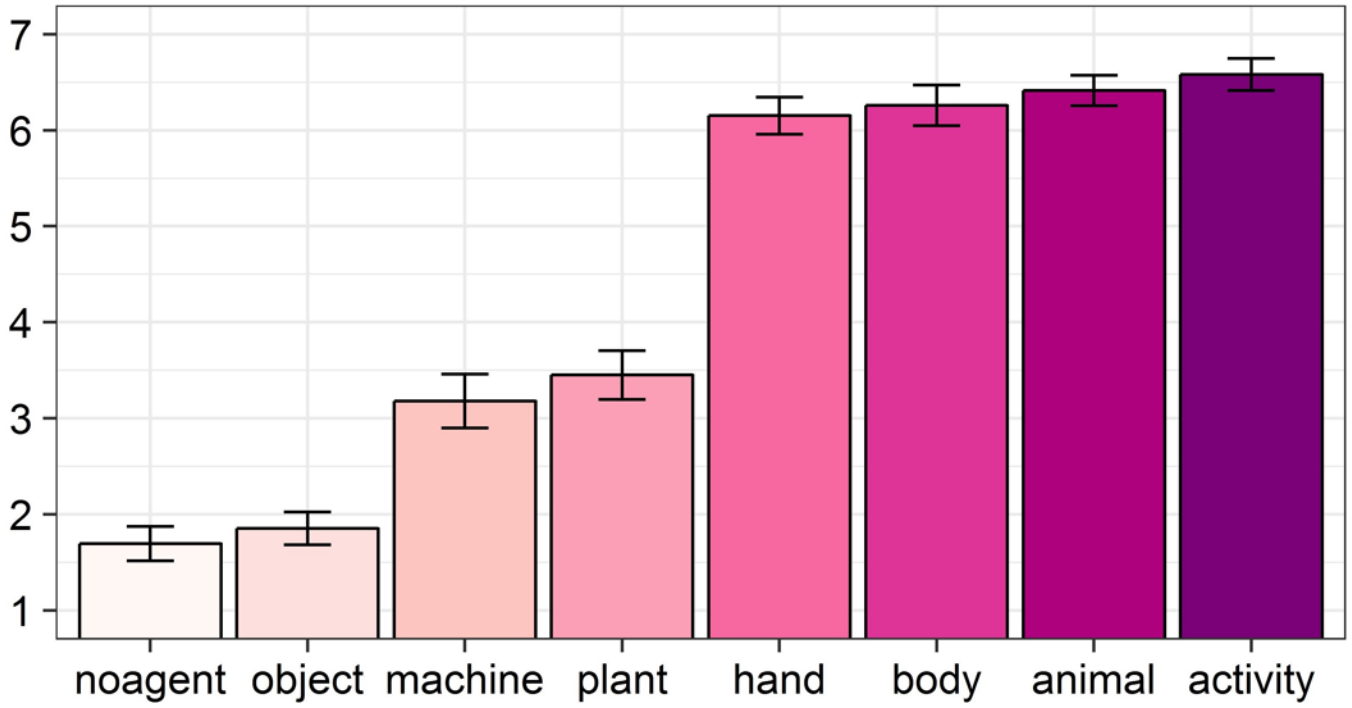

Figure 6.

Average animacy rating on a 1 – 7 point scale for each of the video types in Experiment 4. Error bars show 95% confidence intervals.

Table 1.

Conditions for Experiments 3A-B

Condition	Sentence Type	Video 1	Video 2
1	Active	Body-Agent	Hand-Agent
2	Passive	Body-Agent	Hand-Agent
3	Active	Body-Agent	No Agent [Body]
4	Active	Hand-Agent	No Agent [Hand]
5	Passive	Body-Agent	No Agent [Body]
6	Passive	Hand-Agent	No Agent [Hand]
7	Intransitive	Body-Agent	No Agent [Body]
8	Intransitive	Hand-Agent	No Agent [Hand]

Table 2.

Conditions for Experiment 4 with example videos

Condition	Description	Examples
Body	A person causes a change in an inanimate object; face/torso is visible	A person knocks over a book, a person pushes a ball down a ramp
Hand	A person causes a change in an inanimate object; face/torso is hidden behind a screen	A hand shuts a box, a hand grabs a ball off a box
Activity	A person performs an activity (not on an object); face/torso is visible	A person runs in place, a person jumps up and down
Object	An inanimate object is involved in an event	Water runs in a sink, a top spins on a table
No Agent	An inanimate object undergoes a change on its own	A book tips over, a marker rolls down a ramp
Animal	A non-human animal performs an action	A tiger licks her cub, a frog hops on a log
Plant	A plant grows/changes	Time lapse footage of a flower blooming, time lapse footage of a seed sprouting
Machine	A machine or robot performs a function	A robot walking up a hill, an industrial slicer chopping vegetables

Table 3.

Friedman's post hoc pairwise comparisons for the conditions in Experiment 4.

	Activity	**Animal**	**Body**	**Hand**	**Plant**	**Machine**	**Object**	**No Agent**
Activity								
Animal	16							
Body	18	2						
Hand	28	12	10					
Plant	128	112	110	100				
Machine	125	106	104	94	5			
Object	179	163	161	151	51	56		
No Agent	192	176	174	164	65	70	13	