Speakers and listeners exploit word order for communicative efficiency: A cross-linguistic investigation

P. Rubio-Fernandez[1, 2*], F. Mollica[3], J. Jara-Ettinger[4, 5]

[1]Department of Philosophy, University of Oslo;

[2]Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology;

[3]School of Psychological Sciences, University of Melbourne;

[4]Department of Psychology, Yale University;

[5]Department of Computer Science, Yale University.

*Correspondence to: paula.rubio-fernandez@ifikk.uio.no

Pragmatic theories and computational models of reference must account for people's frequent use of redundant color adjectives (e.g., referring to a single triangle as 'the blue triangle'). We hypothesize that redundant color adjectives are normally intended to facilitate the listeners' visual search for the referent, which makes these language uses efficient. The efficiency view rests on incrementality and is the only account to predict cross-linguistic differences on the production of redundant color adjectives depending on word order. Experiment 1 confirmed that English speakers produced more redundant color adjectives (e.g., 'the blue triangle') than Spanish speakers (e.g., 'el triángulo azul'), but both language groups used more redundant color adjectives in complex displays where the listener's visual search was harder. In Experiments 2a and 2b, we used eye tracking to show that listeners' visual search maximizes communicative efficiency: English speakers established color contrast across categories (BLUE SHAPES > TRIANGULAR ONE), whereas Spanish speakers established color contrast within a category (TRIANGLES > BLUE ONE). Spanish speakers, however, reversed their visual search strategy when tested in English immediately after. Our results show that speakers and listeners of different languages exploit word order to increase communicative efficiency.

Keywords: communicative efficiency, reference, redundant color adjectives, adjective position.

In the last two decades, research in cognitive science has revealed that language is shaped by a pressure to communicate efficiently (Kemp et al., 2018; Gibson et al., 2019; see also Zipf, 1949; Givon, 1979). The idea that language is optimized for fast, easy and reliable information transmission has been highly influential in linguistics, explaining language structure at a phonetic (Aylett & Turk, 2004), morphological (Piantadosi et al., 2011), semantic (Kemp & Regier, 2012) and syntactic level (Gibson et al., 2013). Beyond structural properties, this approach also explains language use, providing normative and descriptive accounts for why speakers ought to convey meaning efficiently (Jaeger, 2010; Mahowald et al., 2013), and how listeners, by holding speakers accountable to this expectation, can go beyond literal meanings and infer things that were left unsaid (Grice, 1975). It follows from this view that if someone asked you 'Could you pass me the blue bowl?', you would assume that there is more than one bowl to choose from, and that not all bowls are blue. Yet if the adjective 'blue' referred to a single bowl (rather than preempting an ambiguity between several bowls), the speaker would have failed to communicate efficiently.

This account of communicative efficiency has great explanatory power, yet it does not always map on to the experimental record. A large number of studies challenge in fact its most basic predictions: speakers are often redundant or overspecific, using descriptive color adjectives that do not add necessary information (Sedivy, 2003, 2004; Maes et al., 2004; Engelhardt et al., 2006; Van der Sluis & Krahmer, 2007; Arts et al., 2011a, 2011b; Engelhardt & Ferreira, 2014; Rubio-Fernandez, 2016, 2019). In response to this puzzle, some have argued that speakers use color adjectives without checking if they are necessary in the context (Pechamann, 1989; Belke & Meyer, 2002; Belke, 2006; Koolen et al., 2013), or alternatively, that they do so strategically to pre-empt a possible ambiguity (Degen et al., under review; Hawkins et al., under review). Here we adopt the view that redundant color adjectives can facilitate the listener's visual search for a referent, being in fact efficient for speaker-listener coordination (Rubio-Fernandez, 2016, 2019).

Overspecification is a pragmatic phenomenon that is normally taken to challenge the Gricean Maxim of Quantity: speakers should make their contribution as informative as is required for the current purpose of the exchange, but not more (Grice, 1975). The purpose of reference is to ensure that the listener uniquely identifies the intended referent, and referential overspecification is therefore understood to run counter to the Gricean maxim (e.g., Engelhardt et al., 2006; Koolen et al., 2013; Engelhardt & Ferreira, 2014; cf. Geurts & Rubio-Fernandez, 2015). However, Rubio-Fernandez (2019) has recently proposed that in face-to-face interaction, the goal of a cooperative speaker should be to provide enough information to ensure the listener identifies the referent rapidly and easily in the visual context. In this view, redundant color adjectives can be more efficient than shorter descriptions if they facilitate the listener's visual search for the referent.

Unlike the traditional pragmatic view, the efficiency account explains overspecification on the basis of informativity and discriminability (Rubio-Fernandez, 2016, 2019). Thus, efficiency is measured not only in relation to the information that is necessary for successful communication, but also on the basis of the time and effort it takes interlocutors to produce and assign reference. This efficiency account is therefore in line with the view that reference is a collaborative process (Clark & Marshall, 1981), where both interlocutors try to spend the least collaborative effort (Clark & Schaefer, 1989). Given our focus on speaker-listener coordination, the present study investigated not only how speakers use color adjectives, but also how color adjectives can guide the listener's visual search for a referent.

If a referential expression's efficiency depends not only on the likelihood of communicative success, but also on the processing effort it requires, then a cooperative speaker should be sensitive to the incremental nature of language processing as it determines the order in which information becomes available to the listener. This is particularly important when looking at language processing cross-linguistically: in a language like English, for example, adjectives are encoded before the noun (e.g., 'blue triangle'), whereas in languages like Spanish, they are encoded after the noun (e.g., 'triángulo azul'). That means that, in processing a color description, an English speaker would search for the referent guided by the adjective (e.g., by color), whereas a Spanish listener would do so guided by the noun (e.g., by shape), with these two search procedures being more or less efficient depending on the visual context.

Here we aimed to investigate precisely how efficiency pressures in reference production and comprehension are shaped by incrementality. We view efficiency and incrementality as interacting forces in communication, with speakers and listeners aiming to communicate efficiently, while being constrained by the incremental nature of language production and comprehension (Rubio-Fernandez & Jara-Ettinger, under review). We predict that speakers and listeners will exploit word order to increase the communicative efficiency of color adjectives, and provide evidence for this hypothesis in three experiments: a language-production experiment investigating how English and Spanish speakers use color adjectives depending on their word order and on the efficiency pressures of the task at hand, and two eye-tracking experiments investigating how adjective position affects the way listeners establish color contrast during language processing.

Contrast and incrementality

According to the standard pragmatic view, the function of color adjectives is to contrast different members of the same category (e.g., in Fig. 1A, ‘blue’ would distinguish the blue triangle from the red one). When this is not the case, the use of a color adjective is considered redundant or non-contrastive (as in Figs. 1B-1C; Sedivy, 2003, 2004). However, from an incremental point of view, color contrast may be established in two ways: within members of a category or across members of different categories (Rubio-Fernandez, 2016). When English speakers process a description such as ‘the blue triangle’, their visual search for the referent is guided by color and refined by shape (i.e. they should look for a blue shape that is triangular), whereas when Spanish speakers process the mirror phrase ‘el triángulo azul’, their visual search is guided by shape and refined by color (i.e. they should look for a triangular shape that is blue). It follows from this basic difference that adjective position affects how listeners establish color contrast during processing: in hearing ‘the blue triangle’ in Figure 1A, English speakers would first contrast blue vs. non-blue items, regardless of their shape[1], while in hearing ‘el triángulo azul’, Spanish speakers would identify the triangles and discriminate them by color. Thus, given their word order, Spanish

[1] Sedivy (2003, 2004) used the visual world paradigm to present listeners with two objects of the same color (e.g., a blue triangle and a blue square), only one of which had a contrast object in the display (e.g., a red triangle, but no other square). In this visual context, listeners are able to anticipate the noun by deriving a contrastive inference (i.e. they anticipate that ‘blue’ will refer to the triangle, not the square). However, their initial search is nonetheless guided by the adjective (i.e. they identify the two blue shapes when hearing ‘blue’; Rubio-Fernandez et al., under review).

speakers are more likely to establish color contrast within a category (TRIANGLES > BLUE ONE), while English speakers would do so across categories (BLUE SHAPES > TRIANGULAR ONE)[2].

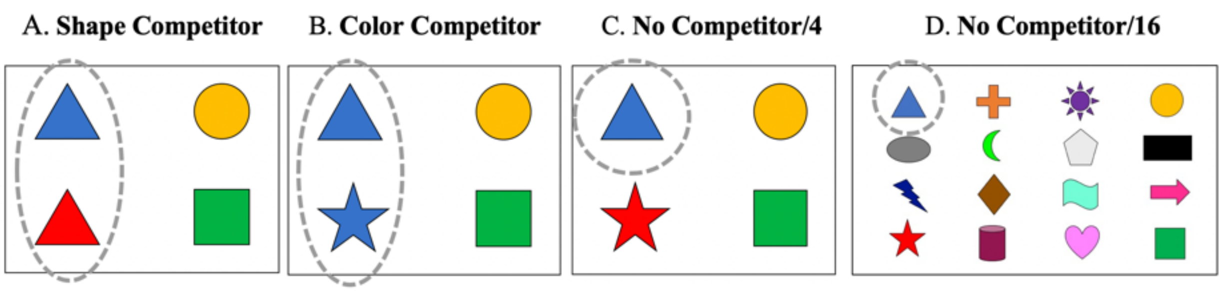

Figure 1: Sample displays from the critical conditions in the study. The target was the blue triangle in all displays. Experiment 1 used the No Competitor/4 and No Competitor/16 conditions in a language-production task and Experiments 2a and 2b used the Shape Competitor and Color Competitor conditions in two eye-tracking tasks.

While perhaps obvious on reflection, the incremental view of color contrast has been overlooked in both pragmatics and psycholinguistics, despite having important theoretical implications for referential communication (for a study on size contrast, see Brown-Schmidt & Konopka, 2008). Contrary to the standard pragmatic view, redundant color adjectives are often contrastive, only that color contrast is established across categories during processing (e.g., blue vs. non-blue shapes in Figs. 1B-1C). The incremental view therefore challenges what counts as a redundant referential expression, since prenominal color adjectives would only be entirely

[2] The incremental view predicts that when English speakers process the color adjective in a postnominal relative clause (e.g., 'the triangle that is blue'), they would establish color contrast within a category.

redundant during processing if they did not have any discriminatory power in the visual display (i.e. if they did not distinguish the target referent from any of its competitors).

Adopting an incremental perspective on color contrast therefore results in a graded notion of redundancy that is sensitive to the discriminatory power of a color word in the entire visual space, as opposed to the theoretical notion of pragmatic contrast, which applies within a category and penalizes all redundant uses as non-contrastive (Rubio-Fernandez, 2016).

The incremental efficiency hypothesis

Since speech unfolds linearly, listeners interpret language incrementally (Eberhard et al., 1995; Spivey et al., 2001). We propose that, in order to explain redundant uses, a color adjective's efficiency ought to be evaluated incrementally, rather than globally. This is what we call the incremental efficiency hypothesis. This hypothesis has implications for both speakers and listeners, if one assumes that they aim to coordinate efficiently. On the production side, the incremental efficiency hypothesis predicts that speakers will be sensitive to adjective position in so far as it affects the efficiency of the color cue for the listener's visual search. On the comprehension side, the same hypothesis predicts that listeners will interpret language incrementally and efficiently, using color information as it becomes available during processing.

This incremental analysis makes different predictions from the standard global analysis (Sedivy, 2003, 2004; Engelhardt et al., 2006; Koolen et al., 2013). From a global standpoint, the word 'blue' is optimally informative in Figure 1A because it is necessary to distinguish the two triangles, but it is overinformative in Figures 1B-1D because there is only one triangle in each display (cf. Degen et al., under review). From an incremental standpoint, however, the relative efficiency of ‘blue’ should be evaluated based on the incomplete phrase ‘the blue …’ rather than on the full phrase ‘the blue triangle’. Under this analysis, ‘blue’ is now efficient in Figures 1A, 1C and 1D because it secures unique reference for the listener, whereas it would be inefficient in 1B because it would create a temporary ambiguity between the blue triangle and the blue star.

From an incremental point of view, mentioning the color of the target triangle in displays 1A, 1C and 1D would be more efficient than in 1B, yet this does not explain why referring to ‘the blue triangle’ may be more efficient than referring to ‘the triangle’ in displays 1C and 1D, given that both expressions secure unique reference incrementally, and the unmodified expression is shorter. Here we appeal to the relative efficiency of different visual search procedures. An extensive literature on visual cognition has shown that color is a highly salient cue for visual search, which plays a key role in object recognition (for a review and meta-analysis, see Bramão et al., 2011). Off-line psycholinguistic studies have revealed that redundant properties such as color, size and position can reduce target identification times (Sonnenschein & Whitehurst, 1982; Mangold & Pobel, 1988; Paraboni et al., 2007; Arts et al., 2011b; Paraboni & Van Deemter, 2014). Recently, Rubio-Fernandez (under review) further showed that redundant color adjectives also reduce target identification times in an eye-tracking task tapping real-time language processing. Thus, searching by color in displays like 1A, 1C and 1D was faster than searching by shape; in other words, participants were faster to identify ‘the blue star’ than ‘the star’. However, as predicted by the incremental efficiency hypothesis, color did not always lead to more efficient visual search: in visual displays where color was not distinctive of the target (as in display 1B), participants were faster to process the shorter description because color created a temporary ambiguity that delayed target identification.

Language production studies also offer support to the incremental efficiency hypothesis. In polychrome displays where the color of the target is distinctive, participants tend to use redundant color adjectives; whereas in monochrome displays where color is a highly inefficient visual cue, they prefer minimal descriptions (Belke, 2006; Koolen et al., 2013; Rubio-Fernandez, 2016, 2019). We take the parallel results observed in reference production and comprehension studies as evidence that reference is a collaborative process (Clark & Marshall, 1981; Clark & Schaefer, 1989).

In summary, psycholinguistic studies have previously suggested that redundant adjectives can facilitate the listener's visual search for a referent. Likewise, eye-tracking studies have also shown that language is interpreted incrementally (Eberhard et al., 1995; Spivey et al., 2001). However, ours is the first pragmatic account to explain overspecification as resulting from efficiency pressures that are shaped by incrementality. Another innovation with our account is that it predicts cross-linguistic differences on overspecification depending on adjective position (Rubio-Fernandez, 2016, 2019). Current pragmatic theories and computational models of reference production have only been tested in languages like English or Dutch, which have prenominal adjectives (e.g., Sedivy, 2003, 2004; Van Gompel et al., 2019; Degen et al., under review) and do not make different predictions for other languages (see Rubio-Fernandez & Jara-Ettinger, under review). However, as we aim to show in this study, adjective position can affect the relative efficiency of a color modifier for visual search, affecting their production as a result.

Cross-linguistic implications

Since language is processed incrementally, the relative efficiency of color for visual search depends on the position of the color adjective relative to the noun it accompanies. Here we aimed to test the incremental efficiency hypothesis by comparing English and Spanish, which have reverse adjective-noun orders and should have reverse strategies to increase the efficiency of color for visual search.

We started by investigating the effect of adjective position on the use of redundant color adjectives. According to the incremental efficiency hypothesis, in languages like English, speakers often use redundant color adjectives to facilitate the listener's visual search, allowing them to quickly identify the color-matching referent, even before they hear the noun (Rubio-Fernandez, under review). However, in languages like Spanish, redundant color adjectives are less efficient because they are processed after the listener has begun searching for the noun. This account predicts that Spanish speakers should produce fewer redundant color adjectives than English speakers, in line with the results of Rubio-Fernandez (2016, 2019).

However, if this cross-linguistic difference is driven by efficiency pressures, the effect of adjective position should be modulated by the density of the display, with redundant color adjectives being generally more efficient in denser displays (e.g., Paraboni et al., 2007; Clarke et al., 2013; Koolen et al., 2015; Gatt et al., 2017; Rubio-Fernandez, 2019). Thus, in Figure 1C, a Spanish listener should be able to identify the target in hearing 'triángulo' (making the ensuing adjective 'azul' inefficient), whereas in Figure 1D, the same listener would probably benefit from learning the color of the target given their harder visual search. It follows from the incremental efficiency hypothesis that the difference in the production of redundant color adjectives between English and Spanish speakers should be reduced in denser displays. This twofold hypothesis was tested with native speakers of English and Spanish in Experiment 1.

We also investigated the effect of adjective position on visual search in order to establish the differential efficiency of prenominal and postnominal color adjectives. The incremental efficiency hypothesis predicts that both English and Spanish speakers should identify the target as soon as they have sufficient information to do so. That means that in Figure 1A, English speakers should identify the target in hearing 'blue', not considering the second triangle when processing the noun. Reversely, in Figure 1B, Spanish speakers should identify the target in hearing 'triángulo', disregarding the other blue shape when processing the adjective. The incremental efficiency hypothesis also predicts which competitors will create interference in processing each language: when English speakers search for 'the blue triangle' in Figure 1B, the blue star (or color competitor) should interfere with their visual search, unlike in Spanish. However, when Spanish speakers search for 'el triángulo azul' in Figure 1A, the other triangle (or shape competitor) should interfere with their visual search, unlike in English. We tested these predictions using the visual world paradigm in Experiment 2a.

To further investigate whether listeners optimize visual search by exploiting word order, the Spanish speakers participated a second time in Experiment 2b, now completing the English version of the task. If the predicted cross-linguistic differences result from a pressure to communicate efficiently, they should be flexible, with Spanish speakers reversing their visual search strategy in English. However, if one's native language determines how color contrast is established during processing (a hypothesis supporting linguistic determinism; Carroll, 1956), Spanish speakers should continue to suffer interference from the shape competitor when tested in English. That is, in processing the shape noun in ‘The blue triangle’, they should consider both triangles in Figure 1A. Given that the Spanish word order supports the canonical color contrast established within members of the same category, it is possible that the frequency of this interpretation in Spanish introduces a bias when processing English as a second language. Such a bias, however, would not be efficient for visual search.

Overall, empirical support for the above hypotheses would confirm that the redundant use of color adjectives is modulated by efficiency pressures on both speakers and listeners (Rubio-Fernandez, 2016, 2019, under review), rather than being pragmatically infelicitous (Engelhardt et al., 2006, 2011). Moreover, the predicted results would support the incremental efficiency hypothesis, according to which a referential expression’s efficiency should be calculated incrementally in relation to the entire visual context, rather than on the informativity of the full message.

Experiment 1

Methods

Participants

A group of 25 undergraduates from University College London and 25 undergraduates from the Universidad de las Islas Baleares (Spain) took part in Experiment 1. The UCL undergraduates were native speakers of English and the UIB undergraduates were native speakers of Spanish. Both groups participated for monetary compensation. All participants reported having normal color vision.

Materials and procedure

Two types of displays were created, one for the Experimenter (consisting of 20 displays of shapes) and another one for the participant (consisting of empty grids with a cross marking the position of the target in the Experimenter's display). The target shapes were the following: circle, cross, diamond, heart, oval, rectangle, square, triangle, star and sun; and came in the following colors: black, blue, brown, green, grey, orange, pink, purple, red and yellow. Target position was counterbalanced across trials. The first block of trials consisted of ten displays from the No Competitor/4 condition (NC/4; Fig 1C) and the second block consisted of ten displays from the No Competitor/16 condition (NC/16; Fig 1D), presented in the same random order.

The displays were shown on a computer monitor placed in front of the Experimenter. The participant sat beside and behind the Experimenter and their task was to ask the Experimenter to click on the target shape in each trial. In order to determine which shape was the target, participants were given printouts of 20 empty grids with a cross indicating the position of the target in the Experimenter's display. The instructions stressed that the Experimenter did not know which shape was the target in each trial, and that all shapes in the displays were different. It followed from this description that color adjectives would be redundant in all trials.

Participants were told that their responses would serve as control data in a study originally designed for children. This was done in order to avoid that participants may become self-conscious and start producing unnatural responses because of the simplicity of the task (this was observed in a pilot study where participants described the shapes in great detail). Participants' requests were recorded and later coded as redundant or not redundant by two blind coders. Only referential expressions including both an adjective and a noun (e.g., 'The blue triangle') were coded as redundant. The task lasted less than 10 minutes.

Results

Figure 2 shows the percentage of times participants used color adjectives as a function of the number of shapes in the display. When there were only four shapes, English speakers used color adjectives 37.3% of time (95% CI:18.2-55.9) whereas Spanish speakers used color adjectives only 2.73% of the time (95% CI: 0-5.0), with these rates being reliably different (difference = 34.57%; 95% CI: 14.09-53.64). In the display with 16 shapes, English speakers now produced color adjectives 80.5% of time (95% CI:67.7-95.9), a reliably higher rate than their production in the four-shape condition (difference = 43.18%; 95% CI:19.55-68.18). Spanish speakers also produced more color adjectives in the 16-shape condition at 61.4% (95% CI: 45.9-77.7), which was reliably higher than their production in the four-shape condition (difference = 58.64%; 95% CI:42.72-75.45) but was no longer lower than the English speakers' rate (difference = 19.09; 95% CI: -2.27 – 40.9).

The overall pattern of results from Experiment 1 was also visible at the subject-level. 50% (n=11) of English speakers used more color adjectives in the 16-shape condition than in the four-shape condition, 50% (n=11) used the same amount (8 of these participants used color adjectives in every trial of the four-shape block, making it impossible for them to use more color adjectives in the 16-shape block), and 0% used fewer color adjectives. Among Spanish speakers, 82% (n=18) used more color adjectives in the 16-shape condition than in the four-shape condition, 18% (n=4) used an equal amount, and 0% used fewer color adjectives (for data visualizations, see Supplemental Materials).

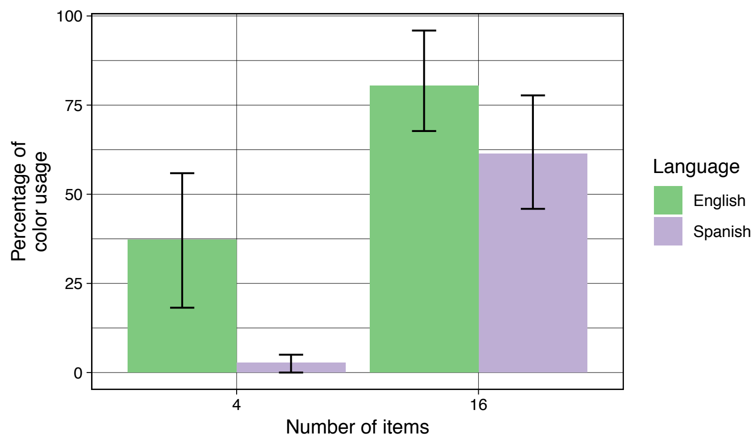

Figure 2. Average percentage of times speakers used redundant color adjectives in each language and condition. Vertical bars show 95% bootstrapped confidence intervals.

The results of Experiment 1 offered support to the incremental efficiency hypothesis, with both English and Spanish speakers producing redundant color adjectives when it would be more efficient for the listener's visual search. However, the question remains as to whether the cross-linguistic differences observed in language production are related to differences in language processing. Such a relation would support the view that reference is a collaborative process between speakers and listeners (Clark & Marshall, 1981; Rubio-Fernandez, 2016, 2019). The aim of Experiments 2a and 2b was to test the incremental efficiency hypothesis using eye tracking during language processing.

Experiments 2a and 2b

Methods

Participants

A new group of 25 undergraduates from each university took part in Experiment 2a. The UCL undergraduates were native speakers of English and the UIB undergraduates were native speakers of Spanish. Both groups participated for monetary compensation. All participants reported having normal color vision.

The same group of 25 undergraduates from UIB who took part in the Spanish version of Experiment 2a were tested again in the English version immediately after. All participants in Experiment 2b reported having an intermediate level of English and none reported speaking English natively. One of the participants did not perform the second task because they did not feel comfortable being tested in English.

Sample size was determined by the time available to collect data at UCL and UIB. A post-hoc power analysis revealed power > 0.9 for our key predictions (see Supplemental Materials).

Materials and procedure

Materials for all three experiments are available at OSF (https://osf.io/9hw68/?view_only=04a2e84459ac46dab007ce2d12c36d6f). In Experiments 2a and 2b, the visual materials consisted of 72 displays of four geometrical shapes, including 12 critical items from the Shape Competitor condition (SC; e.g., Fig 1A), 12 from the Color Competitor condition (CC; e.g., Fig 1B) and 48 fillers. Filler trials included both shape and color competitors and were not analyzed for this study because they speak to a different question (i.e. the derivation of pragmatic inferences; see Sedivy, 2003, 2004). The target shapes (2 per critical condition) were: circle, diamond, rectangle, square, star and triangle; and the colors of the shapes were: blue (x2), brown, green (x2), orange (x2), pink (x2), purple, red and yellow. The position of the target and competitor shapes were counterbalanced across trials, and trials were randomized individually for each participant.

The displays were on the screen for 400ms before the instructions started. This is a relatively short preview window, especially considering that launching a saccade takes 200ms. This short preview window was intended to prevent participants from conceptualizing all the color shapes in the display prior to the start of the instruction, which would have allowed them to identify the target artificially fast.

All instructions were of the form 'Click on the [COLOR ADJECTIVE + SHAPE NOUN]' or 'Haz click en el [SHAPE NOUN + COLOR ADJECTIVE]'. The instructions were recorded by male native speakers of British English and Castilian Spanish, respectively, who did not stress the adjectives contrastively. For data analysis, a critical time window was calculated for each instruction from the onset of the adjective until the offset of the noun in English, and from the onset of the noun until the offset of the adjective in Spanish. The mean duration of the critical time window was 743ms in English and 996ms in Spanish. We decided on this time window because its structure is parallel in the two languages and would allow observing an early effect of adjective position. We did not use RTs to calculate individual time windows because it introduced the issue of having longer time-windows in conditions with an interfering competitor.

Participants were told that they were going to listen to a series of instructions to click on different geometrical shapes and their task was to click on the target as fast and accurately as possible. All participants were first presented with four displays that served as warm-up trials, followed by the remaining 68 trials in a random order. In between trials, participants had to click on a cross in the center of the screen to move on to the next trial. This ensured that participants' gaze and the mouse cursor were always in the center of the screen at the start of each trial.

Results

Data and analysis code for all three experiments are available at OSF (https://osf.io/9hw68/?view_only=04a2e84459ac46dab007ce2d12c36d6f). Eye-tracking data were standardly corrected by +200ms in both experiments in order to account for the time it takes to launch a saccade. Figure 2 shows the percentage of fixations over time. As predicted, English speakers fixated more on the target when there was a shape competitor (53.9%; 95% CI: 49.2-58.6) than when there was a color competitor (42.1%; 95% CI:39.2-45.1), with a reliable difference between the two conditions (Target advantage in CC = 11.8%; 95% CI: 6.05-17.43). By contrast, Spanish speakers fixated less on the target when there was a shape competitor (40.5%; 95% CI: 37.4-43.6) than when there was a color competitor (52.5%; 95% CI: 48.8-56.0), also revealing a significant difference (Target advantage in SC = 12%; 95% CI: 7.07-16.63). Consistent with this, English speakers fixated more on the target than Spanish speakers did when there was a shape competitor (Difference = 13.4%; 95% CI: 7.58-19.07), whereas Spanish speakers fixated more on the target than English speakers did when there was a color competitor (Difference = 10.4; 95% CI: 5.59-15.05).

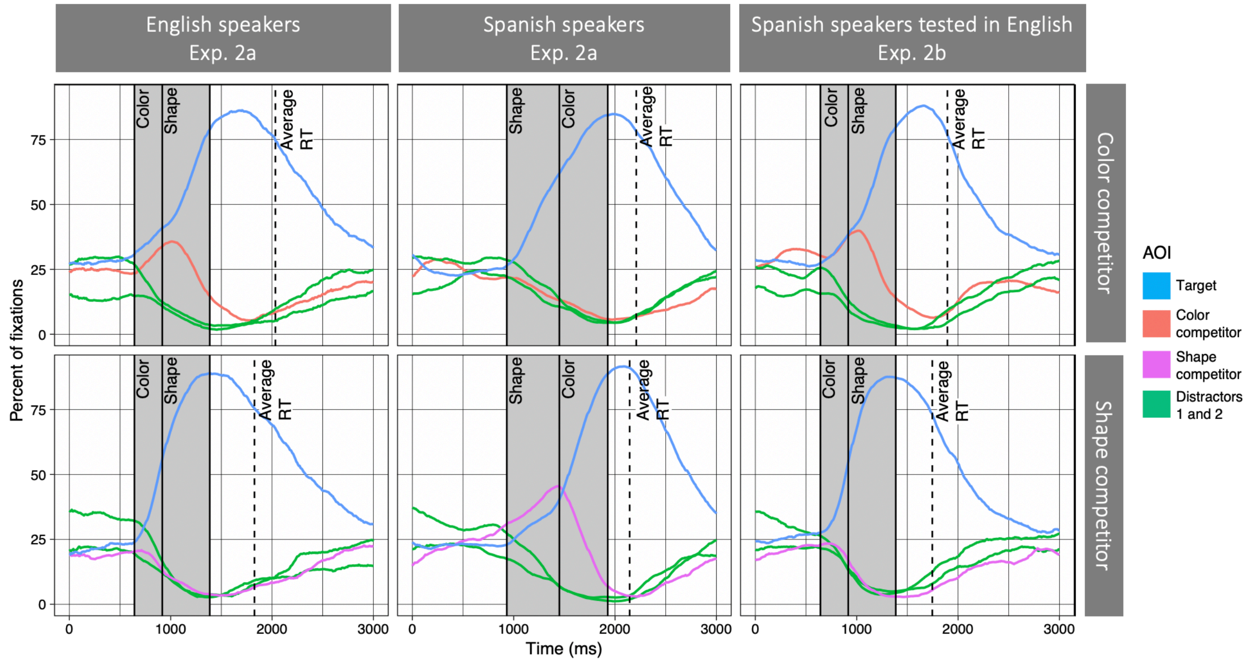

Figure 3. Percentage of eye fixations on the four shapes during the processing of instructions including color and shape words (e.g., 'Click on the blue triangle'). At each time point, the curve represents the percentage of fixations in the previous 200ms. The grey region represents the duration of the critical phrase and the dashed lines represent participants' average reaction times. Experiment 2a tested English and Spanish speakers, and Experiment 2b retested the Spanish speakers in English.

Fixations on the competitor also supported the incremental efficiency hypothesis (for data visualizations, see Supplemental Materials), with Spanish speakers suffering more interference from the shape competitor relative to English speakers (Interference difference = 21.7%; 95% CI: 18.13-25.23). By contrast, English speakers suffered more interference from the color competitor relative to Spanish speakers (Interference difference = 15.4%; 95% CI: 12.14-18.86).

In contrast to the results of Experiment 2a, Spanish speakers' looking pattern was reversed when tested in English (see Fig. 2). Participants in Experiment 2b fixated more on the target when there was a shape competitor (55.6%; 95% CI: 52.4 - 59.1) relative to when there was a color competitor (40.3%; 95% CI:37.2 - 43.6), with a reliable difference between conditions (Target advantage in SC = 15.3%; 95% CI: 10.75 - 19.98). This pattern was reliably different from their looking pattern when tested in Spanish, both when there was a shape competitor (Spanish tested in Spanish vs in English: Difference in SC = 15.1%; 95% CI: 7.58 - 19.07) and a color competitor (Spanish tested in Spanish vs in English: Difference in CC = 12.2%; 95% CI: 7.29-16.94). Consistent with this, participants' looking pattern was no longer different from that of English speakers when there was a shape competitor (Spanish tested in English vs English speakers: Difference in SC = 1.7%; 95% CI: -3.97 - 7.78) or when there was a color competitor (Spanish tested in English vs English speakers: Difference in CC= -1.8%; 95% CI: -6.14 – 2.74). Similarly, Spanish speakers' interference was reversed when tested in English (for data visualizations, see Supplemental Materials). Participants now showed increased interference from the color competitor (Interference increase in CC = 18.5%; 95% CI: 15.51 – 21.81) and reduced interference from the color competitor (Interference decrease in SC = 20.3%; 95% CI: 17.08 - 23.57).

The interaction between language and competitor type in the looking patterns also appeared at the subject-level (see Fig. 3). 88% (n=22) of English speakers and 96% (n=24) of Spanish speakers tested in English fixated more on the target when there was a shape competitor than when there was a color competitor. By contrast, 88.5% (n=23) of Spanish speakers tested in Spanish fixated more on the target when there was a color competitor than when there was a shape competitor. The results of competitor interference were consistent with this. 92% (n=23) of English speakers and 100% (n=25) of Spanish speakers tested in English fixated more on the competitor when it was a color match than when it was a shape match. By contrast, 100% (n=26) of Spanish speakers tested in Spanish fixated more on the competitor when it was a shape match than when it was a color match.

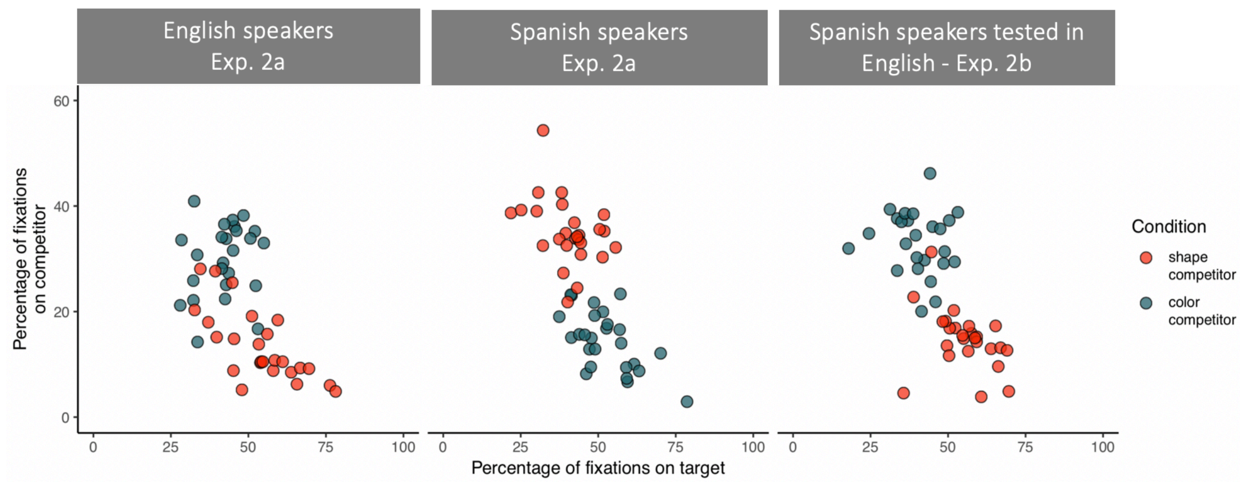

Figure 4: Percentage of fixations on the target and the competitor during the critical window. Each dot represents a participant, with two dots per participant in each plot, one corresponding to their average fixations on the Shape Competitor condition, and the other corresponding to their average fixations on the Color Competitor condition. The x axis shows the percentage of fixations on the target, and the y axis shows the percentage of fixations on the relevant competitor (shape or color, depending on the condition). Supporting the incremental efficiency hypothesis, English speakers and Spanish speakers tested in English showed higher rates of target fixations in the shape competitor trials relative to the color competitor trials. By contrast,

Spanish speakers showed higher rates of target fixations in the color competitor trials relative to the shape competitor trials.

General discussion

Here we tested an incremental account of communicative efficiency, whereby speakers of different languages exploit word order to coordinate more efficiently. Supporting the incremental efficiency hypothesis, English speakers produced more redundant color adjectives than Spanish speakers when referring to shapes in sparse displays (where prenominal color adjectives are more efficient than postnominal color adjectives), but both groups did so more frequently and to comparable rates in denser displays (where the listener's search for the referent was harder and both types of adjectives would be efficient).

Further supporting the incremental efficiency hypothesis, our eye-tracking results revealed that the visual search for a referent was guided by color and refined by shape in English, whereas the search for the same referent was guided by shape and refined by color in Spanish. Importantly, the Spanish speakers reversed their visual search strategy when retested in English, showing efficient incremental processing, rather than a native bias to contrast entities of the same kind when processing color adjectives.

It must be noted, however, that the cross-linguistic differences observed in both the production and comprehension of color adjectives do not support linguistic determinism (Whorf in Carroll, 1956). Spanish speakers did not produce more redundant color adjectives than English speakers across all visual contexts, nor did they establish color contrast within a category regardless of the language they were processing. We therefore conclude that English and

Spanish speakers did not reveal 'differently structured minds' (Pinker, 2007), but flexible effects of word order driven by efficiency pressures.

Our results also inform psycholinguistic analyses of the use and comprehension of color adjectives. Continuous eye-tracking measures revealed that Spanish speakers established color contrast within a category (TRIANGLES > BLUE ONE), whereas English speakers did so across categories (BLUE SHAPES > TRIANGULAR ONE). This difference highlights the importance of adopting an incremental perspective when studying communicative efficiency. Thus, the idea that redundant color adjectives exploit color contrast across categories challenges the general view that they are non-contrastive (Sedivy, 2003, 2004) and explains why redundant color adjectives can facilitate the listener's visual search for a referent (Sonnenschein & Whitehurst, 1982; Mangold & Pobel, 1988; Arts et al., 2011b; Paraboni et al., 2007; Paraboni & Van Deemter, 2014; Rubio-Fernandez, under review) and be efficient (Rubio-Fernandez, 2016, 2019), rather than being pragmatically infelicitous (Engelhardt et al., 2006, 2011).

Our results are also relevant to computational models of reference generation, which so far have only been developed to account for data from pre-nominal languages (e.g., Dale & Reiter, 1995; Frank & Goodman, 2012; Van Deemter et al., 2012; Van Gompel et al., 2019; Degen et al., under review). However, languages where adjectives precede nouns are a minority, with most world languages positioning their adjectives after their nouns (Dryer, 2013). Like pragmatic theories, computational models may not have looked at reference from a cross-linguistic perspective because pragmatics is supposed to apply above and across all languages and communicative situations. However, our research shows that, like other language components, pragmatics is affected by incrementality constraints and efficiency pressures, also requiring a cross-linguistic investigation (Rubio-Fernandez & Jara-Ettinger, under review).

Importantly, incrementality is not the only factor that affects the production of redundant color adjectives. The discriminability of the referent and the density of the display have been shown to affect color overspecification (Rubio-Fernandez, 2019), as well as the typicality of the color (with atypical colors being used redundantly more often than typical colors; Sedivy, 2003; Westerbeek et al., 2015; Rubio-Fernandez, 2016) and the lexical category of the noun (with clothes eliciting higher rates of redundant color adjectives than geometrical shapes; Rubio-Fernandez, 2016, 2019). In addition, not all adjectives are equally efficient in prenominal position: scalar adjectives are more context-dependent than color adjectives because of their relational semantics (Kennedy, 1999, 2007; Kennedy & McNally, 2005), resulting in higher processing costs that may not justify their overspecification (for eye-tracking evidence, see Aparicio et al., 2016; Rubio-Fernandez et al., under review). Likewise, material adjectives are not as visually salient as color adjectives, being less efficient when used redundantly (Rubio-Fernandez & Jara-Ettinger, in preparation). We therefore see overspecification as a multifaceted phenomenon determined not only by informativity considerations, but by the complex interaction of pragmatic, semantic and perceptual factors in speaker-listener coordination.

Under our account, speakers need to determine the relative salience of the target when deciding if a redundant color word will be useful or not. While much work has emphasized that speakers must rely on perspective taking when deciding what to say (e.g., Goodman & Frank, 2016), determining visual salience may be an exception. In situations of co-presence, such as the one created in Experiment 1, speakers are entitled to assume that anything that seems visually salient to them, should in principle be salient to the listener (and likewise, if the speaker's visual search is harder in denser displays, so will it be for the listener). It is thus possible that speakers assume mutual salience and produce referential expressions that are efficient for the listener, without actively adopting their perspective (for discussion, see Clark et al., 1983; Keysar, 1997; Arnold, 2008; Brennan et al., 2010; Barr, 2014; Rubio-Fernandez, 2019). Indeed, related work has shown that our inferences about what other people see are affected by biases in our own visual system, suggesting some overlap between our visual perception and our reasoning about others' visual perception (Bio et al., 2018). Note, however, that mutual salience is not necessarily the same as being egocentric (or insensitive to the listener's perspective), as it is possible that speakers are flexible and only adopt mutual salience when their perspective is matched with the listener's in situations of co-presence (Clark & Marshall, 1981). In addition, a purely egocentric view where speakers choose whether to use color on the basis of visual salience alone would fail to account for the cross-linguistic differences that we report here.

To conclude, our results confirm that speaker-listener coordination is subject to efficiency pressures that affect the production and processing of redundant color adjectives, supporting an incremental view of communicative efficiency. Moreover, these findings highlight the importance of cross-linguistic research for developing nuanced pragmatic theories and computational models that explain how language users maximize communicative efficiency given the constraints and affordances of their languages.

References

Aparicio, H., Xiang, M., & Kennedy, C. (2016), Processing gradable adjectives in context: A visual world study. Semantics and Linguistic Theory, 25, 413-432.

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

Arts, A., Maes, A., Noordman, L., & Jansen, C. (2011a). Overspecification in written instruction. Linguistics, 49, 555–574.

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

Aylett, M., & Turk, A. (2004). The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. Language and Speech, 47, 31-56.

Barr, D. J. (2014). Perspective taking and its impostors in language use: Four patterns of deception. The Oxford Handbook of Language and Social Psychology (pp. 98-110).

Belke, E. (2006). Visual determinants of preferred adjective order. Visual Cognition, 14, 261-294.

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, 237-266.

Bio, B. J., Webb, T. W., & Graziano, M. S. A. (2018). Projecting one's own spatial bias onto others during a theory-of-mind task. Proceedings of the National Academy of Sciences U. S. A. , 115, E1684-E1689.

Bramão, I., Reis, A., Petersson, K. M., & Faísca, L. (2011). The role of color information on object recognition: A review and meta-analysis. Acta Psychologica, 138, 244-253.

Brennan, S. E., Galati, A., & Kuhlen, A. K. (2010). Two minds, one dialog: Coordinating speaking and understanding. Psychology of Learning and Motivation, 53, 301-344.

Brown-Schmidt, S., & Konopka, A. E. (2008). Little houses and casas pequeñas: Message formulation and syntactic form in unscripted speech with speakers of English and Spanish. Cognition, 109, 274-280.

Carroll, B. (Ed.) (1956), Language, Thought, and Reality: Selected Writings of Benjamin Lee Whorf. MIT Press.

Clark, H. H., & Marshall, C. R. (1981). Definite reference and mutual knowledge. In A. K. Joshi, B. Webber and I. A. Sag (Eds.), Elements of discourse understanding. Cambridge: Cambridge University Press (pp. 10-63).

Clark, H. H., & Schaefer, E. F. (1989). Contributing to discourse. Cognitive Science, 13, 259-294.

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

Clarke, A. D., Elsner, M., & Rohde, H. (2013). Where's Wally: The influence of visual salience on referring expression generation. Frontiers in Psychology, 4.

Degen, J., Hawkins, R. X. D., Graf, C., Kreiss, E., & Goodman, N. (under review, 2019). When redundancy is rational: A Bayesian approach to ‘overinformative’ referring expressions. arXiv:1903.08237

Dale, R., & Reiter, E. (1995). Computational interpretations of the Gricean maxims in the generation of referring expressions. Cognitive Science, 18, 233–263.

Dryer, S. (2013). Order of Adjective and Noun. In M. S. Dryer and M. Haspelmath (Eds.), The World Atlas of Language Structures. Max Planck Institute for Evolutionary Anthropology at Leipzig (available online at http://wals.info/chapter/87).

Eberhard, K. M., Spivey-Knowlton, M. J., Sedivy, J., & Tanenhaus, M. K. (1995). Eye movements as a window into real-time spoken language comprehension in natural contexts. Journal of Psycholinguistic Research, 24, 409-436.

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

Engelhardt, P. E., Demiral, Ş. B., & Ferreira, F. (2011). Over-specified referring expressions impair comprehension: An ERP study. Brain and Cognition, 77, 304-314.

Engelhardt, P. E., & Ferreira, F. (2014). Do speakers articulate over-described modifiers differently from modifiers that are required by context? Implications for models of reference production. Language and Cognitive Processes, 29, 975-985.

Frank, M.C., & Goodman, N.D. (2012). Predicting pragmatic reasoning in language games. Science, 336, 998.

Gatt, A., Krahmer, E., Van Deemter, K., & van Gompel, R. P. (2017). Reference production as search: The impact of domain size on the production of distinguishing descriptions. Cognitive Science, 41, 1457-1492.

Geurts, B., & Rubio-Fernandez, P. (2015). Pragmatics and processing. Ratio, 28, 446-469.

Gibson, E., Futrell, R., Piantadosi, S. T., Dautriche, I., Mahowald, K., Bergen, L., Levy, R. P. (in press, 2019). How efficiency shapes human language. Trends in Cognitive Science. doi:10.31234/osf.io/w5m38.

Gibson, E., Piantadosi, S. T., Brink, K., Bergen, L., Lim, E., & Saxe, R. (2013). A noisy-channel account of crosslinguistic word-order variation. Psychological Science, 24, 1079-1088.

Givón, T. (1979), On understanding grammar. New York: Academic Press.

Goodman, N. D., & Frank, M. C. (2016). Pragmatic language interpretation as probabilistic inference. Trends in Cognitive Sciences, 20, 818-829.

Grice, H. P. (1975). Logic and conversation. In P. Cole and J. Morgan (Eds.), Speech acts. Academic Press (pp. 41-58).

Hawkins, R. X. D., Gweon, H., & Goodman, N. D. (under review, 2019). Speakers account for asymmetries in visual perspective so listeners don't have to. arXiv:1807.09000v2

Jaeger, T. F. (2010). Redundancy and reduction: Speakers manage syntactic information density. Cognitive Psychology, 61, 23-62.

Kemp, C., & Regier, T. (2012). Kinship categories across languages reflect general communicative principles. Science, 336, 1049-1054.

Kemp, C., Xu, Y., & Regier, T. (2018). Semantic typology and efficient communication. Annual Review of Linguistics, 4, 109-128.

Kennedy, C. (2007). Vagueness and grammar: The semantics of relative and absolute gradable adjectives. Linguistics and Philosophy, 30, 1-45.

Kennedy, C., & McNally, L. (2005). Scale structure and the semantic typology of gradable predicates. Language, 81, 345–381.

Keysar, B. (1997). Unconfounding common ground. Discourse Processes, 24, 253-270.

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

Koolen, R., Krahmer, E., & Swerts, M. (2015). How distractor objects trigger referential overspecification: Testing the effects of visual clutter and distractor distance. Cognitive Science, 40, 1617-1647.

Long, M. R., Horton, W. S., Rohde, H., & Sorace, A. (2018). Individual differences in switching and inhibition predict perspective-taking across the lifespan. Cognition, 170, 25-30.

Maes, A., Arts, A., & Noordman, L. (2004). Reference management in instructive discourse. Discourse Processes, 37, 117-144.

Mahowald, K., Fedorenko, E., Piantadosi, S. T., &. Gibson, E. (2013). Info/information theory: Speakers choose shorter words in predictive contexts. Cognition, 126, 313-318.

Mangold, R., & Pobel, R. (1988), Informativeness and instrumentality in referential communication. Journal of Language and Social Psychology, 7, 181-191.

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

Paraboni, I., Van Deemter, K., & Masthoff, J. (2007). Generating referring expressions: Making referents easy to identify. Computational Linguistics, 33, 229-254.

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

Piantadosi, S. T., Tily, H., & Gibson, E. (2011). Word lengths are optimized for efficient communication. Proceedings of the National Academy of Sciences of the United States of America, 108, 3526-3529.

Pinker, S. (2007). The Stuff of Thought: Language as a Window into Human Nature. Viking.

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

Rubio-Fernández, P. (under review, 2019). Redundant color adjectives are more efficient than shorter descriptions.

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

Rubio-Fernandez, P., Aparicio, H., Shukla, V., & Jara-Ettinger, J. (under review). Contrastive inferences are sensitive to informativity expectations, visual salience and adjective semantics.

Rubio-Fernandez, P. & Jara-Ettinger, J. (under review). Incrementality and efficiency shape pragmatics across languages: An eye-tracking investigation.

Sedivy, J. (2003). Pragmatic versus form-based accounts of referential contrast: Evidence for effects of informativity expectations. Journal of Psycholinguistic Research, 32, 3-23.

Sedivy, J. (2004). Evaluating explanations for referential context effects: Evidence for Gricean mechanisms in online language interpretation. In J. Trueswell and M. K. Tanenhaus (Eds.), Approaches to studying world-situated language use. MIT Press (pp. 345-364).

Sonnenschein, S., & Whitehurst, G. J. (1982), The effects of redundant communications on the behavior of listeners: Does a picture need a thousand words? Journal of Psycholinguistic Research, 11, 115-125.

Spivey, M. J., Tyler, M. J., Eberhard, K. M., & Tanenhaus, M. K. (2001). Linguistically mediated visual search. Psychological Science, 12, 282-286.

Van Deemter, K., Gatt, A., Van der Sluis, I., & Power, R. (2012). Generation of referring expressions: Assessing the incremental algorithm. Cognitive Science, 36, 799-836.

Van der Sluis, I., & Krahmer, E. (2007). Generating multimodal references. Discourse Processes, 44, 145-174.

Van Gompel, R., van Deemter, K., Gatt, A., Snoeren, R., & Krahmer, E. (2019). Conceptualisation in reference production: Probabilistic modelling and experimental testing. Psychological Review, 126, 345-373.

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

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. Special Issue Models of Reference.

Zipf, G. K. (1949). Human behavior and the principle of least effort. Oxford, England: Addison-Wesley Press.

Acknowledgments

The authors thank Helena Aparicio, Bart Geurts, Ted Gibson and Simon Kirby for helpful discussions of the results of this study. This research was supported by a Young Research Talent

Grant (Research Council of Norway, 230718) awarded to PRF and a Google Faculty Research Award to JJE.