COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 45 (2021) e13065

ISSN: 1551-6709 online
DOI: 10.1111/cogs.13065

Speakers Align With Their Partner's Overspecification During Interaction

Jia E. Loy,[a] Kenny Smith[b]

[a]*Language Science and Technology, Saarland University*
[b]*Linguistics and English Language, University of Edinburgh*

Received 26 January 2021; received in revised form 11 October 2021; accepted 24 October 2021

Abstract

Speakers often overspecify by encoding more information than is necessary when referring to an object (e.g., "the *blue* mug" for the only mug in a group of objects). We investigated the role of a partner's linguistic behavior (whether or not they overspecify) on a speaker's own tendency to overspecify. We used a director–matcher task in which speakers interacted with a partner who either consistently overspecified or minimally specified in the color/size dimension (Experiments 1, 2, and 3), as well as with a partner who switched behaviors midway through interaction (Experiments 4 and 5). We found that speakers aligned with their partner's linguistic behavior to produce overspecific or minimally specific descriptions, and we saw little evidence that the alignment was enhanced by lexical or semantic repetition across prime and target trials. Time-course analyses showed that alignment increased over the course of the interaction, and speakers appeared to track a change in the partner's linguistic behavior, altering their reference strategy to continue matching that of their partner's. These results demonstrate the persistent influence of a partner's behavior on speakers across the duration of an interaction.

Keywords: Reference production; Overspecification; Interaction; Alignment

1. Introduction

When describing an object to a listener, speakers[1] have to decide what information to encode in order for the listener to successfully identify the referent. For instance, a speaker may ask for "the blue mug," where *blue* is useful for disambiguation if we have a rack of

Correspondence should be sent to Jia E. Loy Department of language science and technology, Universität des Saarlandes, Campus C7.2, 66123 Saarbrücken. E-mail: jialoy@coli.uni-saarland.de

different colored mugs. A common observation, however, is that speakers routinely *over-specify* by encoding modifiers such as "blue" when minimal specification (e.g., saying "the mug," if we only have one mug among other objects) would be sufficient for establishing reference. Overspecification poses a problem for theories of language production emphasizing communicative efficiency (e.g., Grice, 1975; Peloquin, Goodman, & Frank, 2020), given the assumption that unnecessary information is more costly to produce and adds no informational value to the utterance. This raises the question of why speakers overspecify. Some researchers have attributed overspecification to egocentric factors such as the ease of encoding certain properties (Koolen, Gatt, van Gompel, Krahmer, & Van Deemter, 2016). Others have argued that overspecification can in fact serve to facilitate reference identification for the addressee (Rubio-Fernández, 2019). Here, we consider another aspect of dailog—that of alignment with an interlocutor's pragmatic behavior (whether or not they overspecify). We examine whether speakers align with their interlocutor to either overspecify or minimally specify, and how this effect unfolds over the course of the interaction.

1.1. Speaker-centric explanations for overspecification

Previous research on overspecification has attributed this behavior to egocentric factors such as ease of processing or encoding for a speaker. This relies on the assumption that speakers tend to direct their attention to information that stands out, for instance, the color of an object (Belke & Meyer, 2002). Indeed, studies consistently show that color is overspecified more than other properties, such as size (Koolen, Krahmer, & Swerts, 2016; Pechmann, 1989), a finding that has been linked to the visual salience and absoluteness of an object's color (Tarenskeen, Broersma, & Geurts, 2015). This has led researchers to link overspecification to a heuristic tendency in speakers to encode attributes that are perceptually salient in an object (Koolen, Goudbeek, & Krahmer, 2013). Specifying information that is salient may also be more efficient from a production perspective. Pechmann (1989) observed that speakers frequently started producing overspecific descriptions before they had finished scanning a display, suggesting that such behavior may be useful in facilitating production before visual processing is complete. Correspondingly, overspecification also increases with scene complexity, such as when there are more distractors, or greater variation in color within a scene (Koolen et al., 2013, 2016).

1.2. Partner-oriented explanations for overspecification

Egocentric explanations for overspecification typically do not consider this behavior in connection with the addressee's perspective during interaction. However, other researchers have argued that overspecification may arise as an addressee-oriented process. This builds on the view in referential production that speakers design utterances in order for an addressee to efficiently establish reference (e.g., Arnold, 2008). Thus, speakers may overspecify if it is beneficial for their interlocutor in identifying the intended referent (Arts, Maes, Noordman, & Jansen, 2011; Rubio-Fernández, 2016). This was observed by Rubio-Fernández (2016), who found that speakers only overspecified when objects were atypically colored (e.g., a pink banana; cf. Westerbeek, Koolen, & Maes, 2015), and were more likely to do so when instructions emphasized the possibility of communication breakdown with the addressee. Speakers have also been found to overspecify less when an addressee-oriented motivation for overspecifying (minimizing ambiguity for the hearer) is removed from the descriptive context (Davies & Katsos, 2009).

A partner-oriented view is supported by empirical evidence from comprehension, which shows that addressees expect speakers to use prenominal adjectives rationally: On hearing a scalar adjective (e.g., “tall…”), addressees begin to narrow their visual search down to a set of contrastive objects (e.g., a tall and a short glass; Sedivy, Tanenhaus, Chambers, & Carlson, 1999). This effect disappears with a speaker who uses modification unreliably (Grodner & Sedivy, 2011). Taken together, the findings from production and comprehension suggest that referential specification depends on both parties in an exchange: Addressees have expectations about how specific a speaker should be, which speakers in turn take into account when producing a description. Partner-oriented overspecification is also not limited to modifier use: Speakers can be more or less informative in their choice of a lexical name. Speakers in Brennan and Clark (1996) who had entrained on a subordinate term (e.g., “pennyloafer” for a shoe) with their partner continued using this term on later trials despite the simpler, basic-level name “shoe” being sufficient. With a new partner, however, they tended to revert to basic-level terms, demonstrating that such overspecification is partner dependent. Thus, overspecification appears to be more than a simple case of egocentricity; rather, it takes into account interactional demands such as a conversation partner’s identity or perspective.

Here, we consider another aspect of the interaction—that of a partner’s linguistic behavior. Specifically, a speaker’s tendency to overspecify may be influenced by their partner’s referential strategy—whether or not they overspecify. Although previous research has shown that overspecification is sensitive to partner-dependent factors, the majority of studies have focussed on one-sided communication with hypothetical listeners or fixed speaker/addressee roles; comparatively little work has examined the role of an interlocutor’s behavior on overspecification in bidirectional interaction.

Outside of overspecification, referential communication studies involving dailog show that interlocutors often adopt their partner’s forms of expression over time, leading to coordinated behavior at many levels of linguistic structure (Branigan, Pickering, & Cleland, 2000; Branigan, Pickering, Pearson, McLean, & Brown, 2011; Cleland & Pickering, 2003; Garrod & Anderson, 1987; Hartsuiker, Pickering, & Veltkamp, 2004; Kim, Horton, & Bradlow, 2011). Branigan et al. (2000) provide a seminal demonstration of this at the syntactic level. In their study, they showed that participants were more likely to produce either Prepositional Object (PO) descriptions (e.g., “the pirate giving the banana to the burglar”) or Double Object (DO) descriptions (e.g., “the pirate giving the burglar the banana”) depending on which form their partner used on the turn before. Branigan et al. explained their findings in terms of priming processes—a subconscious mechanism by which recently encountered structures receive higher activation and hence are more likely to be reused by speakers (cf. Bock, 1986).

A more general account of this is provided by the Pickering and Garrod (2004) Interactive Alignment model. Under this account, conversation partners come to align on multiple levels of linguistic representation, from low-level phonological representations to high-level situation models, via a resource-free priming mechanism. Goudbeek and Krahmer (2012) provide a demonstration of such alignment at the conceptual level in the form of overspecification. In their Experiment 3, participants were exposed to overspecific prime descriptions containing a preferred and a dispreferred attribute (e.g., "the *red* chair seen *from the front*") when either attribute would suffice. They found that participants were more likely to produce similarly overspecific target descriptions compared to participants from an earlier experiment who had heard primes with only one attribute. Thus, overspecification, at least in the form of encoding more attributes than is necessary, appears to be influenced by a partner's linguistic behavior.

A key feature of Interactive Alignment is the assumption of a largely automatic and unconscious process, where interconnections between levels allow alignment at lower levels to facilitate the effect at higher levels. In line with this, syntactic alignment is known to be enhanced by a lexical boost—lexical repetition between prime and target (e.g., *give–give* vs. *give–show*; Branigan et al., 2000), as well as a semantic boost—semantic relatedness between prime and target (e.g., *sheep–goat* vs. *sheep–knife*; Cleland & Pickering, 2003). This effect is well documented at the syntactic level of representation (Branigan et al., 2000; Cleland & Pickering, 2003; Hartsuiker et al., 2004), but has also been observed at other levels such as the semantic, phonetic, and conceptual representations (Branigan et al., 2011; Garrod & Anderson, 1987; Kim et al., 2011). Although Goudbeek and Krahmer (2012) did not systematically manipulate lexical or semantic repetition in their study, they interpreted their results within an Interactive Alignment framework based on participants' posttest reports suggesting that alignment was largely unconscious. However, it remains unclear whether alignment in a pragmatic behavior such as modifier overspecification is sensitive to lexical or semantic effects. Thus, in the current study we examine whether speakers align with their partner in overspecification, as well as whether this is enhanced by lexical and/or semantic boost effects.

It should be noted that there are routes other than Interactive Alignment through which speakers may align with a partner in overspecification. In particular, in contrast to the automatic, unconscious mechanism emphasized by Interactive Alignment, speakers may adopt a partner's linguistic behavior via more conscious or strategic processes. For instance, speakers may converge with various aspects of a partner's production for social motivations, such as a wish to affiliate with their partner. Participants in Weatherholtz, Campbell-Kibler, and Jaeger (2014) were more likely to produce syntactic constructions (PO or DO) in line with a speaker who appeared to share the same political orientation as their own. In a similar vein, Kim et al. (2011) found that speakers showed greater phonetic alignment with a partner who spoke the same regional dialect than one who spoke a different dialect or native language. Alternatively, speakers may assume, via reasoning about their partner's production preferences, that adopting the same referential strategy may be beneficial for communication. Such inferences may be particularly relevant in the context of overspecification, where speakers are known to produce utterances that may be more effortful from their perspective but helpful for their partner (e.g., Rubio-Fernández, 2016). It is conceivable, therefore, that speakers may reason about their partner's motivation for overspecifying (e.g., to be helpful), and engage in similar behavior with the goal of facilitating communication.

Crucially, regardless of the underlying mechanism, the common factor across these explanations is the influence of a partner's behavior on alignment. Here, we similarly focus on the role of partner behavior, looking at the phenomenon of overspecification. Specifically, we ask whether a partner's referential behavior (overspecification or minimal specification) affects a speaker's own tendency to overspecify. Although we do not set out to distinguish between the possible mechanisms underlying such alignment, we note that a lexical or semantic boost on any partner effects we observe would be in line with an Interactive Alignment account; more conscious or strategic processes on the other hand do not make particular predictions about lexical or semantic effects.

1.3. Time course of effects on overspecification

A related question concerns the time course of effects on overspecification in extended interaction. In particular, if speakers align with their partner in overspecification behavior, it is of interest as to whether this partner influence persists over the course of the interaction. Although a large body of research demonstrates how an interlocutor's utterance affects a speaker's production in an immediately subsequent utterance, comparatively little work has examined how this effect may evolve over time. This is an important question, however, as interactions between individuals in naturalistic discourse are typically not limited to a single-turn exchange.

Previous work on overspecification has addressed time course effects by looking at within-speaker consistency across trials. These studies suggest that speakers typically adopt a referential strategy at the start (e.g., overspecification or minimal specification), which they maintain over the production of subsequent utterances (cf. Rubio-Fernández, 2016, 2019; Tarenskeen et al., 2015). This pattern of behavior echoes findings from cumulative priming research, which demonstrate the longevity of syntactic priming effects within individuals, such that effects increase in magnitude the more primes a speaker is exposed to. These effects have been observed in both empirical studies on written production (Kaschak, 2007; Kaschak, Kutta, & Schatschneider, 2011) as well as computational modeling of syntactic priming in monologue (Reitter, Keller, & Moore, 2011). Similar results have also been found in corpus-based analyses of conversational interaction (Jaeger & Snider, 2008). Together, the findings demonstrate that priming effects do not remain constant over time, but increase cumulatively as a function of prior exposure to a structure. However, theories on cumulative priming make little reference to the role of a partner during interaction, and instead center around a production bias toward syntactic repetition within a speaker (Bock, 1986; Branigan, 2007). Hence, it remains to be seen whether similar effects arise from alignment with a partner in interaction.

If speakers align with a partner's overspecification behavior at the start of the interaction, we would expect that they continue with the same referential strategy on subsequent descriptions. However, the source of this consistency remain unclear. One possibility is that the partner effect is long-lasting, and extends across the duration of the interaction; alternatively, speakers may be initially influenced by their partner, but continue with the same strategy simply due to a speaker-internal tendency toward consistency, regardless of what their partner does. In order to tease apart these possibilities, we examine whether a change in the partner's referential strategy (e.g., from overspecification to minimal specification) midinteraction influences the speaker's overspecification behavior.

1.4. The current study

We investigate the role of a partner's linguistic behavior—whether or not they overspecify—on a speaker's own tendency to overspecify. We have two main questions in mind. First, we examine whether speakers align with a partner to overspecify or minimally specify, and whether this alignment is enhanced by lexical and/or semantic repetition.[2] While alignment effects are well documented in psycholinguistic research on dailog, existing studies have mainly focussed on the syntactic, and to a smaller extent lexical and phonetic, levels of representation. Separately, research on priming within monologue conditions has shown that these effects occur in pragmatic comprehension (whether or not to assign an enriched meaning to quantifier and number statements; Bott & Chemla, 2016). Thus, our results will extend findings in both fields by asking whether alignment in dailog extends to a pragmatic phenomenon—in this case overspecification in reference production. Our second question addresses the time course of the hypothesized partner effects on overspecification during an interaction. Specifically, we ask whether any partner effects we observe on speakers' overspecification behavior persist over the course of the interaction.

Experiments 1–3 are designed to focus on the first question. Experiments 1 and 2 look at overspecification featuring a single, prenominal color adjective, comparing interaction with a partner who is consistently overspecific to one who is consistently minimally specific. Experiment 3 turns to overspecification with a different property—that of size, which speakers are typically less inclined to overspecify (e.g., Koolen et al., 2013). Finally, Experiments 4 and 5 further probe the question of partner effects over time by investigating interaction with a partner whose linguistic behavior changes (e.g., from overspecification to minimal specification) midway through the interaction. Across all five experiments we find that speakers tended toward overspecification or minimal specification in line with their partner's behavior. We also observe qualitatively different trends in alignment over time with color and with size modification. Finally, we demonstrate that speakers track alignment across a change in the partner's linguistic behavior, adjusting their production pattern to continue to match their partner's production behavior. This holds for both a shift from overspecification to minimal specification as well as vice versa.[3] We discuss our results in the context of current theories on overspecification and also alignment in interaction.

2. Experiment 1

Experiment 1 investigated whether speakers are influenced by a partner's linguistic behavior to similarly produce overspecific object descriptions using prenominal color modifiers, and whether such alignment is sensitive to lexical and semantic boost effects. The experiment was a director–matcher task in which participants alternated between describing an image for their partner, and matching their partner's description to the correct image. On critical trials, participants heard a partner-produced prime, which either contained an overspecific color modifier (overspecific partner) or featured only the bare noun (minimally specific partner). On a subsequent target trial, participants produced an object description for their partner, where color was never necessary to distinguish between the objects in the display. We manipulated whether the prime and target objects were of the *same* or *different* colors, and whether they were *within* the same semantic category or *across* categories. We analyzed participants' overall overspecification behavior, as well as their change in this behavior over time, comparing interaction with a partner who always overspecified to one who minimally specified. Preregistration details for the experiment can be found here: https://osf.io/m2r9u/

2.1. Method

2.1.1. Participants

Sixty-eight self-reported native speakers of English between the ages of 18 and 35 took part in the experiment. Participants were recruited from the career service website of the University of Edinburgh and paid £6 to take part.

2.1.2. Materials

Critical stimuli. Critical images consisted of four different colored variants (red, blue, green, yellow) of each of eight monosyllabic noun items. These were from two semantic categories: items of clothing (cap, glove, scarf, sock) and items of kitchenware (bowl, fork, mug, pot). This yielded a total of 32 unique objects. Each object was associated with voice recordings specifying the noun item and color (e.g., "the red cap"), as well as only the bare noun (e.g., "the cap"). The recordings were produced by a male native speaker of English, who spoke with a relatively standard English accent.

Filler stimuli. The experiment included two other types of images used in filler trials: images of easily recognizable natural objects and photographs of human facial expressions. For natural object fillers, 32 images (16 animals; 16 fruit/vegetables) with high nameability (name agreement $\geq 90\%$ and H value ≤ 0.5) were selected from the Bank of Standardised Stimuli (BOSS; Brodeur, Dionne-Dostie, Montreuil, & Lepage, 2010), a set of normative color images created for psycholinguistic research. For each object, we created two recordings of its modal name (as determined through the study's norming procedure, e.g., "the giraffe"). The recording played on each match trial was selected randomly from the set of two. For facial expression fillers, eight photographs (four men; four women) for each of six facial emotions (angry, disgusted, frightened, happy, sad, surprised) were taken from the Karolinska Directed Emotional Faces (KDEF) database (Lundqvist, Flykt, & Ohman, 1998). The images were converted to grayscale to discourage participants from relying on color-modified facial features such as hair and eye color in their descriptions. For each facial emotion, we created a variety of recordings (between four and six) to describe the subject's expression (e.g., "the woman who looks angry," "the angry-looking woman," "the angry one"), with the recording played on each match trial selected randomly from this set. All filler recordings were produced by the same speaker who recorded the critical stimuli.

2.1.3. Design

We manipulated color (same vs. different for prime and target) and semantic category (within vs. across category for prime and target) within subjects. In addition, partner behavior (overspecific vs. minimally specific) was manipulated between subjects.

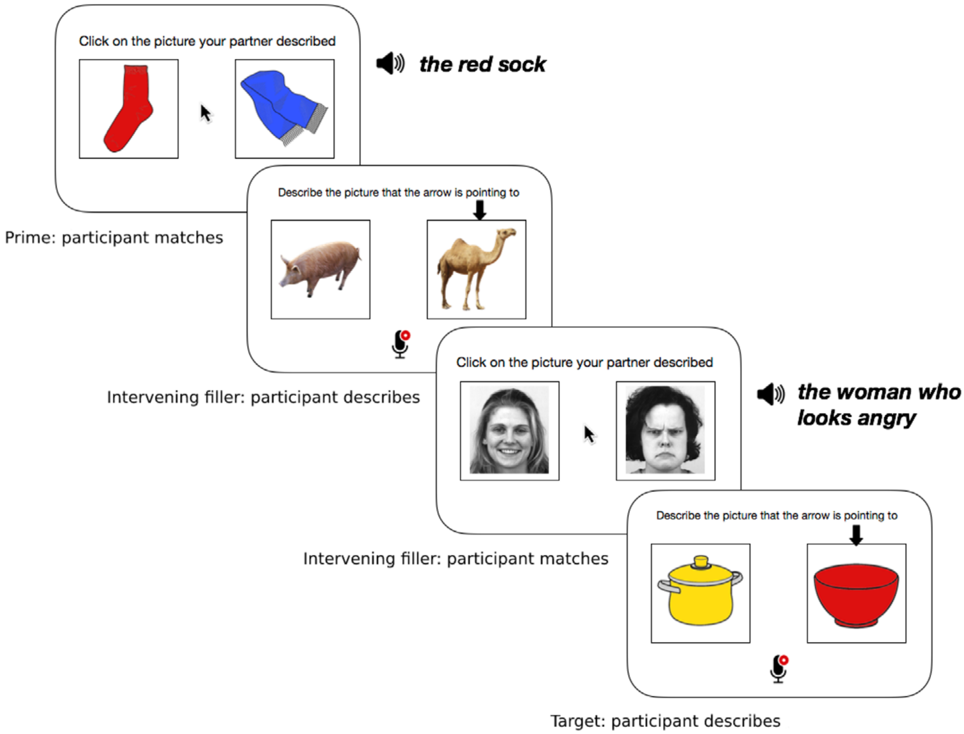

Fig. 1. Example of a four-trial sequence that makes up a critical item. In this case the prime and target objects feature the same color but are drawn from different categories.

The experiment included 32 critical items—eight per within-subjects condition. Critical items consisted of a four-trial sequence: a prime (participant matching), an intervening filler (participant describing), an intervening filler (participant matching), and a target (participant describing) trial. Fig. 1 shows an example of four trials within a critical item. Within each condition, each of the eight noun items appeared once as a prime (matching) and once as a target (describing) referent. Color was varied within objects such that the four objects from each category each appeared once in a different color across the eight prime and eight target trials.

On critical prime and target trials, the display featured the target image alongside a distractor, chosen randomly from the full set of 32 on the constraint that it differed from the target in both color and noun item, that is, color was never necessary to distinguish between the two. Thus, producing a color adjective would always constitute an overspecific description. Intervening filler trials similarly presented two images, both either natural objects or facial expressions. These filler trials were included to reduce the connection between the prime and target trials (cf. Goudbeek & Krahmer, 2012). The relative positions of target and distractor on each trial were randomized, with the target appearing equally often on each side in each condition.

Outside of critical items, the experiment included an additional 128 filler trials: 48 natural object trials, 48 facial expression trials, and 32 trials featuring the same images used in critical trials. For natural object fillers, the target and distractor images were always drawn from the same superordinate category (i.e., animals or fruit/vegetables). These fillers were included to elicit easily nameable, unmodified referring expressions. For facial expression fillers, the target and distractor images always presented two faces from the same gender to encourage participants to focus on the facial expression in their description. These fillers were included primarily as a more open-ended description task to distract participants from the actual focus of the experiment. For fillers featuring the same images as those in critical trials, half the trials presented two different objects of the same color while the other half presented two of the same object in different colors. These were included to prevent participants from resorting to specifying only color or only the noun item over the course of the experiment, and never occurred within critical items to avoid influencing participants before a target trial. For all three types of fillers, half of the trials served as participant-describing trials and the other half as participant-matching trials.

The order of presentation of items was randomized on each run with the constraints that at least six filler trials (i.e., three participant match–describe sequences) preceded the first critical item, and at least two filler trials occurred between critical items.

2.1.4. Procedure

Participants were tested individually in sound-attenuated booths. Prior to the experiment, participants were told they would be playing a picture description and matching game over a networked connection with a participant located at a partner university. To reinforce the cover story, the experiment included wait screens and partner delays, as well as disfluencies in the partner's speech to simulate realistic partner behavior (described below).

After participants were briefed and began the experiment, a "connecting" screen appeared for 15 s to simulate a wait for their partner to connect to the network. This was followed by an audio check phase where they were instructed to click on an object as indicated by their partner, and then to record a similar instruction ostensibly to send to their partner. After this, the experiment began, with the participant matching on the first trial and alternating between description and match trials. On description trials, the target and distractor appeared side-by-side with an arrow pointing to the target. After 500 ms, microphone symbol turned red to signify that the participant was being recorded. Participants were instructed to click on the microphone when they had finished speaking to send their description to their partner. A "wait" message then appeared on the screen for a set interval to simulate their partner selecting an object. This was fixed at 2,000 ms on critical trials and randomly variable between 1,800 and 3,000 ms on filler trials, with the delay decreasing as the experiment progressed. On match trials, the target and distractor images appeared side-by-side. Playback of the audio recording for the target image began after a set delay to simulate the partner's utterance onset latency. This was fixed at 2,000 ms on critical trials and randomly variable between 1,800 and 3,000 ms on filler trials, with the delay decreasing as the experiment progressed. After the audio playback, the mouse pointer appeared at the screen center along with the instruction text "Click on the picture your partner described" below the images. No feedback was provided on match accuracy and the experiment progressed on to the next trial once participants clicked on an image.

To simulate naturalistic partner behavior, the first two partner utterances were of the form "Click on the <*referent*>"; subsequent utterances included only the article and the noun item ("the <*referent*>"). In addition, the partner audio included a variety of disfluencies (filled pauses, repetitions, prolongations; e.g., "*thee*… the woman who looks angry"). The first three partner descriptions were disfluent, after which approximately 10% of filler descriptions were disfluent, with the probability of a disfluency occurring being higher on facial expression fillers, and decreasing as the experiment progressed.

At the end of the experiment participants were debriefed, during which the false partner interaction was revealed. After this, participants were verbally asked whether they had suspected that they were not interacting with their partner in real time. Only participants who explicitly confirmed that they had believed the interaction was real were included in our analyses.

2.2. Results

2.2.1. Data and analysis

We excluded from our analyses 20 participants who indicated during debrief suspicion about the nature of the interaction. Thus, the final data set consisted of 48 participants (24 per partner condition).

We coded participants' descriptions for overspecification. This was defined as including both the color modifier and the noun item in the description. Trials on which participants produced an invalid description (e.g., cutting off the recording while speaking, or providing hints rather than naming the object) were coded as NA and excluded. These occurred on fewer than 3% of trials in all data sets. Statistical analyses were conducted in R (R Core Team, 2020) using logistics mixed effects regression (Bates, Maechler, Bolker, & Walker, 2014). We analyzed the binary outcome variable of whether or not participants produced an overspecific description or not on each critical trial. We had two main analyses. The first analysis examined the influence of the partner's behavior on participants' overall likelihood of overspecifying. We modeled the outcome variable by the fixed effects of target color (same/different), semantic category (within/across), and partner behavior (overspecific/minimally specific), with all predictors sum-coded. The second analysis examined the influence of the partner's behavior on participants' overspecification rate over time. This was operationalized as trial number, looking only at critical items and ignoring fillers (i.e., the first critical target trial was 0, the second was 1, and so on). We used the same model structure as in the first analysis, with the addition of the fixed effect of trial as a continuous variable (coded such that the intercept was the first critical trial). For all models here as well as in later experiments, we started with a maximal random effects structure (participant and target object intercepts and slopes for all within-subjects predictors; Barr, Levy, Scheepers, & Tily, 2013), removing items progressively if needed until convergence was achieved. Full details for the random effects structure used in each model can be found in our analyses scripts: https://osf.io/sqjv2/?view_only=2bb7627ba87e4e83b60e315bad4b8c75.

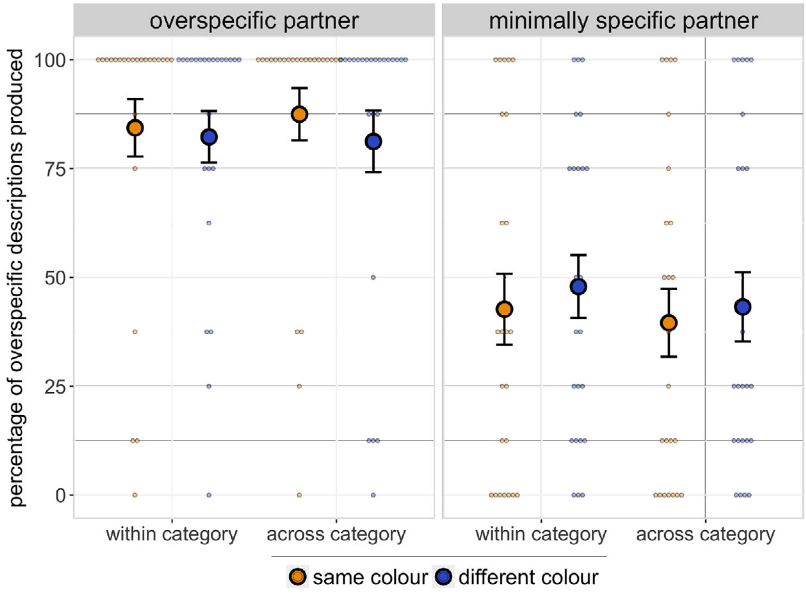

Fig. 2. Mean percentages of overspecific descriptions recorded in each condition with the overspecific partner (left) and minimally specific partner (right). Error bars represent ±1 standard error of by-participant means. Dots represent individual participant proportions.

2.2.2. *Overall overspecification*

Across the 1,534 valid critical descriptions recorded, 977 (63.7%) were coded as overspecific: 644 with the overspecific partner and 333 with the minimally specific partner . Fig. 2 shows the mean percentage of overspecific descriptions recorded in each condition.

There were main effects of partner, showing more overspecification with an overspecific partner, $\beta = 4.95$, $SE = 1.16$, $p < .001$; of color, showing less overspecification when prime and target objects differed in color, $\beta = -0.75$, $SE = 0.37$, $p = .04$; as well as a partner by color interaction, $\beta = -2.24$, $SE = 0.73$, $p = .002$, suggesting a lexical boost effect of color with the overspecific partner.[4] This was confirmed by follow-up analyses revealing less overspecification when prime and target objects differed in color in the overspecific partner condition (i.e., a lexical boost effect of color), $\beta = -5.96$, $SE = 3.68$, $p = .01$, but no difference in the minimally specific partner condition ($p = .6$). There was no effect of category or its interaction with any other predictors; in other words, participants did not overspecify more when prime and target objects were drawn from the same category than from different categories.

2.2.3. *Change in overspecification over time*

Fig. 3 shows the mean percentage of overspecific descriptions recorded in each condition over the course of the experiment.

There were main effects of partner, with more overspecification with an overspecific partner, $\beta = 4.46$, $SE = 1.34$, $p = .001$; of color, with less overspecification when prime and

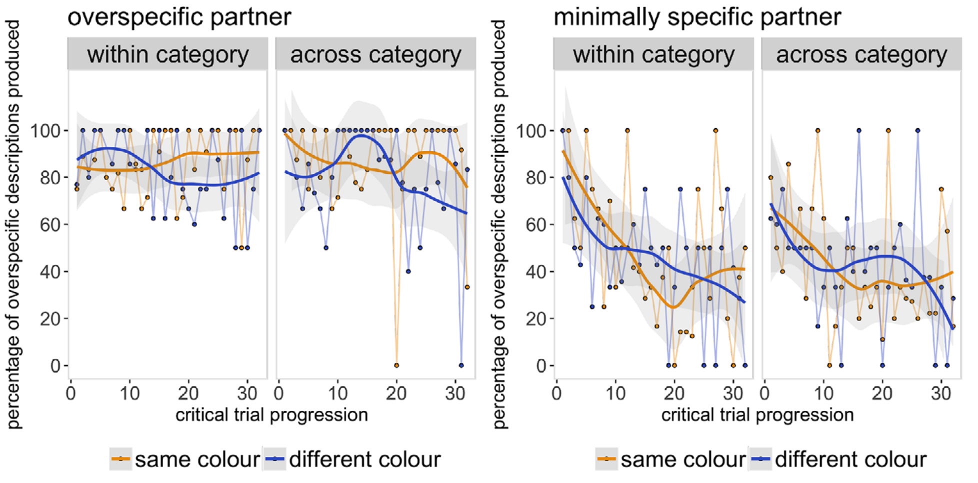

Fig. 3. Mean percentages of overspecific descriptions recorded by critical trial progression. Curved lines represent the data with a loess smoothing curve fitted. Gray ribbons represent 95% confidence intervals around the smooth.

target objects differed in color, $\beta = -2.25$, $SE = 0.69$, $p = .001$; and a partner by color interaction reflecting a lexical boost effect of color with the overspecific partner, $\beta = -3.19$, $SE = 1.16$, $p = .006$. Turning to the rate of overspecification over time, we found a main effect of trial, $\beta = -0.73$, $SE = 0.13$, $p < .001$, indicating that participants were less likely to overspecify as the experiment progressed. Trial number interacted with color, $\beta = 0.75$, $SE = 0.27$, $p = .005$, reflecting a smaller decrease in overspecification when prime and target objects differed in color. Importantly, trial number interacted with partner, $\beta = 0.61$, $SE = 0.25$, $p = .01$, reflecting differences in overspecification over time in the two partner conditions. This was confirmed by follow-up analyses looking at the effect of trial in each partner condition: This revealed a trial effect in both partner conditions, which was larger with the minimally specific partner, $\beta = -1.00$, $SE = 0.14$, $p < .001$ (i.e., greater decrease in overspecification) than the overspecific partner condition, $\beta = -0.60$, $SE = 0.28$, $p = .03$. There was no effect of category or its interaction with any other predictors.

2.3. Discussion

Experiment 1 showed that a speaker's tendency to overspecify is influenced by their partner's linguistic behavior: Speakers were more likely to overspecify by using redundant color modifiers when their partner consistently overspecified than when their partner minimally specified. This demonstrates that, as with many other aspects of linguistic behavior, overspecification is susceptible to the tendency to align with a partner, providing evidence that alignment can manifest in a pragmatic behavior.

The alignment we observed was enhanced when the color modifier was repeated across prime and target trials. This is in line with predictions by Interactive Alignment that lexical repetition at lower levels can enhance alignment at higher levels. However, this could indicate that lexical enhancement extends to alignment on a pragmatic aspect of production—that of overspecification. However, we note that this effect turns out not to be reliable across later experiments, hence we refrain from over-interpreting it here. In addition, we found no evidence for a semantic boost effect of category relatedness between prime and target trials. This appears at odds with Cleland and Pickering (2003), who found a semantic boost effect on syntactic alignment, and may suggest that syntactic encoding is influenced by semantically related concepts in a way that influences on pragmatic behavior are not. A simpler explanation, however, is that the categories we used (clothes vs. kitchenware) were not distinct enough in comparison to those used by Cleland and Pickering (animate vs. inanimate objects), reducing the likelihood for a within-category advantage to emerge.

An unexpected result we found was the direction of trends in alignment over time in the two partner conditions. Although the pattern of alignment differed by partner behavior, rather than increase overspecification with an overspecific partner, speakers in both partner conditions tended to overspecify at the outset, and decreased this behavior dramatically with a minimally specific partner. In other words, speakers appeared to be driven toward minimal specification rather than overspecification by their partner. The fact that overspecification was close to ceiling at the start supports the egocentric view that overspecification, at least with color modification, arises from an intrinsic tendency to encode information that is visually salient or cognizable (e.g., Pechmann, 1989). It is also in line with an addressee-oriented view in which speakers specify properties to facilitate their partner's referential identification, although we note that this seems less likely given the simplicity of our arrays. Importantly, though, it could explain the lack of increase in overspecification with the overspecific partner, since the near-ceiling rates at the start likely limited opportunity for any observable increase in this behavior as the experiment progressed. We note, however, that this should not rule out the possibility that speakers might also align with an overspecific partner over time. This may be more likely in a context where speakers are less prone to overspecification, such as with a different property (e.g., size adjectives). We address this question in Experiment 3.

3. Experiment 2

In order to facilitate online data collection for the remaining experiments, we set out to first replicate the results of Experiment 1 using a web-based paradigm. Thus, Experiment 2 was a replication of Experiment 1 investigating whether speakers are influenced by a partner's linguistic behavior to similarly produce overspecific object descriptions using prenominal color modifiers. We increased the sample size here (and in the remaining experiments) to counteract the issue of reduced control over the experimental situation in web-based paradigms (Reips, 2000). The experiment was conducted online using Amazon's crowdsourcing platform Mechanical Turk (AMT).[5] Preregistration details for the experiment can be found here: https://osf.io/ekcbf

3.1. Method

3.1.1. Participants

One hundred and nine participants took part in the experiment in exchange for $6. Participants were recruited online on AMT with filters set to target workers based in the United

States who have at least a 97% approval rating and 1,000 approved Human Intelligence Tasks (HITs).

3.1.2. Materials and design

The materials and design were identical to those of Experiment 1 with the exception of the partner recordings. We recorded new partner recordings from a female native speaker of Standard Canadian English.

3.1.3. Procedure

This was largely the same as Experiment 1 with minor changes to facilitate web-based presentation. Participants accessed the experiment in their web browser via an external link on AMT's website. The instructions informed them that they would be paired with another AMT worker to complete a picture description and matching task. After accepting the HIT, participants were reminded to enable mic access in their browser for the experiment to work. As in Experiment 1, a "waiting" screen was displayed for 15 s to simulate a wait for a partner worker to sign up to. During the experimental task, match trials were identical to those of Experiment 1. On description trials, participants controlled when their audio was ostensibly being streamed to their partner by clicking "start" and "stop" buttons on screen. A microphone symbol turned orange to indicate when audio streaming was supposedly taking place, and gray to indicate that this had stopped. At the end of the experiment, participants were directed to a posttest questionnaire, which included a link to the debrief explaining the purpose of the experiment and the false partner manipulation. Participants were asked to provide any comments they had on the experiment, with particular emphasis placed on their impression of their partner and the interaction.

3.2. Results

3.2.1. Data and analysis

We excluded data from participants who indicated in the posttest questionnaire suspicion about their partner or the nature of the interaction (10), or that they were a nonnative speaker of English (4). An additional eight participants were excluded due to problems saving audio recordings to the server (audio files were silent). Thus, the final data set consisted of 87 participants (43 and 44 in the overspecific and minimally specific partner conditions, respectively).

We followed the same coding and analysis procedures as in Experiment 1.

3.2.2. Overall overspecification

Across the 2,783 valid critical utterances recorded, 1,550 (55.7%) were coded as overspecific: 1,187 with the overspecific partner and 363 with the minimally specific partner. Fig. 4 shows the mean percentage of overspecific descriptions recorded in each condition.

There was a main effect of partner showing more overspecification with an overspecific partner, $\beta = 7.51$, $SE = 1.06$, $p < .001$. There was also a three-way partner by color by category interaction, $\beta = -1.34$, $SE = 0.65$, $p = .04$, suggesting differences in a lexical boost effect with the overspecific partner depending on category. This was confirmed by

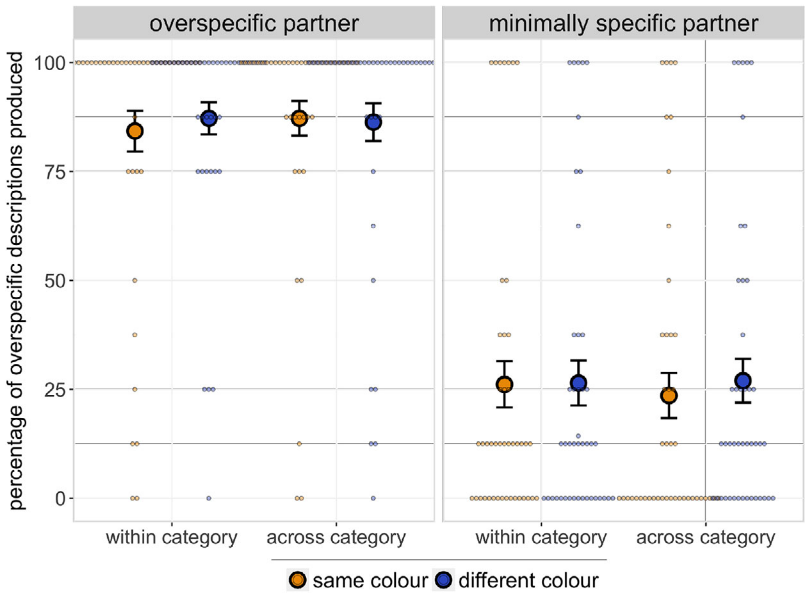

Fig. 4. Mean percentages of overspecific descriptions recorded in each condition with the overspecific partner (left) and minimally specific partner (right). Error bars represent ±1 standard error of by-participant means. Dots represent individual participant proportions.

follow-up analyses indicating a color by category interaction in the overspecific partner condition, indicating a weaker boost effect when prime and target objects were across categories than within the same category, $\beta = -1.08$, $SE = 0.54$, $p = .05$. Separate analyses by category revealed that the effect was marginally significant within categories, $\beta = 0.70$, $SE = 0.40$, $p = .08$, and not significant across categories $\beta = 0.66$, $SE = 1.36$, $p = .6$. Unlike Experiment 1, there was no partner by color interaction suggesting a lexical boost effect (or otherwise) in the overspecific partner condition. Although the figure for the overspecific partner condition suggests slightly less overspecification when prime and target objects were different compared to the same color, separate analyses by partner revealed no effect of color in either partner condition (all $p > .1$).

3.2.3. *Change in overspecification over time*

Fig. 5 shows the mean percentage of overspecific descriptions recorded in each condition over the course of the experiment.

There were main effects of partner, with more overspecification with an overspecific partner, $\beta = 6.30$, $SE = 1.03$, $p < .001$; and of trial number, showing an overall decrease in overspecification as the experiment progressed, $\beta = -0.19$, $SE = 0.09$, $p = .03$. Trial number interacted with partner, reflecting a smaller decrease in overspecification over time with the overspecific partner, $\beta = 0.70$, $SE = 0.18$, $p < .001$. As before, this was confirmed by follow-up analyses which showed that overspecification decreased across trials

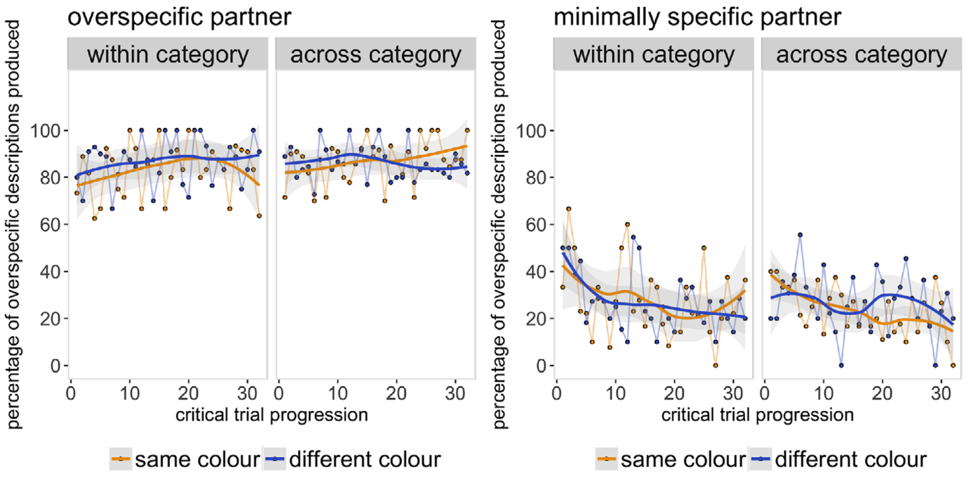

Fig. 5. Mean percentages of overspecific descriptions recorded by critical trial progression. Curved lines represent the data with a loess smoothing curve fitted. Gray ribbons represent 95% confidence intervals around the smooth.

with the minimally specific partner ($\beta = -0.57$, $SE = 0.11$, $p < .001$), but did not change significantly with the overspecific partner ($\beta = 0.17$, $SE = 0.15$, $p = .2$).

3.3. Discussion

The results of Experiment 2 replicate the partner effect from Experiment 1. Speakers were more likely to overspecify with an overspecific than a minimally specific partner, and this behavior decreased across trials to demonstrate alignment over time with the minimally specific partner. Unlike Experiment 1, we found little evidence that alignment with the overspecific partner was enhanced by lexical repetition of color across prime and target trials. Although a color by category interaction suggested differences in the boost effect modulated by semantic category, further analyses showed that this was only marginally significant within categories, and not significant across categories.

4. Experiment 3

Experiment 3 investigated whether the alignment we observed in Experiments 1 and 2 extends to another property that is less prone to overspecification—that of size. Previous research on overspecification in noninteractive contexts has demonstrated systematic differences in speakers' tendencies to overspecify across different properties. In particular, color adjectives stand out as being frequently overspecified. The same is not true for other adjectives such as size, material, or pattern, where far lower rates of overspecification are observed (Koolen et al., 2013; Pechmann, 1989; Tarenskeen et al., 2015), with an exception being Van Gompel, Gatt, Krahmer, and Van Deemter (2014), who showed that the color preference over size can disappear with the right context, such as a large enough size contrast. This

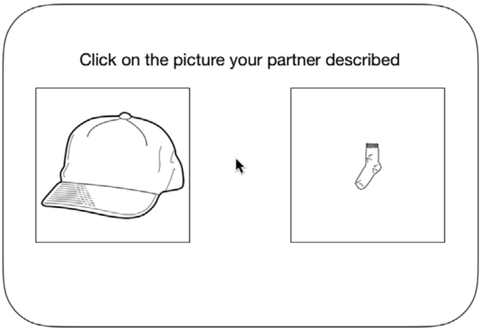

Fig. 6. Example of a prime trial display featuring a big cap and a small sock.

asymmetry has been attributed to the “special status” of color adjectives due to its visual salience and therefore contrastive properties (Tarenskeen et al., 2015), and may in turn explain the high rates of overspecification we found in Experiments 1 and 2, especially with the overspecific partner. We should therefore expect a lower preexisting bias in speakers toward overspecification of size, which allows us to test if this behavior increases during interaction with an overspecific partner.

Thus, in Experiment 3 we replaced the color variation with size, using black and white images with two size variants—big and small. Based on previous research, we would expect to find overall lower rates of overspecification. Based on results from Experiments 1 and 2 we would expect speakers to align with their partner’s behavior and thus expect to see overspecification increase with an overspecific partner. An absence of such a partner effect could suggest that speakers are only influenced toward minimal specification and not overspecification. Preregistration details for the experiment can be found here: https://osf.io/ytvew.

4.1. Method

4.1.1. Participants

One hundred and ten participants took part in the experiment in exchange for $6. Participants were recruited online on AMT with filters set to target workers based in the United States who have at least a 97% approval rating and 1,000 approved HITs. Participants who had taken part in Experiment 2 were excluded from taking part.

4.1.2. Materials

For the critical images, we used black and white versions of the same eight noun items as before. Each of these occurred in two size variants—big and small. The ratio of the heights of the big and small variants was 3:1. Fig. 6 shows an example of a trial display featuring a big and a small object. Each object was associated with voice recordings that specified the noun item and size modifier (e.g., “the small sock”), as well as only the bare noun (e.g., “the sock”). The recordings were produced by the same speaker who produced the recordings for Experiment 2.

For fillers, the size of the natural object and facial expression images were adjusted to be in between the big and small variants of the critical images. All other aspects of the filler stimuli remained the same.

4.1.3. Design

The design of the experiment was the same as Experiment 2 with the exception that the color manipulation was replaced by size. As before, this had two levels — same versus different across prime and target trials.

4.1.4. Procedure

This was identical to Experiment 2.

4.2. Results

4.2.1. Data and analysis

We excluded data from participants who indicated in the posttest questionnaire suspicion about their partner or the nature of the interaction (seven), or that they were a nonnative speaker of English (six). An additional three participants were excluded due to problems saving audio recordings to the server. Thus, the final data set consisted of 94 participants (46 and 48 in the overspecific and minimally specific partner conditions respectively).

We followed the same coding and analysis procedures as in Experiment 1.

4.2.2. Overall overspecification

Across the 2926 valid critical utterances recorded, 1,221 (41.7%) were coded as overspecific: 952 with the overspecific partner and 269 with the minimally specific partner. Fig. 7 shows the mean percentage of overspecific descriptions recorded in each condition.

There was a main effect of partner with more overspecification with the overspecific partner, $\beta = 3.58$, $SE = 0.46$, $p < .001$. None of the other predictors or their interactions were significant (all $p > .1$).

4.2.3. Change in overspecification over time

Fig. 8 shows the mean percentage of overspecific descriptions recorded in each condition over the course of the experiment.

There were main effects of trial, with overspecification increasing as the experiment progressed, $\beta = 0.29$, $SE = 0.14$, $p = .04$, and of partner, with more overspecification with the overspecific partner, $\beta = 2.01$, $SE = 0.45$, $p < .001$. Trial number interacted with partner, $\beta = 1.57$, $SE = 0.32$, $p < .001$, suggesting that the pattern of overspecification varied depending on partner behavior. This was confirmed by follow-up analyses revealing that overspecification increased over time with the overspecific partner, $\beta = 0.75$, $SE = 0.09$, $p < .001$, but did not change significantly with the minimally specific partner, $\beta = 0.05$, $SE = 0.09$, $p = .6$. There was also a trial by category interaction, $\beta = 0.29$, $SE = 0.14$, $p = .03$. Follow-up analyses showed that this was driven by an increase in overspecification across trials that was greater when prime and target objects were from different categories,

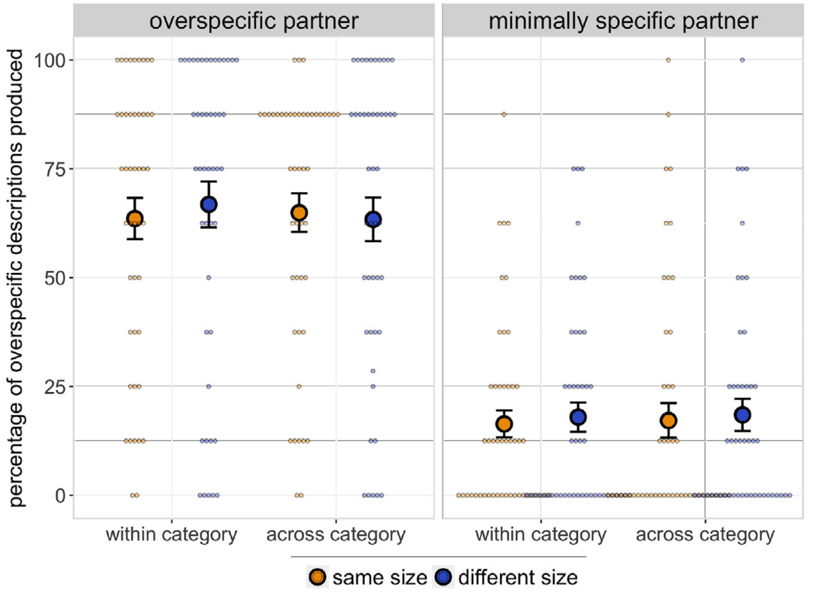

Fig. 7. Mean percentages of overspecific descriptions recorded in each condition with the overspecific partner (left) and minimally specific partner (right). Error bars represent ±1 standard error of by-participant means. Dots represent individual participant proportions.

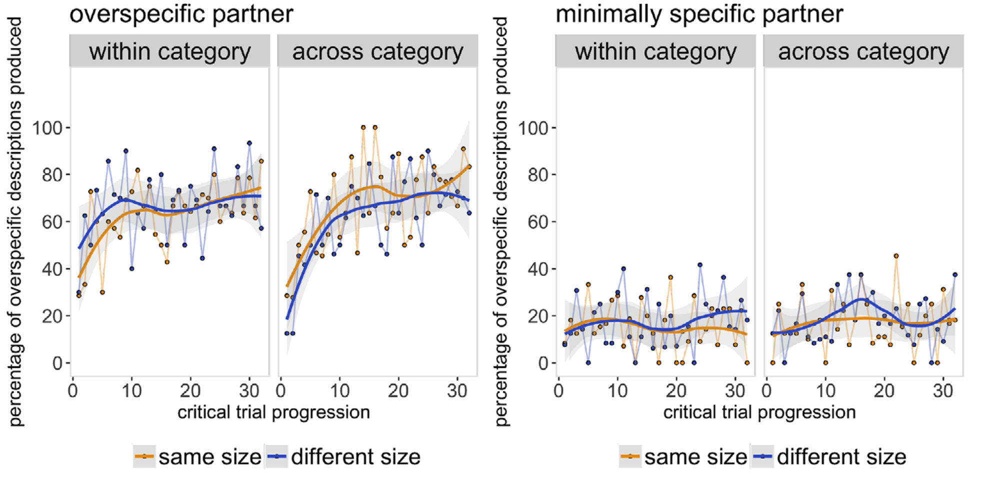

Fig. 8. Mean percentages of overspecific descriptions recorded by critical trial progression. Curved lines represent the data with a loess smoothing curve fitted. Gray ribbons represent 95% confidence intervals around the smooth.

$\beta = 0.78$, $SE = 0.36$, $p = .03$, than when they were from the same category, $\beta = 0.37$, $SE = 0.27$, $p = .2$.

4.3. Discussion

The results of Experiment 3 confirm our findings from Experiments 1 and 2 that a speaker's tendency to overspecify is influenced by their partner's linguistic behavior: Speakers are more likely to overspecify if their partner is also overspecific, with the effect increasing over the course of the interaction. Furthermore, the experiment demonstrates that this alignment is not limited to color modification but extends to size, which speakers are typically less prone to overspecify.

In line with earlier studies on color and size overspecification, we observed lower overall rates of overspecification here compared to Experiments 1 and 2. However, our overspecification rates are notably still higher than those of previous studies (e.g., Belke & Meyer, 2002; Tarenskeen et al., 2015, Experiments 1–3). This demonstrates that given the right context, for example, a partner who uses size redundantly, size overspecification can be triggered in speakers (cf. Tarenskeen et al., 2015; Van Gompel et al., 2014, Experiment 4). We saw no evidence of more overspecification in the same-size condition with the overspecific partner to suggest a lexical boost effect with size modification. This is in contrast to Experiment 1, which showed a lexical boost with color modification, and may suggest different modulating effects of lexical repetition with size and color adjectives on alignment. However, we note that the lexical boost effect in Experiment 1 was relatively weak—in particular the effect was small compared to that of partner behavior. In Experiment 2 the lexical boost was only within categories and only marginally significant then. Thus, collectively we have little evidence for a lexical boost effect on alignment in overspecification.

Across Experiments 1–3, one striking result was the time course of alignment with color and size overspecification. Comparing the two properties, we see that speakers tended to overspecify color from the outset, with alignment over time more pronounced with a minimally specific partner; for size modification we saw the opposite, with low initial rates of overspecification that increased only with the overspecific partner. This result likely reflects different baseline tendencies with respect to encoding of the two properties: Speakers appear to readily overspecify color, whereas this seems to be less the case with size. This pattern is consistent with reports from previous studies (Belke & Meyer, 2002; Pechmann, 1989), and highlights the role of a property's salience and absoluteness in triggering whether or not a speaker includes it (Tarenskeen et al., 2015).

One unexpected finding was an increase in overspecification over time when prime and target objects were across categories, but not within. This is at odds with previous studies on semantic boost effects, which suggest a within-category advantage in the alignment of syntactic structure (Cleland & Pickering, 2003) as well as modifier choice (Heller & Chambers, 2004). However, we note that these studies did not examine the effect of such category effects across trials. The current results also contrast with those of Experiments 1 and 2, which found no effect of category on alignment. One explanation could be that replacing the color manipulation, which was likely very salient in our earlier experiments, with size here made the category difference stand out more for participants. However this is speculation, and a more systematic manipulation of color, size, and category, in conjunction with their effects over time, would be needed to tease apart these variables.

Although we observed qualitatively different trends over time with color (Experiments 1 and 2) and with size (Experiment 3), the results provide converging evidence that alignment increases the more speakers are exposed to the relevant behavior from their partner. This suggests that partner effects of alignment in overspecification are not static but evolve over time as an interaction unfolds. It is important to consider, however, that this pattern of alignment could also be explained by cumulative priming in the absence of any extended partner influence. That is, speakers may have been initially biased by their partner's linguistic behavior to overspecify or minimally specify, but subsequently influenced by their own prior productions to continue doing so (cf. Bock, 1986). This may have been the case regardless of their partner's subsequent productions. This would be in line with predictions within the framework of cumulative priming, which relies on speaker-internal tendencies toward repeating recently heard utterances (Kaschak, 2007; Kaschak et al., 2011). Consequently, we cannot say definitively that our partner effects extended across the duration of the interaction. Experiments 4 and 5 were designed to address this question.

5. Experiment 4

Experiment 4 investigated what effect (if any) a change in the partner's linguistic behavior midway through the interaction has on the pattern of alignment in overspecification. As before, participants completed a picture description and matching task in which they alternated between describing an image for their partner, and matching their partner's description to the correct image. In the first half of the experiment, all speakers interacted with a partner who only produced overspecific prime descriptions; in the second half, the *consistent* partner continued to overspecify, while the *variable* partner switched descriptive strategies to that of minimal specification. If overspecification in speakers is indeed sensitive to partner effects over time, we should see a change in the pattern of overspecification with the variable partner in the second half reflecting this shift in the partner's linguistic behavior. Conversely, if the effects over time are simply due to accumulated speaker-internal effects rather than any influence of partner behavior, we should see no difference in the time course between the two partner conditions since speakers should continue with the strategy they had established from the first half. Preregistration details for the experiment can be found here: https://osf.io/taecd.

5.1. Method

5.1.1. Participants

One hundred and twenty-four participants took part in the experiment in exchange for $6. Participants were recruited online on AMT with filters set to target workers based in the United States who have at least a 97% approval rating and 1,000 approved HITs. Participants who had taken part in Experiments 2 and 3 were excluded from taking part.

5.1.2. Materials and design

We used the same visual and audio stimuli as in Experiment 2 (i.e., we manipulated color rather than size).

The length of the experiment remained the same (256 trials), but now consisted of two consecutive blocks of 128 trials each. The 32 critical items were organized such that 16 occurred in the first block and the other 16 in the second block. Within each block, each of the 32 unique color–noun item combinations appeared once as a prime and once as a target referent. The experiment manipulated partner behavior between subjects. In the consistent partner condition, speakers interacted with a partner who consistently produced overspecific descriptions on all critical prime trials; this condition is identical to the overspecific partner condition in Experiment 1 with the exception that here the number of critical items and the color manipulation were balanced across the two blocks (see below). In the variable partner condition, the partner produced overspecific prime descriptions in the first block and minimally specific prime descriptions in the second block. The same recordings were used for the first block in both partner conditions. In addition, we also manipulated color (same vs. different for prime and target) within subjects: Half of the critical items within each block featured prime and target referents that shared the same color, while the other half featured referents of different colors.

The order of presentation of items was randomized on each run with the constraints that at least six filler trials (i.e., three participant match–describe sequences) preceded the first critical item in each block, and at least two filler trials occurred between critical items during the experiment.

5.1.3. Procedure

The procedure of the experiment remained the same. Other than the change in the partner's linguistic behavior in the variable partner condition, there was nothing to indicate a transition between blocks from the participant's perspective.

5.2. Results

5.2.1. Data and analysis

We excluded data from participants who indicated in the posttest questionnaire suspicion about their partner or the nature of the interaction (nine), or that they were a nonnative speaker of English (four). An additional seven participants were excluded due to problems saving audio recordings to the server. Thus, the final data set consisted of 104 participants (50 and 54 in the consistent and variable partner conditions, respectively).

We followed the same coding procedure as in the earlier experiments. To evaluate the effect of a change in the partner's linguistic behavior on participants' overspecification, we analyzed the data from the second block of the experiment, in which the partner had either changed their descriptive strategy (variable partner condition) or continued with the same strategy (consistent partner condition). We constructed two models—the first examining overall overspecification and the second examining overspecification over time. For overall overspecification, we included target color (same vs. different) and partner (consistent vs. variable) as fixed effects, with both predictors sum-coded. For overspecification over time, we used the same model structure with the addition of trial number (counting only critical trials and coded such that the intercept was the first critical trial of the block) as a predictor.

We also conducted a preliminary analysis of the data from block 1 of the experiment, when the partner's linguistic behavior was the same in the two conditions. This analysis was identical to the analysis on the second block, and simply served as a sanity check to verify that participants' behavior did not differ before partner behavior diverged. Thus, we expected participants to behave similarly in the two partner conditions here, and any differences can only be attributed to variation in the sampled participants. We first report the results of the preanalysis (on block 1), followed by the results for overall overspecification and overspecification over time for the second block.

5.3. *Preanalysis of block 1*

We first checked whether participants behaved similarly in the two partner conditions in the first block, where both partners only produced overspecific prime descriptions.

Across the 1,651 valid target descriptions recorded, 1,431 (86.7%) were coded as overspecific: 669 with the consistent partner and 762 with the variable partner. These percentages are comparable with those of the overspecific partner condition in Experiments 1 and 2. There was no effect of partner on participants' overall likelihood of overspecifying ($p = .96$), suggesting that participants' rate of overspecification was similar across the two partner conditions in block 1. There was also no effect of color ($p = .6$), nor did color interact with partner ($p = .2$).

The analysis over time showed a main effect of trial, $\beta = 3.65$, $SE = 1.49$, $p = .01$, indicating that overspecification increased as the experiment progressed. There was no effect of partner or its interaction with any other predictors (all $p \geq .1$), indicating that participants' behavior was similar across the two partner conditions in block 1.

5.3.1. *Overall overspecification in block 2*

Our main analysis focused on how a change in the partner's linguistic behavior influenced overspecification in the second block, where the two partners diverged in their behavior. Across the 1,653 valid critical descriptions recorded in this block, 1,034 (62.6%) were coded as overspecific: 655 with the consistent (overspecific) partner and 379 with the variable partner. Fig. 9 shows the mean percentage of overspecific descriptions recorded in each condition.

There was a main effect of partner in this block, $\beta = -5.83$, $SE = 1.04$, $p < .001$, showing less overspecification with the variable (minimally specific) partner than the consistent (overspecific) partner. There was also a marginal effect of color, $\beta = -0.69$, $SE = 0.39$, $p = .08$, showing more overspecification when prime and target objects differed in color. However, color did not interact with partner ($p > .1$), providing no evidence for a lexical boost effect of color.[6]

5.3.2. *Change in overspecification over time in block 2*

Fig. 10 shows the mean percentage of overspecific descriptions recorded in each condition over the course of the experiment.

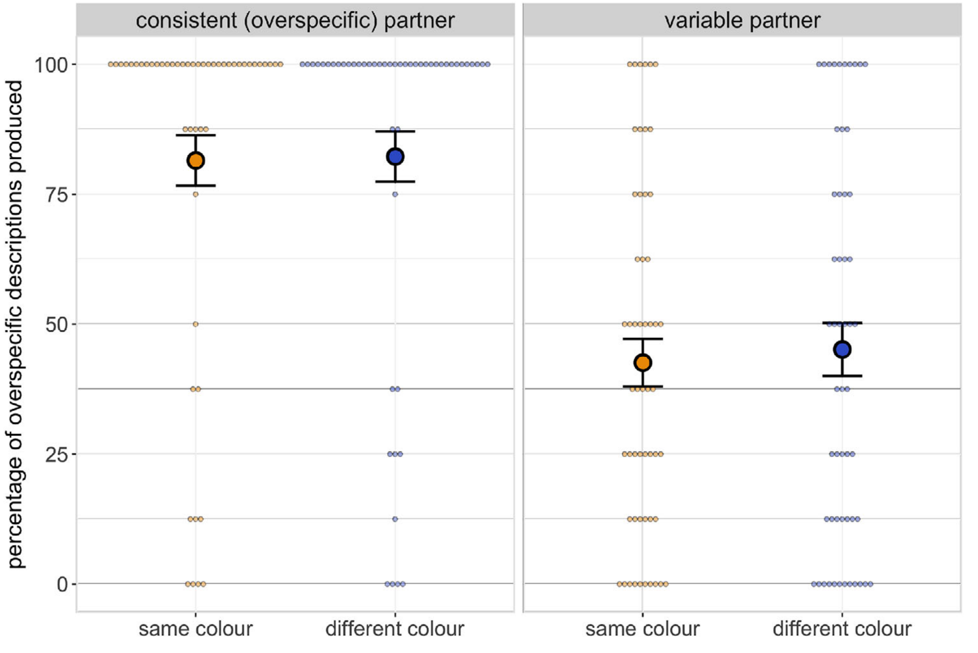

Fig. 9. Mean percentages of overspecific descriptions recorded in each condition with the consistent (overspecific) partner (left) and variable (overspecific–minimally specific) partner (right) in the second block of the experiment. Error bars represent ±1 standard error of by-participant means. Dots represent individual participant proportions.

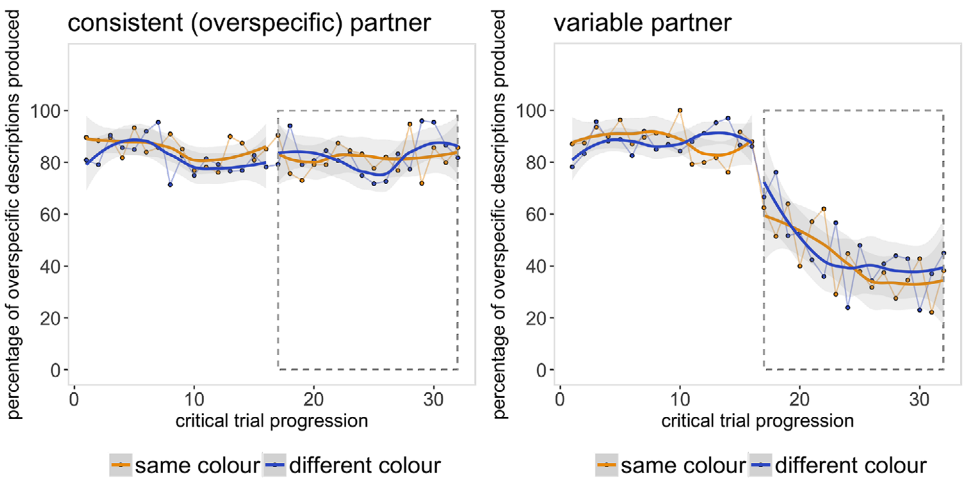

Fig. 10. Mean percentages of overspecific descriptions recorded over critical trial progression. Curved lines represent the data with a loess smoothing curve fitted (curves were fit separately for each block to better illustrate trends within a block). Gray ribbons represent 95% confidence intervals around the smooth. The dotted lines indicate critical trials, which occurred in the second block of the experiment (trials 128–256), where partner behavior diverged.

There was a main effect of partner, with less overspecification with the variable partner, $\beta = -4.54$, $SE = 1.26$, $p < .001$, and a partner by trial interaction $\beta = -2.93$, $SE = 1.12$, $p = .009$, indicating different trends in participants' tendency to overspecify depending on their partner's behavior. This was confirmed by follow-up analyses, which revealed that overspecification decreased over time with the variable partner, $\beta = -1.72, SE = 0.38, p < .001$, but increased over time with the consistent partner, $\beta = 0.93$, $SE = 0.48$, $p = .05$.

5.4. Discussion

Experiment 4 corroborates our main result from earlier experiments that overspecification behavior is susceptible to alignment with one's partner in dailog. In addition, we found that this partner effect was sensitive to a change in the partner's linguistic behavior, as evidenced from the patterns of overspecification we observed across the first and second blocks. We observed similarly high rates of overspecification here with both partners in the first block as with the overspecific partner in Experiments 1 and 2, reflecting the influence of a partner who consistently overspecified. However, a difference emerged in the second block as partner behavior diverged, such that overspecification decreased with the variable partner, who stopped overspecifying in this block. Notably, this result emerged after a period of interaction in which speakers established a tendency to overspecify with their partner (block 1), indicating a shift in speakers' referential strategies to align with the change in their partner's. The fact that alignment was tracked across a change that occurred even as the interaction was underway is indicative of long-lasting effects of a partner's behavior on speakers' productions. Moreover, this effect continued to increase as the rest of the interaction unfolded, highlighting the cumulative nature of a partner's influence on alignment.

Unlike Experiments 1 and 2, we found no evidence that the lexical repetition of color enhanced alignment with the overspecific partner here. This weakens our evidence for a lexical boost effect of color on alignment in overspecification.

6. Experiment 5

Having established that speakers are influenced by a change in their partner's linguistic behavior to switch from overspecifying to minimally specifying, we turned our attention to whether they would be similarly influenced to switch to overspecifying. Previous research has implied overspecification to be a suboptimal speaker strategy as it involves encoding more information than is necessary, and may at times even impede communication (e.g., Engelhardt, Bailey, & Ferreira, 2006). Thus, it is perhaps unsurprising that we found that speakers continued to align with a partner whose referential strategy switched to minimal specification, which may have been seen as more optimal from a production perspective. Such a view also implies that we might expect an asymmetry in which speakers adjust their descriptive strategy to continue to align with a partner who switches to minimal specification, but do not do so with a partner who switches to overspecification. Indeed, recent research implies an asymmetry in speakers' tendencies to modify their behavior toward overspecification or minimal specification in response to changes in the communicative context: Long, Rohde, and Rubio-Fernandez (2020) observed that the tendency for speakers to start overspecifying following a change from monochrome to polychrome displays was smaller than the tendency for speakers to stop overspecifying when going from polychrome to monochrome displays. This may suggest that the tendency to switch to a minimally specific strategy may be greater than that of switching to an overspecific strategy.

Thus, Experiment 5 tested whether we would see alignment over time with a partner whose linguistic behavior shifted from minimal specification to overspecification. As in Experiment 4, speakers completed a picture description and matching task with their partner across two blocks. However, here all speakers were exposed to a minimally specific partner in the first block; in the second block, the *consistent* partner continued to minimally specify, while the *variable* partner switched to an overspecific descriptive strategy. If speakers align with their partner to switch from minimal to overspecification, we should see this reflected in a difference in overspecification trends between the two partner conditions. Preregistration details for the experiment can be found here: https://osf.io/9zja3/.

6.1. Method

6.1.1. Participants

One hundred and twenty-six participants took part in the experiment in exchange for $6. Participants were recruited online on AMT with filters set to target workers based in the United States who have at least a 97% approval rating and 1,000 approved HITs. Participants who had taken part in Experiments 2–4 were excluded from taking part.

6.1.2. Materials and design

We used the same materials and design as Experiment 4 (i.e., we manipulated color again) with the only change being that to the partner's behavior. Both partners minimally specified in the first block; in the second block, the consistent partner continued to do so while the variable partner switched descriptive strategies to overspecification.

6.1.3. Procedure

This was identical to Experiment 4.

6.2. Results

6.2.1. Data and analysis

We excluded data from participants who indicated in the posttest questionnaire suspicion about their partner or the nature of the interaction (nine), or that they were a nonnative speaker of English (one). An additional 14 participants were excluded due to problems saving audio recordings to the server. Thus, the final data set consisted of 102 participants (51 each in the consistent and the variable partner conditions).

We followed the same coding and analysis procedures as in Experiment 4.

6.2.2. Preanalysis of block 1

We first checked whether participants behaved similarly in the two partner conditions in the first block, where both partners always produced minimally specific critical prime descriptions. Across the 1,610 valid target descriptions in this block, 565 (35.1%) were coded as overspecific: 279 with the consistent partner and 286 with the variable partner. These percentages are comparable with those of the minimally specific partner condition in Experiments 1 and 2. There was no effect of partner ($p = .6$), color ($p = .9$), or an interaction between the two ($p = .6$), suggesting that participants' overall rate of overspecification was similar across the two partner conditions.

As for overspecification over time, there were main effects of trial, $\beta = -2.44$, $SE = 0.40$, $p < .001$, with overspecification decreasing over time; and of color, $\beta = -0.85$, $SE = 0.32$, $p = .008$, with less overspecification when prime and target objects differed in color. However, color did not interact with partner, providing no evidence for a lexical boost effect of color. There was a trial by color interaction, $\beta = 1.65$, $SE = 0.43$, $p = .001$, reflecting a greater decrease in overspecification when prime and target objects shared the same color. Trial number also marginally interacted with partner, $\beta = 1.33$, $SE = 0.74$, $p = .07$, possibly suggesting partner-dependent differences in overspecification. Visual inspection of the time-course plot suggested that this difference largely centered around the first few trials (see Fig. 12). To follow up on this we repeated the analysis after splitting the data set into trials that occurred in the first half (trials 1–64) and the second half (trials 65–128) of the block. This revealed that the partner effects were driven by a trial by color by partner interaction in the first half, $\beta = -5.21$, $SE = 2.51$, $p = .04$, reflecting a greater decrease in overspecification with the consistent partner, in particular when prime and target objects shared the same color. There was no comparable effect of partner or any of its higher order interactions in the second half (all $p > .1$), suggesting that overspecification behavior was similar across the two partner conditions by the later part of the first block.

6.2.3. Overall overspecification in block 2

Our main analysis focused on the effect of a change in partner behavior on speakers' tendency to overspecify in the second block, where the two partners diverged in their behavior. Across the 1,599 valid critical descriptions recorded in this block, 543 (34.0%) were coded as overspecific: 168 with the consistent (minimally specific) partner and 375 with the variable partner. Fig. 11 shows the mean percentage of overspecific descriptions recorded in each condition.

There was a main effect of partner in this block, $\beta = 3.29$, $SE = 0.96$, $p < .001$, showing more overspecification with the variable (overspecific) partner than the consistent (minimally specific) partner. There was also a main effect of color, $\beta = -0.77$, $SE = 0.31$, $p = .01$, showing less overspecification when prime and target objects differed in color; however, this did not interact with partner ($p > .1$), providing no evidence for a lexical boost effect of color.

6.2.4. Change in overspecification over time in block 2

Fig. 12 shows the mean percentage of overspecific descriptions recorded in each condition over the course of the experiment.

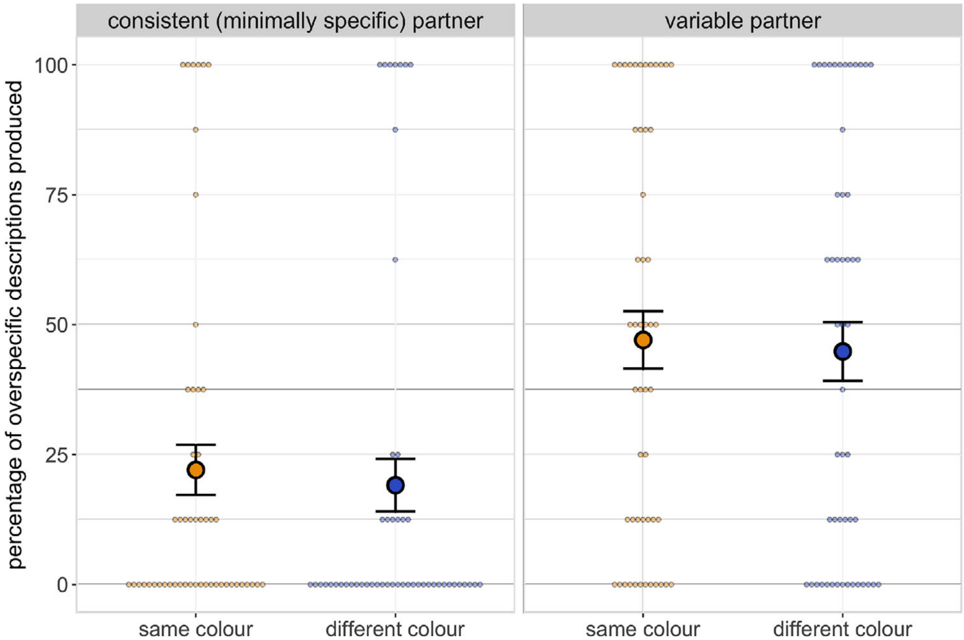

Fig. 11. Mean percentages of overspecific descriptions recorded in each condition with the consistent (minimally specific) partner (left) and variable (minimally specific–overspecific) partner (right) in the second block of the experiment. Error bars represent ±1 standard error of by-participant means. Dots represent individual participant proportions.

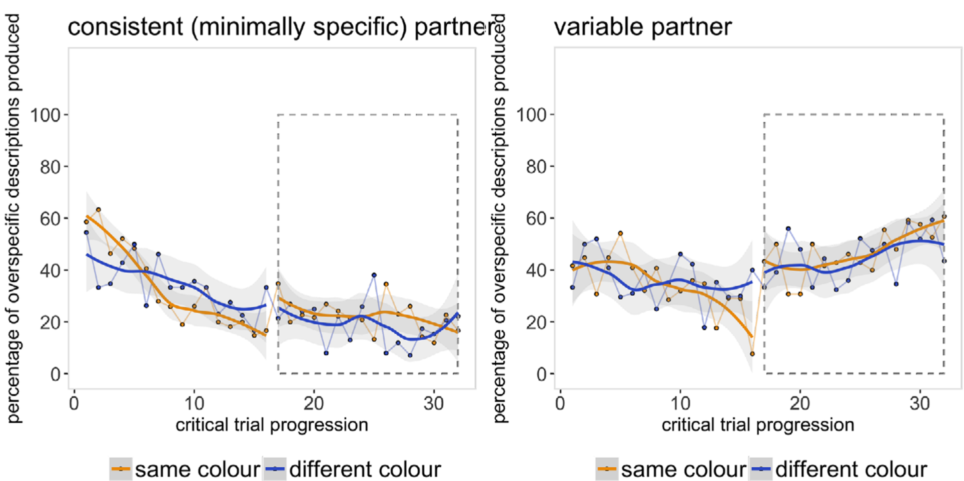

Fig. 12. Mean percentages of overspecific descriptions recorded over critical trial progression. Curved lines represent the data with a loess smoothing curve fitted (curves were fit separately for each block to better illustrate trends within a block). Gray ribbons represent 95% confidence intervals around the smooth. The dotted lines indicate critical trials, which occurred in the second block of the experiment (trials 128–256), where partner behavior diverged.

There was a main effect of color, $\beta = -1.33$, $SE = 0.48$, $p = .006$, indicating less overspecification when prime and target objects differed in color. There was no overall main effect of partner ($p = .1$); however, there was a trial by partner interaction, $\beta = 3.22$, $SE = 0.81$, $p < .001$, reflecting different patterns of overspecification in the two partner conditions. This was confirmed by follow-up analyses, which showed that overspecification decreased over time with the consistent partner, $\beta = -1.17$, $SE = 0.37$, $p = .002$, but increased over time with the variable partner, $\beta = 1.55$, $SE = 0.51$, $p = .002$.

6.3. Discussion

Experiment 5 replicates the finding from Experiment 4 that alignment in overspecification is tracked across a change in the partner's linguistic behavior. This strengthens our evidence for long-lasting partner effects across the duration of an interaction. In addition, we showed that this alignment extends to a change from minimal specification to overspecification. This is evident from the pattern of overspecification across the two blocks in the variable partner condition, which showed an increase in this behavior when the partner began to overspecify. In other words, speakers were influenced by their partner to adjust their referential strategy to one that we may expect to be dispreferred under a traditional view of overspecification (e.g., Grice, 1975). We discuss the implications of this in Section 7.

6.4. Comparison of Experiments 4 and 5

Experiments 4 and 5 established that speakers align with a change in their partner's referential behavior, regardless of whether the change is from overspecification to minimal specification or vice versa. One question that follows is whether they are equally likely to adjust their strategy in both partner conditions (cf. Long et al., 2020).

To address this, we conducted an additional analysis comparing the behavior of speakers in the second block of the variable partner conditions in Experiments 4 and 5. We used logistic mixed effects regression to model the outcome variable of whether or not speakers aligned with their partner's strategy, with fixed effects of trial number (coded such that the intercept was centered on the first critical trial of the block) and experiment (sum-coded). This analysis revealed a main effect of trial number $\beta = 1.60$, $SE = 0.28$, $p < .001$, indicating that alignment increased as the experiment progressed. However, there was no effect of experiment or a trial by experiment interaction (all $p > .4$), suggesting that speakers were equally likely to alter their behavior to align with a partner who switched from overspecification to minimal specification or vice versa.

7. General discussion

In this paper, we examined the role of a partner's behavior in overspecification on a speaker's own tendency to overspecify. We asked whether overspecification is susceptible to alignment with one's partner, whether this alignment is sensitive to lexical and/or semantic boost effects, and how the influence of partner behavior persists across the duration of the interaction. Here, we address each of these points in turn in relation to our findings across all experiments.

7.1. Is overspecification susceptible to alignment?

In all five experiments, our results showed that overspecification is influenced by a partner's linguistic behavior: Speakers were more likely to overspecify when interacting with a partner who overspecified compared to one who minimally specified. This corroborates findings from previous studies, which show partner dependence on overspecification in the production of nominal referring expressions (Brennan & Clark, 1996), as well as the encoding of dispreferred object attributes (Goudbeek & Krahmer, 2012). Together, these results provide converging evidence that speakers adjust their referential strategy during interaction to align with their partner in producing more information than is necessary. In other words, speakers appear to align with their partner in a pragmatic behavior—that of overspecification. We note that our results do not rule out the possibility of an underlying structural explanation to the alignment, since the partner's use of an overspecified or a minimally specified expression would simultaneously make that structure more accessible for speakers. Thus, speakers may be mirroring their partner's form of description (modified vs. bare noun), as opposed to a higher level pragmatic decision of whether or not to overspecify. Nevertheless, the observed alignment demonstrates that a pragmatic phenomenon (tendency to overspecify) is sensitive to a partner's linguistic behavior. This opens up questions for future research on the nature of alignment in pragmatic processes, and how this may interact with other levels of linguistic representation.

Our participants aligned with their partner on both color and size overspecification, demonstrating that the effect is not limited to overspecification in a single dimension. Notably, however, speakers did not behave similarly across the two dimensions. In particular, we observed a stronger tendency for speakers to overspecify color, a pattern in line with the finding in the literature that speakers often encode color adjectives redundantly and preferentially over other properties (Koolen, Gatt, Goudbeek, & Krahmer, 2011; Pechmann, 1989). This likely explains the lack of observable alignment over time with the overspecific partner in Experiments 1 and 2, as overspecification rates were already near ceiling from the start. Compared to color modification, we observed more modest rates of overspecification with size. However, it is worth noting that these were still higher than previous studies reporting size overspecification in monologue contexts (e.g., Belke & Meyer, 2002; Tarenskeen et al., 2015, Experiments 1–3). This highlights the relevance of communicative context, in particular the role of interaction and its associated partner-dependent factors, in affecting overspecification. More generally, it is also in line with studies that indicate measurable differences in speaker and listener behavior between interactive and non-interactive contexts (e.g., Fox Tree, 1999; Pivneva, Palmer, & Titone, 2012; Schober & Clark, 1989).

Differences in predispositions to overspecify in the two dimensions also led to different trends in alignment over time. We found that speakers showed a tendency to align with a minimally specific partner over time with color modification, but an overspecific partner over time with size modification. These provide further evidence of differences in speakers' preexisting biases toward color and size overspecification. However, Experiment 5 showed that speakers can, under the right conditions, be induced to increase color overspecification with their partner—in this case, a change in their partner's linguistic behavior after interlocutors aligned on minimal specification prompted this behavior. This demonstrates that alignment is a flexible and dynamic process that adapts to global changes in the interactional context.

7.2. Is alignment sensitive to lexical and/or semantic effects?

Turning to the effect of lexical and semantic repetition on alignment, our results present a mixed picture. Overall, we found little evidence for a lexical boost on alignment in overspecification, and none for a semantic boost. Although Experiments 1 and 2 showed some evidence of a lexical boost with color (significant effect in Experiment 1; marginally significant and only within-categories in Experiment 2), this was not observed in Experiments 4 and 5, or in Experiment 3 with size repetition.

We note that the lexical boost was tested here by manipulating the color of a target object between prime and target trials, which differs somewhat from syntactic alignment studies, which typically manipulate the repetition of a verb that speakers have to use (e.g., Branigan et al., 2000). This also differs from the semantic boost manipulation, which we tested by manipulating a category overlap status between the prime and target objects. Such manipulations may be too nuanced for clear differences to emerge in a phenomenon such as over/minimal specification, in particular for the lexical boost, which depends on the decision of whether or not to encode an (optional) adjective.

As noted in Section 1, it is not our primary aim to distinguish between the possible mechanisms underlying alignment here, however, it is worth acknowledging that our results can be informative for theories of alignment. In particular, the lexical and semantic effects we observed are largely inconsistent with predictions by the Interactive Alignment model, which posits that alignment at lower levels should enhance the effect at higher levels—an effect well documented in the syntactic alignment literature. It is possible, however, that alignment in a pragmatic behavior such as overspecification may be more encapsulated from such effects. On the other hand, the lack of a lexical boost effect in Experiments 3–5 could also be due to a number of reasons. First, the smaller number of size compared to color modifiers used (two vs. four) meant that both size adjectives would have been highly activated for speakers in Experiment 3, reducing the likelihood of a boost effect. As for Experiments 4 and 5, the analyses were conducted on half the number of items compared to earlier experiments, weakening the chances of detecting a lexical boost. To account for this second possibility, we repeated the same analyses for Experiments 4 and 5, combining data from the two blocks in each experiment to maximize power. However, this analysis similarly revealed no evidence of a lexical boost in either experiment. Thus, overall our findings from the lexical and semantic boosts effects are not consistent with an Interactive Alignment framework, although it is possible that the lack of effects could in part be due to experimental design.

Conversely, our results may indicate that the alignment was the result of some mechanism other than Interactive Alignment. As noted earlier, studies have suggested that overspecification can be influenced by more strategic, partner-specific considerations (e.g., Rubio-Fernández, 2016). The overspecification we saw could be attributed to similar motivations. Speakers may, for instance, have adopted an overspecific partner's descriptive strategy on the assumption that this behavior would facilitate their partner's search task, perhaps via reasoning about the perceptual salience of color from their partner's perspective (cf. Davies & Katsos, 2013; Tarenskeen et al., 2015). This is supported by the fact that some participants mentioned during debrief a metalinguistic awareness that their partner was using color modifiers unnecessarily, suggesting that speakers may have overspecified despite recognizing that they were being redundant.[7]

It is worth pointing out that the various mechanisms are not necessarily mutually exclusive. Haywood, Pickering, and Branigan (2005) proposed, for instance, that syntactic priming effects may occur alongside other influences on production, such as a tendency to produce utterances that can be easily understood by a particular audience. In other words, priming within an Interactive Alignment framework may operate hand-in-hand with more strategic adaptive mechanisms, such as audience design (Clark, 1996). It is possible that the alignment we observed could be similarly attributed to multiple mechanisms. Overall, our finding that alignment was only weakly sensitive to lexical boost effects but responsive to partner effects provides more support for strategic mechanisms taking into account addressee-specific considerations. The interplay between how different mechanisms contribute to alignment in overspecification would be useful for future research to investigate.

7.3. Influence of partner over time

The robustness of partner effects on overspecification are also reflected in our time-course results. Across all five experiments we found that speakers' alignment with their partner's behavior increased as the interaction progressed. This is broadly in line with findings from cumulative priming that such mechanisms are sensitive to the accumulated frequency of experience in individuals. Importantly, Experiments 4 and 5 showed that speakers aligned even across a change in the partner's referential strategy. This implies that our time course effects were not simply due to a speaker-internal tendency toward consistency; rather, speakers continued to be influenced by their partner over time, and adjusted their own referential strategy in response to a change in their partner's. Moreover, speakers altered their production behavior to align with overspecification (Experiment 4) as well as minimal specification (Experiment 5). This indicates that they were not merely driven to produce expressions that were less effortful (minimal specification), as the partner influence extended to the production of utterances containing more information than is strictly necessary for object identification. This finding provides a point of comparison to studies investigating speakers' sensitivity to changes in the communicative context in referential description paradigms. Here, we observed that speakers were equally willing to switch to an overspecific or a minimally specific referential strategy in response to a change in their partner's strategy. This stands in contrast to Long et al. (2020), who observed an asymmetric tendency in speakers to switch to overspecification or minimal specification following a change in properties relating to the visual display. This inconsistency may be due to differences in the relevant contextual trigger (visual vs. social); however, another key difference between the two studies is the degree of interaction between the participant and the addressee. Long et al. (2020) relied on unidirectional communication in which participants were assigned a fixed role of speaker, whereas our task involved bidirectional interaction in which participants alternated between speaking and listening to their partner. This may point to the role of interaction in highlighting an interlocutor's informational needs and communicative preferences, motivating speakers to tailor their production behavior in response (cf. Grigoroglou & Papafragou, 2019).

Our time-course results support two preliminary conclusions. First, the fact that speakers continued to align across a change in partner strategy suggests that the tendency to align with a partner may outweigh speakers' internal tendency to maintain the same strategy, at least in the production of overspecific modified expressions. More broadly, speakers' willingness to align with an overspecific partner illustrates that overspecification may not be as dispreferred as traditional notions of rational communication imply it is—a view shared by studies analyzing the phenomenon from the addressee's perspective (Arts et al., 2011; Rubio-Fernández, 2016). Here, our results highlight a different aspect of the role of the addressee on overspecification—that of speakers' perception of their partner's referential preferences based on their partner's productions. Together, these findings point toward a general need to reevaluate the phenomenon of overspecification within the broader situational context (cf. Rubio-Fernández, 2019). However, it should be noted that in the context of our study, overspecification may have been more effortful from a production perspective, but was nevertheless still felicitous for communication. In other words, in aligning with their partner, whether overspecific or minimally specific, speakers were never jeopardizing successful communication. An interesting question to address is whether speakers would be similarly ready to align with a partner whose referential strategy is less felicitous, such as one who consistently underspecifies, or who switches from overspecification to underspecification. Research from referential production and comprehension shows that speakers routinely penalize underspecific utterances more than overspecific ones (Davies & Katsos, 2010; Engelhardt et al., 2006; Pogue, Kurumada, & Tanenhaus, 2016), reflecting an asymmetric tolerance for specificity violations. It stands to reason, therefore, that speakers may be less willing to align with an underspecific partner in interaction.

Finally, we note that the paradigm we employed is still somewhat limited in terms of ecological validity. Although we designed our recordings to simulate realistic partner behavior, and limited our analyses to participants who reported having believed they were engaged in live interaction, such methods still depart from authentic interaction between speakers. However, recent studies demonstrate successful implementation of similar methods for web-based interaction experiments, with results suggesting that a simulated partner was sufficiently convincing for participants to believe in the authenticity of the interaction (Buz, Jaeger, & Tanenhaus, 2014; Duran & Dale, 2011). Notably, Out, Goudbeek, and Krahmer (2020) also recently replicated (Goudbeek & Krahmer, 2012) Experiment 1, finding that speakers aligned with their hypothetical partner's choice of modifier (color vs. orientation) in more a naturalistic dailog setting with a real interlocutor. This demonstrates that alignment effects are comparable across real and simulated interaction, at least with a speaker's choice of modifier encoding.

Thus, we would likely expect to find similar effects in the alignment of modifier overspecification in a more authentic interactive context.

8. Conclusions

We investigated the influence of a partner's linguistic behavior on a speaker's tendency to overspecify. Our results show that speakers align with their partner to produce overspecific or minimally specific object descriptions, with the effect seen in both color and size modification. The time course of alignment revealed that partner effects increased as the interaction progressed, and alignment was sensitive to a change in the partner's linguistic behavior. This demonstrates the persistence of a partner's influence on a speaker's production tendencies across the duration of an interaction. However, we found little evidence that the alignment was affected by lexical or semantic boost effects, suggesting that these pressures may be less relevant in pragmatic alignment.Reip

Acknowledgments

This research received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement 681942), held by KS. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the article.

Funding Statement: Open Access funding enabled and organized by Projekt DEAL.

Notes

1 Of course not all languages are spoken, and we have no reason to expect that similar effects would not be seen in sign languages; however, in the absence of widely accepted modality-neutral terms we will use "speaker" and "listener" to denote the two roles in interaction.

2 "Minimal specification," also called "optimal specification" in some studies, is used in contrast to overspecification, and refers to the minimal description that would be sufficiently informative for the listener to identify the correct referent. We opt for the former, more neutral, term here based on recent research (in addition to the current results), which highlights that rather than being suboptimal, overspecification may in fact be purposeful in communication (Degen, Hawkins, Graf, Kreiss, & Goodman, 2020; Rubio-Fernández, 2019).

3 All experiments were preregistered individually on the Open Science Framework (links provided in each section). With the exception of the time-course analysis for Experiment 1, which was exploratory, all analyses were confirmatory and preregistered.

4 Note that the minimally specific partner did not produce color adjectives, hence the lexical boost effect is only relevant in the overspecific partner condition.

5 https://www.mturk.com/

6 As in the minimally specific partner condition in Experiments 1 and 2, the variable partner did not produce any color adjectives in this block, hence a lexical boost effect would only be relevant in the consistent (overspecific) partner condition, and indicated by a color by partner interaction.

7 Although we note that a number of other participants reported not having picked up on the manipulation at all after it was revealed during the debrief, an observation shared by priming experiments (e.g., Pickering & Branigan, 1998).

Open Research Badges

This article has earned Open Data and Open Materials badges. Data are available at https://doi.org/10.17605/OSF.IO/WGZ9T and materials are available at https://doi.org/10.17605/OSF.IO/HU8TA.

References

Arnold, J. E. (2008). Reference production: Production-internal and addressee-oriented processes. *Language and Cognitive Processes*, *23*(4), 495–527.

Arts, A., Maes, A., Noordman, L., & Jansen, C. (2011). Overspecification facilitates object identification. *Journal of Pragmatics*, *43*(1), 361–374.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). *lme4: Linear mixed-effects models using Eigen and S4*. R package version 1.1-6.

Belke, E., & Meyer, A. S. (2002). Tracking the time course of multidimensional stimulus discrimination: Analyses of viewing patterns and processing times during "same"-"different" decisions. *European Journal of Cognitive Psychology*, *14*(2), 237–266.

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

Bott, L., & Chemla, E. (2016). Shared and distinct mechanisms in deriving linguistic enrichment. *Journal of Memory and Language*, *91*, 117–140.

Branigan, H. P. (2007). Syntactic priming. *Language and Linguistics Compass*, *1*(1-2), 1–16.

Branigan, H. P., Pickering, M. J., & Cleland, A. A. (2000). Syntactic co-ordination in dialogue. *Cognition*, *75*(2), B13–B25.

Branigan, H. P., Pickering, M. J., Pearson, J., McLean, J. F., & Brown, A. (2011). The role of beliefs in lexical alignment: Evidence from dialogs with humans and computers. *Cognition*, *121*(1), 41–57.

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(6), 1482–1493.

Brodeur, M. B., Dionne-Dostie, E., Montreuil, T., & Lepage, M. (2010). The Bank of Standardized Stimuli (BOSS), a new set of 480 normative photos of objects to be used as visual stimuli in cognitive research. *PloS One*, *5*(5), e10773.

Buz, E., Jaeger, T. F., & Tanenhaus, M. K. (2014). Contextual confusability leads to targeted hyperarticulation. In Bello, P., Guarini, M., McShane, M., & Scassellati, B. (Eds.), *Proceedings of the 36th Annual Meeting of the Cognitive Science Society*, volume 36, (pp. 1970–1975). Quebec City, Canada.

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

Cleland, A. A., & Pickering, M. J. (2003). The use of lexical and syntactic information in language production: Evidence from the priming of noun-phrase structure. *Journal of Memory and Language*, *49*(2), 214–230.

Davies, C. & Katsos, N. (2009). Are interlocutors as sensitive to over-informativeness as they are to under-informativeness. In K. van Deemter, A. Gatt, R. van Gompel, and E. Krahmer (Eds). *Proceedings of the Workshop on Production of Referring Expressions: Bridging Computational and Psycholinguistic Approaches*, (pp. 282–287). PRE-CogSci-09, Amsterdam.

Davies, C., & Katsos, N. (2010). Over-informative children: Production/comprehension asymmetry or tolerance to pragmatic violations? *Lingua*, *120*(8), 1956–1972.

Davies, C., & Katsos, N. (2013). Are speakers and listeners "only moderately Gricean"? An empirical response to Engelhardt et al.(2006). *Journal of Pragmatics*, *49*(1), 78–106.

Degen, J., Hawkins, R. D., Graf, C., Kreiss, E., & Goodman, N. D. (2020). When redundancy is useful: A Bayesian approach to "overinformative" referring expressions. *Psychological Review*, *127*(4), 591–621.

Duran, N., & Dale, R. (2011). Spatial cognition adapts to social context. In Carson, L., Hölscher, C., & Shipley, T. (Eds.), *Proceedings of the 33rd Annual Meeting of the Cognitive Science Society*, volume 33, (pp. 1111–1116). Boston, USA.

Engelhardt, P. E., Bailey, K. G., & Ferreira, F. (2006). Do speakers and listeners observe the Gricean Maxim of Quantity? *Journal of Memory and Language*, *54*(4), 554–573.

Fox Tree, J. E. (1999). Listening in on monologues and dialogues. *Discourse Processes*, *27*(1), 35–53.

Garrod, S., & Anderson, A. (1987). Saying what you mean in dialogue: A study in conceptual and semantic co-ordination. *Cognition*, *27*(2), 181–218.

Goudbeek, M., & Krahmer, E. (2012). Alignment in interactive reference production: Content planning, modifier ordering, and referential overspecification. *Topics in Cognitive Science*, *4*(2), 269–289.

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

Grigoroglou, M., & Papafragou, A. (2019). Children's (and adults') production adjustments to generic and particular listener needs. *Cognitive Science*, *43*(10), e12790.

Grodner, D. & Sedivy, J. C. (2011). The Effect of Speaker-Specific Information on Pragmatic Inferences. In E. Gibson & N. Pearlmutter (Eds.). *The processing and acquisition of reference* (Vol. 2327, pp. 239-272). Cambridge, MA: MIT Press.

Hartsuiker, R. J., Pickering, M. J., & Veltkamp, E. (2004). Is syntax separate or shared between languages? Cross-linguistic syntactic priming in Spanish-English bilinguals. *Psychological Science*, *15*(6), 409–414.

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

Heller, D., & Chambers, C. G. (2004). Would a blue kite by any other name be just as blue? Effects of descriptive choices on subsequent referential behavior. *Journal of Memory and Language*, *70*, 53–67.

Jaeger, T. F., & Snider, N. (2008). Implicit learning and syntactic persistence: Surprisal and cumulativity. In Love, B. C., McRae, K., & Sloutsky, V. M. (Eds.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society*, (pp. 1061–1066). Cognitive Science Society Washington, DC.

Kaschak, M. P. (2007). Long-term structural priming affects subsequent patterns of language production. *Memory & Cognition*, *35*(5), 925–937.

Kaschak, M. P., Kutta, T. J., & Schatschneider, C. (2011). Long-term cumulative structural priming persists for (at least) one week. *Memory & Cognition*, *39*(3), 381–388.

Kim, M., Horton, W. S., & Bradlow, A. R. (2011). Phonetic convergence in spontaneous conversations as a function of interlocutor language distance. *Laboratory Phonology*, *2*(1), 125–156.

Koolen, R., Gatt, A., Goudbeek, M., & Krahmer, E. (2011). Factors causing overspecification in definite descriptions. *Journal of Pragmatics*, *43*(13), 3231–3250.

Koolen, R., Gatt, A., van Gompel, R. P., Krahmer, E. & Van Deemter, K. (2016). Viewing time affects overspecification: Evidence for two strategies of attribute selection during reference production. In A. Papafragou, D. Grodner, Daniel Mirman, J.C. Trueswell (Eds). *Proceedings of the 38th Annual Conference of the Cognitive Science Society*, (pp. 159–164). Philadelphia, USA.

Koolen, R., Goudbeek, M., & Krahmer, E. (2013). The effect of scene variation on the redundant use of color in definite reference. *Cognitive Science*, *37*(2), 395–411.

Koolen, R., Krahmer, E., & Swerts, M. (2016). How distractor objects trigger referential overspecification: Testing the effects of visual clutter and distractor distance. *Cognitive Science*, *40*(7), 1617–1647.

Long, M., Rohde, H., & Rubio-Fernandez, P. (2020). The pressure to communicate efficiently continues to shape language use later in life. *Scientific Reports*, *10*(1), 1–13.

Lundqvist, E., Flykt, A., & Ohman, A. (1998). The Karolinska Directed Emotional Faces - KDEF. CD ROM from Department of Clinical Neuroscience, Psychology section, Karolinska Institutet, ISBN 91-630-7164-9.

Out, C., Goudbeek, M., & Krahmer, E. (2020). Do speaker's emotions influence their language production? Studying the influence of disgust and amusement on alignment in interactive reference. *Language Sciences*, *78*, 101255.

Pechmann, T. (1989). Incremental speech production and referential overspecification. *Linguistics*, *27*(1), 89–110.

Peloquin, B. N., Goodman, N. D., & Frank, M. C. (2020). The interactions of rational, pragmatic agents lead to efficient language structure and use. *Topics in Cognitive Science*, *12*(1), 433–445.

Pickering, M., & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language*, *39*(4), 633–651.

Pickering, M., & Garrod, S. (2004). Towards a mechanistic psychology of dialogue. *Behavioural and Brain Sciences*, *27*, 169–226.

Pivneva, I., Palmer, C., & Titone, D. (2012). Inhibitory control and L2 proficiency modulate bilingual language production: Evidence from spontaneous monologue and dialogue speech. *Frontiers in Psychology*, *3*, 57.

Pogue, A., Kurumada, C., & Tanenhaus, M. K. (2016). Talker-specific generalization of pragmatic inferences based on under-and over-informative prenominal adjective use. *Frontiers in Psychology*, *6*, 2035.

R Core Team. (2020). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Reips, U. D. (2000). The Web experiment method: Advantages, disadvantages, and solutions. In M. H. Birnbaum (Eds). *Psychological experiments on the Internet* (pp. 89-117). New York: Academic Press. New York: Academic Press.

Reitter, D., Keller, F., & Moore, J. D. (2011). A computational cognitive model of syntactic priming. *Cognitive Science*, *35*(4), 587–637.

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

Rubio-Fernández, P. (2019). Overinformative speakers are cooperative: Revisiting the Gricean Maxim of Quantity. *Cognitive Science*, *43*(11), e12797.

Schober, M. F., & Clark, H. H. (1989). Understanding by addressees and overhearers. *Cognitive Psychology*, *21*(2), 211–232.

Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G., & Carlson, G. N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition*, *71*(2), 109–147.

Tarenskeen, S., Broersma, M., & Geurts, B. (2015). Overspecification of color, pattern, and size: Salience, absoluteness, and consistency. *Frontiers in Psychology*, *6*, 1703.

Van Gompel, R., Gatt, A., Krahmer, E. & Van Deemter, K. (2014). Overspecification in reference: Modelling size contrast effects. In Poster presented at the Architectures and Mechanisms for Language Processing conferenece (AMLaP). Edinburgh, UK.

Weatherholtz, K., Campbell-Kibler, K., & Jaeger, T. F. (2014). Socially-mediated syntactic alignment. *Language Variation and Change*, *26*(3), 387.

Westerbeek, H., Koolen, R., & Maes, A. (2015). Stored object knowledge and the production of referring expressions: The case of color typicality. *Frontiers in Psychology*, *6*, 935.